第28卷 第1期
2006年3月
西北地震学报
NORTHWESTERN SEISMOLOGICAL JOURNAL
Vol. 28 No. 1
March, 2006

综述

支持向量机及其在地震预报中的应用前景

王 炜，林命週，马钦忠，赵利飞
（上海市地震局，上海 200062）

摘 要：统计学习理论（SLT）是研究小样本情况下机器学习规律的理论。支持向量机（SVM）基于统计学习理论，可以处理高度非线性分类和回归等问题，不但较好地解决了小样本、过学习、高维数、局部最小等实际难题，而且具有很强的泛化（预测）能力。本文介绍了支持向量机的分类、回归方法，分析了这一方法的特点，讨论了该方法在地震预报中的应用前景。

关键词：统计学习理论；支持向量机；分类；回归；地震预报
中图分类号：P315.75 文献标识码：A 文章编号：1000-0844(2006)01-0078-07

Support Vector Machines and Its Application Future in Earthquake Predication

WANG Wei, LIN Ming-zhou, MA Qin-zhong, ZHAO Li-fei
(*Earthquake Administration of Shanghai, Shanghai 200062, China*)

Abstract: Statistical learning theory (SLT) is a small-sample statistics theory. Support vector machine (SVM) is a new machine learning method based on statistical learning theory. It can process the high nonlinear problems with classification and regression. SVM not only can solve some problems, such as small-sample, over-fitting, high-dimension and local minimum, but also has higher generalization (forecasting) ability than that of the artificial neural networks. In this paper, the classification and regression methods of SVM are introduced, the characters of the methods are analyzed, and the application future of SVM in earthquake prediction is discussed also.

Key word: **Statistical learning theory; Support vector machine; Classification; Regression; Earthquake prediction**

0 引言

机器学习是现代智能技术中的重要方面。统计学习理论[1]（Statistical Learning Theory，简称SLT）是一种专门研究小样本情况下机器学习规律的理论。V. Vapnik等人[1]从上世纪六、七十年代开始致力于此方面研究；到九十年代中期，随着其理论的不断发展和成熟，也由于神经网络等学习方法在理论上缺乏实质性进展，统计学习理论开始受到越来越广泛的重视。统计学习理论建立在一套较坚实的理论基础之上，为解决小样本学习问题提供了一个统一框架，能解决许多原来难以解决的问题；同时，在该理论基础上发展了一种新的通用学习方法——支持向量机（Support Vector Machine，简称SVM），它已表现出很多优于已有方法的性能。不少学者认为，SLT和SVM正在成为继神经网络研究之后新的研究热点，并将有力地推动机器学习理论和技术的发展。

收稿日期：2005-08-29
基金项目：地震科学联合基金（104090）
作者简介：王 炜（1947—），男（汉族），江苏南京人，研究员，主要从事地震预报及其研究工作.

SVM方法的基本思想是:基于Mercer核展开定理[2],可以通过非线性映射φ,把样本空间映射到一个高维乃至于无穷维的特征空间(Hilbert空间),使在特征空间中可以应用线性学习机的方法解决样本空间中的高度非线性分类和回归等问题。降维是人们处理复杂问题时常用的简化方法之一,这样做可以降低计算的复杂性。而升维,即向高维空间做映射,一般只会增加计算的复杂性,甚至引起"维数灾",因而人们很少问津。但是作为分类、回归等问题来说,很可能在低维样本空间无法线性处理的样本集在高维特征空间却可以通过一个线性超平面实现线性划分(或回归),而与特征空间的线性划分(或回归)相对应的却是样本空间的非线性分类(或回归)。这样就自然发生的两个问题,即如何求得非线性映射φ和解决算法的复杂性。SVM方法巧妙地解决了这两个难题:由于应用了核函数展开定理[2],所以根本不需要知道非线性映射的显式表达式;由于是在高维特征空间中应用线性学习机的方法,所以与线性模型相比几乎不增加计算的复杂性,这在某种程度上避免了"维数灾"。

支持向量机是近年来倍受瞩目的一种处理高度非线性分类、回归等问题的新方法。研究发现支持向量机的各项性能、尤其是泛化(预测)能力优于传统的统计方法和人工神经网络方法,因此近几年来已在文字识别、人脸识别、医疗诊断、故障诊断、基因分类、遥感图像分析、函数逼近、勘察测井、气象、化学计量[3]等领域得到了一些十分有意义的应用。一些学者还研究了支持向量机在混沌时间序列预测方面的应用[4-6]。各方面的应用结果表明支持向量机比传统的分类和预测方法具有更佳的性能。本文介绍了支持向量机的分类、回归方法,讨论了支持向量机方法在地震预报中的可能应用前景。

2　支持向量机分类方法[2,3,7-10]

2.1　线性支持向量机

2.1.1　线性可分情况

支持向量机是从线性可分情况下的最优分类面发展而来的,其基本思想可用图1的两维情况说明。图中实心点和空心点代表两类样本;H为分类超平面;H_1、H_2分别为过各类中离分类超平面最近的样本且平行于分类超平面的平面,它们之间的距离称为分类间隔(margin)。所谓最优分类面就是要求分类面不但能将两类正确分开(训练错误率为0),而且使分类间隔最大。

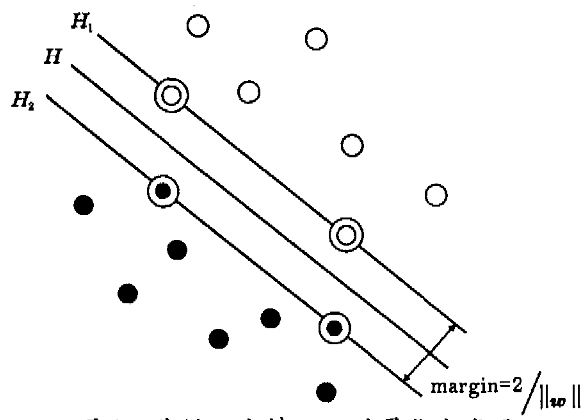

图1　线性可分情况下的最优分类面

Fig. 1　Optimization classification plane in linear discriminability.

最大间隔和最优超平面只由落在边界上的样本点完全确定,我们称这样的样本点为支持向量。此时只需由少数训练样本点(支持向量)就能把最大间隔和最优超平面完全确定,其余非支持向量的样本点均不起作用。这具有重要的意义,它说明间隔最大化原则下的最优划分不是依赖于所有点,而只是由支持向量决定。求最优超平面和最大间隔等同于确定各个样本点是否为支持向量,这预示着该方法具有较好的鲁棒性(robust)。由支持向量确定的线性分类机称为线性支持向量机。

设样本为n维向量;样本集为(x_i, y_i), $i=1,2,\cdots,k$, $x_i \in R^n$(R为实数域);$y_i \in \{-1,+1\}$。这里先讨论线形可分的二类划分的SVM分类器。此时n维空间中线性判别函数的一般形式为$f(x) = w \cdot x + b$,其分类超平面H的方程为

$$w \cdot x + b = 0 \tag{1}$$

这里(·)表示内积;w、b为待定参数;w为超平面H的法向量。将判别函数归一化,使两类样本都满足$|f(x)| \geqslant 1$,即使离分类面最近样本的$|f(x)| \geqslant 1$。则对所有样本x_i应满足

$$w \cdot x_i + b \geqslant 1 \qquad 若\ y_i = 1$$

$$w \cdot x_i + b \leqslant -1 \qquad 若\ y_i = -1$$

可将上述不等式的规范形式合并为如下紧凑型式:

$$y_i(w \cdot x_i + b) \geqslant 1 \qquad i = 1,2,\cdots,k \tag{2}$$

可以证明此时分类间隔margin $=2/\|w\|$,使间隔最大等价于使$\|w\|^2$最小。满足条件(2)且使$\|w\|^2/2$最小的分类面称为最优分类面,H_1、H_2上的训练样本点就称作支持向量。因此构造最优分类超平面的问题可转化为在满足式(2)约束条件下对(3)式的最小化的问题:

$$\Phi(w,b) = \frac{1}{2} w \cdot w \tag{3}$$

在线性条件下(3)式的优化问题是典型的二次规划问题,其最优解为拉格朗日(Lagrange)函数的鞍点:

$$L(w,b,\alpha)=\frac{1}{2}\|w\|^2-\sum_{i=1}^{k}\alpha_i[y_i(w\cdot x_i+b)-1] \tag{4}$$

式中 α_i 为非负 Lagrange 乘子。在鞍点处由于 w 和 b 的梯度均为零,则可得到

$$\sum_{i=1}^{k}\alpha_i y_i=0,\quad w=\sum_{i=0}^{k}\alpha_i y_i x_i \tag{5}$$

因为只有支持向量可以在 w 的展开式中具有非零系数 α_i,这时的支持向量就是使式(2)成立的向量,即只有支持向量影响最终的分类结果。根据优化理论的 KKT(Karush – Kuhn – Tucher)条件,最后可以得到最优超平面方程为

$$\sum_{i=1}^{k}\alpha_i y_i x_i\cdot x+b=0 \tag{6}$$

由此可得最优超平面的分类函数为(7)式,这里 sgn 为符号函数:

$$f(x)=\mathrm{sgn}(\sum_{i=1}^{k}\alpha_i y_i x_i\cdot x+b) \tag{7}$$

2.1.2 线性不可分情况

显然,以上介绍的方法是在样本能全部正确分类的情况下得到的。考虑到可能存在一些样本不能全部被超平面正确分类,因此引入松驰变量 ξ_i 来允许错分样本的存在:

$$\xi_i\geqslant 0,\qquad i=1,2,\cdots,k \tag{8}$$

显然,当分类出现错误时 ξ_i 大于零,$\sum_{i=1}^{k}\xi_i$ 是分类错误数量的一个上界。这时约束条件(2)式变为

$$y_i(w\cdot x_i+b)\geqslant 1-\xi_i,\quad i=1,2,\cdots,k \tag{9}$$

为此引入错误惩罚分量 C,因此构造广义最优分类超平面问题就转化为在(8)、(9)式约束条件下的最小化函数

$$\Phi(w,b)=\frac{1}{2}w\cdot w+C\sum_{i=1}^{k}\xi_i \tag{10}$$

的问题。上式中 C 为一正常数,C 越大,对错误的惩罚越重。其中第 1 项为样本到超平面的距离,应尽量大,从而提高泛化能力;第 2 项使误差尽量小。即折衷考虑最少错分样本和和最大分类间隔,得到广义最优分类面。

式(10)的优化问题同样是典型的二次规划问题,其最优解为下面拉格朗日(Lagrange)函数的鞍点:

$$L(w,b,\alpha,\xi,\gamma)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^{k}\xi_i-\sum_{i=1}^{k}\alpha_i[y_i(w\cdot x_i+b)-1+\xi_i]-\sum_{i=1}^{k}\gamma_i\xi_i \tag{11}$$

式中 α_i,γ_i 为非负 Lagrange 乘子。由于 w、b、ξ 的梯度均为零,则可得到

$$\sum_{i=1}^{k}\alpha_i y_i=0,\qquad w=\sum_{i=0}^{k}\alpha_i y_i x_i,$$
$$C-\alpha_i-\gamma_i=0 \tag{12}$$

将式(12)代入式(11),可得优化问题的对偶形式,即在式(14)约束条件下,对 α_i 求解式(13)的最大值

$$Q(\alpha)=\sum_{i=1}^{k}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{k}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j) \tag{13}$$

其约束条件为

$$\sum_{i=1}^{k}\alpha_i y_i=0,\qquad 0\leq\alpha_i\leq C \tag{14}$$

这是一个二次函数寻优的问题,存在唯一解。容易证明解中只有一部分(通常是少部分)的 $\alpha_i>0$,所对应的样本为支持向量。求解上述问题的最优分类函数与式(7)相同。

2.2 非线性支持向量机

SVM 方法真正有价值的应用是用来解决非线性问题,方法是通过一个非线性映射 φ,把样本空间映射到一个高维乃至于无穷维的特征空间,使在特征空间中可以应用线性支持向量机的方法解决样本空间中的高度非线性分类和回归等问题。图 2 对此给出二维样本数据的直观示意图。

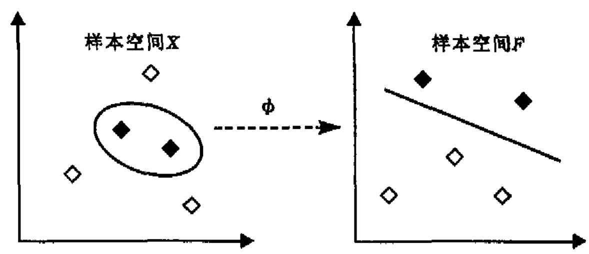

图 2 样本空间到特征空间的非线性映射示意

Fig.2 Sketch of nonlinear mapping from sample space to characteristic space.

在特征空间中应用线性支持向量机的方法将 $\varphi(x)$ 和 $\varphi(x_i)$ 代替了 x 和 x_i,分类决策函数式(7)变为

$$f(x)=\mathrm{sgn}(\sum_{i=1}^{k}\alpha_i y_i(\varphi(x_i)\cdot\varphi(x))+b) \tag{15}$$

直接确定非线性映射 φ 的形式是较困难的,且计算量随特征空间维数增加呈指数递增。根据 Hilbert-Schmidt 原理[2],处理高维特征空间的计算

问题可以避开求解空间映射 φ 的显式形式，即通过引入所谓核函核 $K(x_i, x)=(\varphi(x_i)\cdot\varphi(x))$，将变换空间中的内积转化为原空间中某个函数的计算，从而间接求解输入空间向高维特征空间的映射 φ。分类决策函数为

$$f(x)=\operatorname{sgn}\left(\sum_{i=1}^{k}\alpha_i y_i K(x_i,x)+b\right) \quad (16)$$

任意满足 Mercer 核[2]条件的对称函数均可作为核函数。常用的核函数有：①多项式核函数 $K(x_i,x_j)=(x_i\cdot x_j+1)^d, d=1,2,\cdots$，所得到的是 d 阶多项式；②径向基核函数(RBF) $K(x_i,x_j)=\exp\{-\|x_i-x_j\|^2/2\sigma^2\}$；③Sigmoid 核函数 $K(x_i,x_j)=\tanh[b(x_i\cdot x_j)+c]$ 等。在核函数中，RBF 函数因其优秀的局部逼近特性在 SVM 中应用最为广泛。

式(16)就是非线性支持向量学习机的最终分类决策函数。用支持向量机求得的分类决策函数在形式上类似于一个神经网络(图3)，其输出为若干层中间节点的线性组合。每个中间节点对应一个支持向量，因此也被称为支持向量网络。通过核函数的计算即可得到原来样本空间的非线性划分输出值。这样我们就通过核函数和线性 SVM 方法解决了非线性 SVM 问题。而线性 SVM 的算法归结为一个二次凸规划问题，对此已有很多成熟的算法和应用软件可供使用。

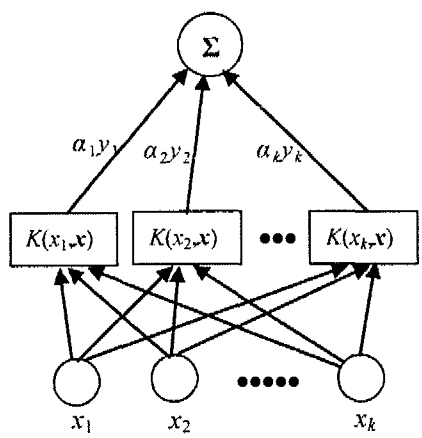

图3 支持向量机示意图

Fig. 3 Sketch of support vector machine.

3 支持向量机回归方法[2-6]

回归分析要解决的问题是：设样本为 n 维向量，样本集为 $(x_i,y_i), i=1,2,\cdots,k, x_i\in R^n$($R$ 为实数域)；$y_i\in R$。其中 x_i 为预报因子值，y_i 为预报对象值，寻求一个反映样本数据的最优函数关系 $y=f(x)$。如果所得函数关系 $y=f(x)$ 是线性函数，则为线性回归，否则为非线性回归(图4)。与 SVM 分类问题不同的是：SVM 回归的样本点只有一类，所寻求的最优超平面不是使两类样本点分得"最开"，而是使所有样本点离超平面的"总偏差"最小。这时样本点都在两条边界线之间，求最优回归超平面同样等价于求最大间隔，推导过程与 SVM 分类情况相同。

用 SVM 进行回归，其基本思想同样是通过一个非线形映射 $\varphi: R^N\to R^M (M>N)$，将输入空间映射到高维特征空间，在高维特征空间中进行线性回归。从几何意义上讲，就是寻找一个最优拟合样本集的超平面。在线性情况下其超平面函数为 $f(x)=w\cdot x+b$。当映射到高维特征空间时，拟合样本集为 $\{\varphi(x_i), y_i\}, (i=1,2,\cdots,k)$。这样以向量形式的回归函数表达为

$$f(x,w)=(w\cdot\varphi(x))+b \quad (17)$$

w，$\varphi(x)$ 为 n 维向量；b 为阈值；$(\cdot)$ 表示内积。定义不灵敏函数 ε 作为误差函数(图4)，当所有样本点到所求超平面的距离都可以不大于 ε 时(这相当于 SVM 分类时的线性可分的情况)，寻求最优回归超平面的问题转化为求解如下一个二次凸规划问题：

$$\min_{w,b}\Phi(w,b)=\frac{1}{2}w\cdot w \quad (18)$$

约束条件

$$\begin{cases} y_i-w\cdot\varphi(x_i)-b\leqslant\varepsilon, & i=1,2,\cdots,k \\ -y_i+w\cdot\varphi(x_i)+b\leqslant\varepsilon, & i=1,2,\cdots,k \end{cases}$$

同样，当个别样本点到所求超平面的距离大于 ε 时(这相当于 SVM 分类时的线性不可分的情况)，考虑到允许拟合误差的情况，引入正数松弛变量 ξ 和 ξ^*，寻求最优回归超平面的二次凸规划问题转化为带约束条件的优化问题：

$$\min_{w,b,\xi,\xi^*}\Phi(w,b,\xi,\xi^*)=\frac{1}{2}w\cdot w+C\sum_{i=1}^{k}(\xi_i+\xi_i^*) \quad (19)$$

约束条件

$$\begin{cases} y_i-w\cdot\varphi(x_i)-b\leqslant\varepsilon+\xi_i^*, & i=1,2,\cdots,k \\ -y_i+w\cdot\varphi(x_i)+b\leqslant\varepsilon+\xi_i, & i=1,2,\cdots,k \\ \xi_i\geqslant 0, \quad \xi_i^*\geqslant 0 \end{cases}$$

$w\cdot w$ 代表模型的复杂程度，其作用将使函数更为平坦，从而提高泛化能力。第二项则为减小误差，C 为惩罚因子，用于调整对超出拟合误差 $\varepsilon(\varepsilon>0)$ 的惩罚程度。传统 SVM 求解过程、即函数(19)的优

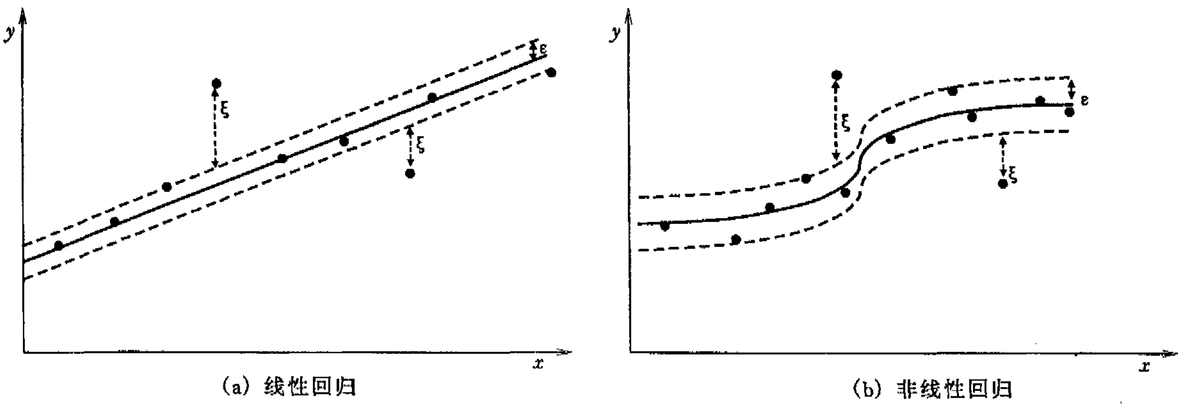

图4 线性回归、非线性回归及其不灵敏函数 ε、松弛变量 ξ

Fig. 4 Linear regression, nonlinear regression, non - sensitive function ε and laxity variable ξ.

化问题是典型的二次规划问题。引入拉格朗日(Lagrange)函数 $L(w,b,\xi,\xi^*,\alpha,\alpha^*,\gamma,\gamma^*)$。分别用 w,b,ξ,ξ^* 对函数 L 求最小化,构造式(17)的对偶形式,并解该凸函数的鞍点。可得非线性回归函数为

$$f(x) = \sum_{i=1}^{k}(\alpha_i - \alpha^*)(\varphi(x_i)\cdot\varphi(x)) + b \tag{20}$$

这里 α 和 α^* 为 Lagrange 乘子。根据核函核定理 $K(x_i, x) = (\varphi(x_i) \cdot \varphi(x))$,上式变为

$$f(x) = \sum_{i=1}^{k}(\alpha_i - \alpha^*)K(x_i,x) + b \tag{21}$$

构造形如式(21)映射函数的学习机器被称作支持向量机,它将构造输入空间的非线性映射函数转化为构造高维特征空间的线性映射函数,而且通过把原问题转化为对偶问题,使得计算的复杂度不再取决于空间维数,而是取决于样本数,特别是支持向量的个数。通常,式(21)中系数($\alpha - \alpha^*$)只有一小部分数目是非零值,它们所对应的数据点就是支持向量。

支持向量机回归分析中,标准 SVM 的结构形式上仍类似于一个神经网络(图3)。这时的隐层节点是支持向量,网络权重变为($\alpha - \alpha^*$),输出是隐节点的线性组合,b 相当于网络的阈值。这些量均可由算法自动产生,无需像神经网络构架网络过程中步步经验试算。

4 支持向量机在地震预报中的应用及其前景

在地震预报中各类地震预测数据往往样本不多、同时具有很强的非线性特征,各类异常与地震之间也具有很强的非线性关系,因此支持向量机在地震预报中将会得到很好的应用。

4.1 中国大陆强震预测的支持向量机方法①

许多学者的研究表明,我国大陆强震与全球主要板块边界的强震活动之间具有一定的关系,但是这种关系具有较强的非线性。

文献[11]曾根据全球的强震活动与板块边界的分布将全球分为16个强震活动区,使用BP神经网络研究了这16个区域的7级以上强震活动频次与次年中国大陆地震活动的关系。将这16个区域在一年中的 $M_S \geq 7.0$ 地震次数 N_i 作为神经网络的输入元,而将次年中国大陆是否发生7级以上强震作为输出,得到较好的预测效果。

为比较支持向量机方法与BP神经网络的预测效果,文章①使用与文章[11]相同的资料、方法以及相同的学习样本和检验样本应用支持向量机进行预测。该文选取1925－2003年的资料,使用支持向量机对我国大陆1926－2004年是否发生7级以上强震进行学习和预测。通过对上述65个样本的学习,对学习样本的内符检验全部正确;对14个待检验样本有12个样本的检验结果正确,两个报错(1969、2003年),报准率为12/14＝0.86。预测检验结果表明支持向量机要优于BP神经网络。

4.2 中国大陆强震时间序列预测的支持向量机方法②

使用支持向量机的回归方法可以进行时间序列的预测。文献[12]选择两种不同的方法使用BP神经网络对中国大陆最大地震时间序列进行预测。为比较两种方法的预测效果,文章②使用与文献[12]

① 王炜,刘悦,李国正,等.我国大陆强震预测的支持向量机方法[J].地震学报,(待发表).

② 王炜,刘悦,李国正,等.中国大陆强震时间序列预测的支持向量机方法[J].地震,(待发表).

相同的资料和方法采用支持向量机进行预测,结果表明该方法预报效果优于BP神经网络。

文献[12]中的方法(1)依次取前13年中我国大陆每年的最大地震震级 $M_{i+1},M_{i+2},\cdots,M_{i+13}$,预测第14年的我国大陆最大地震震级 M_{i+14},($i=1,2,\cdots,m$)。这样支持向量机的输入项为前13年中每年的最大地震;输出项为1个,即第14年的我国大陆最大地震震级 M_{i+14}。文章②使用同样的方法选取1900－2004年的中国大陆最大地震震级时间序列通过支持向量机进行学习与检验,共选取1900－1989年的地震资料78个学习样本,选取1990－2003年资料14个待检验样本,预测1991－2004年的最大地震震级。使用支持向量机对学习样本的内符检验效果很好,与实测值之差全部小于0.2级;而对14个外推检验样本的检验结果表明,如果取预测与实际最大地震震级之差小于等于0.5为报准,则有12个样本的检验结果正确,2个报错,报准率为12/14＝0.86。

考虑到我国大陆强震活动还受到一些外界因素的影响,如全球强震活动、太阳黑子等。因此在文献[12]中的方法(2)中,除使用与前述相同的时间序列模型,即输入项为前13年中每年的最大地震震级外,还增加8个输入项,分别是全球地震活动第12和13年的年频次,第13年的年释放能量,第12和13年与前一年的释放能量的差分值,第13年的太阳年平均黑子数,第12和13年与前一年的太阳黑子数差分值。这样支持向量机的输入项共为21个;输出项为1个,仍为所预测的我国大陆第14年的最大地震震级。文章②使用同样的方法选取1900－2004年的中国大陆最大地震震级时间序列进行支持向量机的学习与检验,对14个外推检验样本的检验结果表明,如果取预测与实际最大地震震级之差小于等于0.5为报准,则有13个样本的检验结果正确,1个报错。

这样报准率为13/14＝0.93。这表明我国大陆强震活动除了与强震时间序列本身有关外,还与全球的强震活动、太阳黑子活动等有密切的关系。

与文献[12]中的神经网络预测方法比较,文章②所用的支持向量机方法无论在报准率、预测的平均误差、均方差这三个方面,都要优于神经网络方法。

4.3　支持向量机方法在地震预测中的应用前景

除以上介绍了的一些工作成果外,支持向量机方法在地震预测中还有很多可应用的领域。

(1)地震前兆异常识别与干扰排除。我们可以使用支持向量机的分类方法对地震前兆异常和正常数据进行分类,从而识别和提取地震前兆异常,甚至还可以寻找新的地震前兆等,也可以对前兆干扰因素进行排除。可以使用支持向量机对正常的地震前兆数据进行回归,或者建立起前兆数据与各类影响因子之间关系的模型,从而识别前兆数据的异常或干扰;还可以对前兆时间序列、尤其是具有混沌特性的前兆时间序列进行预测等。

(2)地震学预报方法中的应用。通过支持向量机的分类方法可以对地震序列和地震活动图像进行深入研究,发现异常或正常地震序列及其异常或正常地震活动图像,也可以对异常或正常地震活动参数进行识别等。我们还可以使用支持向量机回归方法对某一地区的地震时间序列进行预测。

(3)地震综合预报方法中的应用。在综合预报中可以将各类异常作为支持向量机输入项,而将所期望的预测结果作为输出项,比如建立起某一地区的各类主要预报指标与未来地震三要素的预测模型。

(4)地震活动大形势预报方法中的应用。在大形势的分析中主要是根据一些与地震活动大形势有关的相关因素进行预报,这些相关因子与实际结论之间往往有着很明显的非线性关系。我们可以将各类相关因子作为支持向量机输入项,而将所期望的预测结果作为输出项。使用支持向量机的回归或分类方法建立起各种影响因子预测我国或某一地区未来地震活动大形势的模型等。

(5)地震震后预测中的应用。震后预测包括当前序列是否为前震或余震序列?地震序列的类型?下一次强余震三要素的预测等内容。在震后趋势判定和预测中我们往往是使用多项指标进行预测和判断。我们可以通过支持向量机的分类方法对地震序列类型进行判定;使用支持向量机建立起前震序列或余震序列判定的模型;使用支持向量机对下一次强余震三要素进行预测等。

5　结论

本文介绍了支持向量机的分类、回归方法以及在地震预测中的一些应用,讨论了支持向量机方法在地震预报中的可能应用前景。

大量研究的结果表明,支持向量机有如下一些特点:(1)基于结构风险最小化(SRM,Structural Risk Minimization)原则,具有良好的泛化(预测)能

力;(2)巧妙地解决了算法复杂度与输入向量维数密切相关的问题;(3)应用核技术,将输入空间中的非线性问题,通过非线性函数映射到高维特征空间中,在高维空间中构造线性判别函数;(4)与传统统计学不同,支持向量机专门针对小样本情况,它的最优解基于已有样本信息,而不是样本数趋于无穷大时的最优解;(5)算法最终转化为一个凸二次规划的优化问题,保证了算法的全局最优性,解决了神经网络无法避免的可能出现局部最小的问题;(6)神经网络、遗传算法等传统方法的实现中带有的很大的经验成分,而支持向量机具有更严格的理论和数学基础。

[参考文献]

[1] Vapnik V 著.张学工译. 统计学习理论的本质[M]. 北京:清华大学出版社, 2000. 1-50.

[2] Cristianini N, Shawe-Taylor J 著.李国正,等 译.支持向量机导论[M]. 北京:电子工业出版社,2004. 82-106.

[3] Chen Nian-Yi, Lu Wen-Cong, Yang Jie, et al.. Support vector machines in Chemistry[M]. Singaore: World Scientific Publishing Company, 2004.

[4] 崔万照,朱长纯,保文信星,等.混沌时间序列的支持向量机预测[J].物理学报,2004,**53**(10):3304-3309.

[5] Mukherjee S, Osuma E, Girosi F. Nonlinear Prediction of Chaotic Time Series Using Support Vector Machines[J]. Neural Networks for Signal Processing Ⅶ,1997: 511-520.

[6] 孙德山,吴今培,侯振挺,等.基于 SVR 的混沌时间序列预测[J].计算机工程应用,2004,**54**(2):54-56.

[7] 陈永义,俞小鼎,高学浩,等.处理非线性分类和回归问题的一种新方法(Ⅰ)——支持向量机方法简介[J].应用气象学报,2004,**15**(3):345-354.

[8] 王国胜,钟义信.支持向量机的若干新进展[J].电子学报,2001,**29**(10): 1397-1400.

[9] 陈念贻,陆文聪,叶晨周,等.支持向量机及其他核函数算法在化学计量学中的应用[J].计算机与应用化学,2002,**19**(6):82-106.

[10] 祁亨年.支持向量机及其应用研究综述[J].计算机工程,2004,**30**(10): 6-10.

[11] 王炜,曹雪峰,宋先月.使用人工神经网络判断未来我国地震形势[J].地震学刊,2001,**3**(21):10-14.

[12] 王炜,谢端,宋先月,等.使用人工神经网络进行我国大陆强震时间序列预测[J].西北地震学报,2002,**24**(4):315-320.