计算机科学 2006Vol. 33№. 6

核函数的性质及其构造方法*)

王国胜[1,2]

(德州学院计算机系　德州 253023)[1]　(北京邮电大学信息工程学院　北京 100876)[2]

摘　要　支持向量机是一项机器学习技术,发展至今近10年了,已经成功地用于模式识别、回归估计以及聚类等,并由此衍生出了核方法。支持向量机由核函数与训练集完全刻画。进一步提高支持向量机性能的关键,是针对给定的问题设计恰当的核函数,这就要求对核函数本身有深刻了解。本文首先分析了核函数的一些重要性质,接着对3类核函数,即平移不变核函数、旋转不变核函数和卷积核,提出了简单实用的判别准则。在此基础上,验证和构造了很多重要核函数。

关键词　支持向量机,核函数,机器学习,核方法

Properties and Construction Methods of Kernel in Support Vector Machine

WANG Guo-Sheng[1,2]

(Department of Computer Science and Technology, Dezhou University, Dezhou 253023)[1]

(School of Information Engineering, Beijing University of Posts & Telecommunications, Beijing 100876)[2]

Abstract　Support vector machine, which has been successfully applied to pattern recognition, regression estimation, cluster and so on, is a typical instance of kernel method. It is completely characterized by kernel function and training set. The key to enhance performance of support vector machine is to choose an appropriate kernel function for the given problem; therefore deep understanding to kernel itself is needed. Firstly, this paper analyzes some important properties of kernel, and then proposes criterions for judgment of three classes of kernel function, i. e. translation invariant, rotation invariant and convolution kernels. By them, a lot of important kernel functions are constructed some of which are commonly employed in practice.

Keywords　Support vector machine, Kernel, Kernel method, Machine learning

1　引言

支持向量机自20世纪90年代初提出,迄今已经历了10年的发展。它在应用方面,如模式识别、预测以及聚类等,成功的例子屡见不鲜[1,3,4],在理论上也取得了很大进展[1~3]。支持向量机由核函数与训练集完全刻画。核函数本质上是一个内积(或点积)。核函数的思想后来发展成核方法,即通过引入核函数,把基于内积运算的线性算法非线性化。核方法为处理许多问题提供了一个统一的框架。除支持向量机外,还有基于核的主成分分析、基于核的Fisher判决等[3]。

核函数理论已有很长的历史,Mercer定理[1]可追溯到1909年;再生核(Reproduction Kernel) Hilbert理论[2]是20世纪40年代发展起来的;1975年,Poggio[5]首次用到了多项式核函数。但是,直到Boser,Guyon和Vapnik[6]将之用于支持向量机之前,它的重要性没有受到充分重视。

支持向量机创立之初,人们多关心基于核函数的算法的设计。后来认识到,提高支持向量机性能的关键之一,是设计适合给定问题的核函数。这就要求对核函数本身有深入了解。本文第2节分析了核函数的一些基本性质,第3节针对3类重要核函数,即平移不变核函数、旋转不变核函数和卷积核,提出了简单实用的判别准则。在此基础上,验证和构造了很多重要核函数,其中一些是我们经常用到的。

本文假设输入空间 $X \subset \mathbb{R}^d$(d 维 Euclid 空间),除非另有所指,不再申明。

2　核函数及基本性质

2.1　核函数与正定矩阵

定义 2.1.1　称二元函数 $k: X\times X\to\mathbb{R}$ 是核函数,如果存在某个内积空间(或 Hilbert 空间)$(H,\langle\cdot,\cdot\rangle)$,以及映射 $\Phi: X\to H$,使得

$$k(x,x')=\langle\Phi(x),\Phi(x')\rangle$$

称 H 为特征空间,Φ 为特征映射。

定义 2.1.2　称 $k: X\times X\to\mathbb{R}$ 是正定的,如果它是对称的,即 $k(x,x')=k(x',x)$,并且对任意 $m\in\mathbb{N}$(正整数集合),任意 $x_1,x_2,\cdots,x_m\in X,\alpha_1,\alpha_2,\cdots,\alpha_m\in\mathbb{R}$,都有

$$\sum_{i,j=1}^{m}\alpha_i\alpha_j k(x_i,x_j)\geq 0$$

即对任意训练数据 $x_1,x_2,\cdots,x_m\in X$,$K=(k(x_i,x_j))$ 是正定矩阵。

定理 2.1.1　$k: X\times X\to\mathbb{R}$ 是核函数当且仅当它是正定的。

证明:设 k 是核函数,则存在内积空间 $(H,\langle\cdot,\cdot\rangle)$,以及映射 $\Phi: X\to H$,使得 $k(x,x')=\langle\Phi(x),\Phi(x')\rangle$。对任意 $m\in\mathbb{N}$,任意 $x_1,x_2,\cdots,x_m\in X,\alpha_1,\alpha_2,\cdots,\alpha_m\in\mathbb{R}$,有

*)山东省教育厅科技计划项目(No. J03P52)、德州市科技计划项目(No. 042103)。王国胜　副教授,博士研究生,主要研究方向为计算智能、支持向量机、模式识别。

$$\sum_{i,j=1}^{m}\alpha_i\alpha_j k(x_i,x_j)=\sum_{i,j=1}^{m}\alpha_i\alpha_j\langle\Phi(x_i),\Phi(x_j)\rangle$$
$$=\langle\sum_{i=1}^{m}\alpha_i\Phi(x_i),\sum_{i=1}^{m}\alpha_i\Phi(x_i)\rangle\geqslant 0$$
又内积显然是对称的，从而 k 正定。

反之，设 k 正定。对任一 $x\in X$，对应（通过 k）X 上一个实函数

$$\Phi_x(\cdot)=k(x,\cdot)$$

所有这样定义的实函数的线性组合构成一个线性空间，记作 H_0。对 H_0 中任意两点，

$$f(\cdot)=\sum_{i=1}^{m}\alpha_i k(x_i,\cdot),g(\cdot)=\sum_{j=1}^{m'}\beta_j k(x'_j,\cdot)$$

其中 $m,m'\in\mathbb{N}$，$\alpha_i,\beta_j\in\mathbb{R}$，$x_i,x'_j\in X$，定义

$$\langle f,g\rangle:=\sum_{i=1}^{m}\sum_{j=1}^{m'}\alpha_i\beta_j k(x_i,x'_j)$$

可验证，这个定义不依赖于 f,g 的表示方式，并且 $\langle\cdot,\cdot\rangle$ 满足对称性、双线性性。因为 k 正定，可得 $\langle f,f\rangle\geqslant 0$，并且 $\langle f,f\rangle=0\Leftrightarrow f=0$。因此，$\langle\cdot,\cdot\rangle$ 是 H_0 上的一个内积。H_0 的完备化称为再生核 Hilbert 空间（Reproducing Kernel Hilbert Space，RKHS），记作 H。根据定义，k 满足

$$k(x,x')=\langle k(x,\cdot),k(x',\cdot)\rangle$$

定义特征映射

$$\Phi: X\to H,\Phi(x)=k(x,\cdot)$$

则 $k(x,x')=\langle\Phi(x),\Phi(x')\rangle$。证毕。

2.2 核函数的基本性质

定理 2.2.1（封闭性） 若 $k_1,k_2,\cdots$ 是核函数，则

(1) k_1+k_2 是核函数，

(2) αk_1，$\alpha\geqslant 0$ 是核函数，

(3) $k_1\cdot k_2$ 是核函数，

(4) 若 $k(x,x')=\lim\limits_{n\to\infty}k_n(x,x')$ 存在，则 $k(x,x')$ 是核函数。

定理的结论(1)、(2)、(4)可由定义直接验证，此三条结论表明核函数的全体构成一个闭凸锥。因为正定矩阵关于 Schur 乘积，即对应元素直接相乘，是封闭的[7]，根据定理 2.1.1，便得结论(3)。

定理 2.2.2 设 $k: X\times X\to\mathbb{R}$ 是核函数，$f: X\to\mathbb{R}$ 是任意函数，则

(1) $f(x)f(x')$ 是核函数，

(2) $f(x)k(x,x')f(x')$ 是核函数。

证明：先证(1)。显然 $f(x)f(x')$ 是对称的。对任意 $m\in\mathbb{N}$，任意 $x_1,x_2,\cdots,x_m\in X$，$\alpha_1,\alpha_2,\cdots,\alpha_m\in\mathbb{R}$，有

$$\sum_{i,j=1}^{m}\alpha_i\alpha_j f(x_i)f(x_j)=(\sum_{i=1}^{m}\alpha_i f(x_i))(\sum_{j=1}^{m}\alpha_j f(x_j))=(\sum_{i=1}^{m}\alpha_i f(x_i))^2\geqslant 0$$

所以，$f(x)f(x')$ 正定。由(1)及定理 2.1.2(3)可得结论(2)。证毕。

形如定理 2.1.1(1)中的核函数，称为可分离的核函数。

把一些已知的核函数作为基本模块，根据定理 2.1.2 和 2.1.3，便能构造出许多新核函数。

例 2.1 核函数的凸组合。设 k_i，$1\leqslant i\leqslant n$ 是核函数，$a_i\geqslant 0$ 且 $\sum_{i=1}^{n}a_i=1$，则 $\sum_{i=1}^{n}a_i k_i$ 是核函数，称为 k_i，$1\leqslant i\leqslant n$ 的凸组合。如果基本核函数 k_i 选取适当，那么新核函数能表现出更强的泛化能力，详见文[8,9]。

例 2.2 核函数的标准化。若核函数满足 $k(x,x)>0$，$\forall x\in X$，则

$$\tilde{k}(x,x')=\frac{k(x,x')}{\sqrt{k(x,x)}\sqrt{k(x',x')}}$$

是核函数，称为 $k(x,x')$ 的标准化。

例 2.3 核函数的零扩张。若 $S\subset X$，k 是 $S\times S$ 上的核函数，则

$$\tilde{k}(x,x')=\begin{cases}k(x,x'), & x\in S \text{ and } x'\in S\\ 0, & \text{otherwise}\end{cases}$$

是 $X\times X$ 上的核函数，称为 k 的零扩张。

例 2.4 核函数的零置换。设 k 是 $X\times X$ 上的核函数，$S\subset X$，则

$$\tilde{k}(x,x')=\begin{cases}k(x,x'), & x\in S \text{ and } x'\in S\\ 0, & \text{otherwise}\end{cases}$$

是 $X\times X$ 上的核函数，称为 k 的零置换。

证明：$\tilde{k}(x,x')=k(x,x')I_{S\times S}(x,x')=I_S(x)k(x,x')I_S(x')$，由定理 2.1.3(2)，$\tilde{k}(x,x')$ 是核函数。证毕。

2.3 核函数反映了输入数据之间的相似性

设 $k: X\times X\to\mathbb{R}$ 是核函数，Φ 是 k 的特征映射，则 k 在输入空间 X 上诱导了一个伪距离：

$$\rho_k(x,x')=\|\Phi(x)-\Phi(x')\|$$
$$=\sqrt{\langle\Phi(x),\Phi(x)\rangle-2\langle\Phi(x),\Phi(x')\rangle+\langle\Phi(x'),\Phi(x')\rangle}$$
$$=\sqrt{k(x,x)-2k(x,x')+k(x',x')}$$

这个伪距离可解释成 x 与 x' 之间的相似性度量，这一点对理解核函数和支持向量机很重要。实际上，核函数能定义在更抽象、更复杂的输入空间上，如树、图、字符串等，因此就可以对诸如此类的复杂结构定义一种相似性，从而把这些差别迥异的对象纳入统一的处理框架。另一方面，选择核函数相当于定义输入空间中元素的相似性，因此要构造好的核函数，必须把握待处理问题的具体特点，这与我们对问题领域的先验知识及其利用程度有关，是目前支持向量机研究中的主要问题之一。

2.4 核函数对应的特征空间和特征映射不是唯一的

看下面的简单例子。

设数据 $x,y\in R^2$，则 $k(x,y)=\langle x,y\rangle^2$ 是核函数。其对应的特征空间 H 和特征映射 Φ 有，

$$H=R^3,\Phi(x)=(x_1^2,\sqrt{2}x_1x_2,x_2^2);$$
$$H=R^3,\Phi(x)=\frac{1}{\sqrt{2}}((x_1^2-x_2^2),2x_1x_2,(x_1^2+x_2^2));$$
$$H=R^4,\Phi(x)=(x_1^2,x_1x_2,x_1x_2,x_2^2)。$$

这里 $x=(x_1,x_2)$。可以直接验证，它们都满足 $k(x,y)=\langle\Phi(x),\Phi(y)\rangle$。

虽然核函数对应的特征空间和特征映射不唯一，却有如下结论。

定理 2.4.1[11] 设 $k: X\times X\to\mathbb{R}$ 是核函数，$\Phi_1: X\to H_1$，$\Phi_2: X\to H_2$ 是 k 的两个特征映射，则对任一 $\omega_1\in H_1$，存在 $\omega_2\in H_2$，使得 $\|\omega_2\|\leqslant\|\omega_1\|$，且

$$\langle\omega_1,\Phi_1(x)\rangle=\langle\omega_2,\Phi_2(x)\rangle,\forall x\in X$$

此定理说明：在支持向量机中，最优分离曲面（在输入空间中）仅与核函数本身有关，与特征空间和特征映射的具体表示无关。

3 3类重要核函数

3.1 平移不变核

平移不变核是指核函数具有形式 $k(x,x')=f(x-x')$，其中 $f: X\to\mathbb{R}$ 是实函数。例如，高斯核函数

$k(x,x')=e^{-a\|x-x'\|^2},a>0$

试问,当 f 满足什么条件时,$f(x-x')$是核函数?我们有如下判别准则。

定理 3.1.1 设 $f:X\to\mathbb{R}$ 是有界可积连续函数,则 $k(x,x')=f(x-x')$为核函数的充要条件是:$f(0)>0$,且其 Fourier 变换

$\tilde{f}(\omega)=\int_X f(x)e^{-i\langle\omega,x\rangle}dx\geqslant 0$

证明:必要性。因为 $f(x-x')$是核函数,所以正定,并且可推出 $f(0)>0$。又 $f(x)$连续,根据 Boucher 定理,$f(x)/f(0)$是特征函数。由 Fourier 反演公式

$$\frac{1}{(2\pi)^d}\int_X\frac{f(x)}{f(0)}e^{-i\langle\omega,x\rangle}dx\geqslant 0$$

从而,$\tilde{f}(\omega)=\int_X f(x)e^{-i\langle\omega,x\rangle}dx\geqslant 0$。

充分性。设 $\varphi(x)$和 $\tilde{\varphi}(\omega)$分别是 d 维标准正态分布的密度函数和特征函数,设

$$u(\omega)=\frac{1}{(2\pi)^d}\int_X\frac{f(x)}{f(0)}e^{-i\langle\omega,x\rangle}dx\geqslant 0$$

两边乘以 $\tilde{\varphi}(t\omega)e^{i\langle\omega,a\rangle}$,其中 $t>0,a\in\mathbb{R}^d$是参数,

$$u(\omega)\tilde{\varphi}(t\omega)e^{i\langle\omega,a\rangle}=\frac{1}{(2\pi)^d}\int_X\tilde{\varphi}(t\omega)\frac{f(x)}{f(0)}e^{-i\langle\omega,x-a\rangle}dx$$

两边积分,并利用 Fourier 反演公式,得

$$\int_X u(\omega)\tilde{\varphi}(t\omega)e^{i\langle\omega,a\rangle}d\omega=\int_X\frac{f(x)}{f(0)}\varphi\left(\frac{x-a}{t}\right)\frac{dx}{t}\qquad(1)$$

因为对任意 $t>0$,$\varphi(\frac{x-a}{t})\frac{1}{t}$是一个概率密度函数,从而右端是$\frac{f(x)}{f(0)}$的数学期望。又 $f(x)$有界,从而对任意 $t>0$,上面的积分有界。

取 $a=0$,则$\int_X u(\omega)\tilde{\varphi}(t\omega)d\omega$ 有界,因被积函数非负,且 $t\to 0$ 时极限是 $u(\omega)$,由 Fatou 引理,$u(\omega)$可积。

因为 $t\to 0$,$\varphi(\frac{x-a}{t})\frac{1}{t}$趋于集中在 a 的单点分布,在(1)式两端令 $t\to 0$,得

$$\int_X u(\omega)e^{i\langle\omega,a\rangle}d\omega=\frac{f(a)}{f(0)}\qquad(2)$$

取 $a=0$,得

$$\int_X u(\omega)d\omega=1$$

即 $u(\omega)$是概率密度。由 2 式看出,$\frac{f(a)}{f(0)}$是其特征函数,故由 Boucher 定理,$f(x-x')/f(0)$正定。又 $f(0)>0$,所以 $f(x-x')$正定,因此 $k(x,x')=f(x-x')$是核函数。证毕。

利用定理 2.2.1,可方便地验证一些重要核函数。

例 3.1 高斯核函数

$k(x,x')=e^{-a\|x-x'\|^2},a>0$

这里 $f(x)=e^{-a\|x\|^2}$。显然,$f(x)$是$\mathbb{R}^d$上的有界可积连续函数,且 $f(0)=1$;周知,高斯型函数的 Fourier 变换还是高斯型函数,从而是非负的。由定理 2.2.1,$k(x,x')$是核函数。

例 3.2 指数径向基核

$k(x,x')=e^{-a\|x-x'\|},a>0$

这里 $f(x)=e^{-a\|x\|}$。若不计一个正的常数因子(此因子与维数 d 有关),$f(x)$的 Fourier 变换具有形式$(a^2+\|\omega\|^2)^{-\frac{d+1}{2}}$,因而是正的。由定理 2.2.1,$k(x,x')$是核函数。

3.2 旋转不变核

旋转不变核是指核函数具有形式 $k(x,x')=f(\langle x,x'\rangle)$,其中 $f:D\to\mathbb{R}$ 是一元实函数($D\subset\mathbb{R}$)。例如,齐次多项式核函数 $k(x,x')=\langle x,x'\rangle^p,p\in\mathbb{N}$。

试问,当 f 满足什么条件时,$f(\langle x,x'\rangle)$是核函数?我们有如下判别准则。

定理 3.2.1 设 $f(t)$在$-r<t<r$ 上有定义($0<r\leqslant\infty$),若它的各阶导数都存在,且

$f^{(n)}(t)\geqslant 0,0<t<r$

则 $k(x,x')=f(\langle x,x'\rangle)$是核函数。

证明:满足定理条件的函数 $f(t)$,称为 $0<t<r$ 上的绝对单调函数。$f(x)$是 $0<t<r$ 上绝对单调函数的充要条件是[13]:$f(t)$具有非负系数的幂级数展开,即

$f(t)=a_0+a_1t+a_2t^2+\cdots\cdots$

其中 $a_i\geqslant 0,-r<t<r$。根据定理 2.1.2,$k(x,x')=f(\langle x,x'\rangle)$是核函数。证毕。

例 3.3 非齐次多项式核

$k(x,x')=(\langle x,x'\rangle+1)^p,p\in\mathbb{N}$

这里 $f(t)=(t+1)^p,-\infty<t<\infty$。当 $0<t<+\infty$时,

$$f^{(n)}(t)=\begin{cases}p(p-1)\cdots(p-n+1)(t+1)^{p-n}>0,n\leqslant p\\0,n>p\end{cases}$$

由定理 2.2.2,$k(x,x')=(\langle x,x'\rangle+1)^p$ 是核函数。多项式核函数是常用的核函数之一。

例 3.4 $k(x,x')=e^{\langle x,x'\rangle}$

这里 $f(t)=e^t,-\infty<t<\infty$。$f^{(n)}(t)=e^t>0$,所以,$k(x,x')=e^{\langle x,x'\rangle}$是核函数。

例 3.5 感知器核

$k(x,x')=\tanh(\rho\langle x,x'\rangle+c),\rho,c>0$

感知器核是常用核函数之一。这里强调的是,它仅对某些 $\rho,c>0$ 才是正定的。至于使得 k 正定,这些参数所应满足的充要条件,至今还没有答案。感知器核还有其他特点。比如,虽然对某些 $\rho,c>0$,它不是正定的,却是条件正定的;又如对某些参数,它与高斯核函数很近似。这些可能是它在应用中表现不俗的原因之一。详细的讨论见文[10]。

例 3.6 设 $X=[\{x\in\mathbb{R}^d:\|x\|<1\},a>0$,则 $k(x,x')=(1-\langle x,x'\rangle)^{-a}$是 $X\times X$ 上核函数。因为$|t|<1$ 时

$$(1-t)^{-a}=\sum_{n=0}^{\infty}\binom{-\alpha}{n}(-1)^nt^n$$

并且$\binom{-\alpha}{n}(-1)^n\geqslant 0,\forall n\geqslant 0$。

3.3 卷积核

另一个构造核函数的方法是卷积核。按照这种方法,我们甚至能够构造集上的核函数、序列上的核函数、树上的核函数等复杂结构上的核函数。先介绍一个概念。

定义 3.3.1 设 $X_1,\cdots,X_D,X$ 是非空集合,其中 D 是正整数;R 是 $X_1\times\cdots\times X_D\times X$ 的一个子集,若$(x_1,\cdots,x_D,x)\in R$,则称 $x_1,x_2,\cdots,x_D,x$ 有关系 R,记作 $R(x_1,x_2,\cdots,x_D,x)$。

例如,设 Λ 是有限字符集,$X_1=X_2=X$ 是 Λ 上有限长字符串的全体构成的集合。定义 $R(x_1,x_2,x)$当且仅当 $x_1\circ x_2=x$,这里$\circ$表示两字符串串接起来。

关系是函数概念的推广。为书写方便,令 $\vec{x}=x_1,x_2,\cdots,x_D$,简记$R(x_1,x_2,\cdots,x_D,x)$为 $R(\vec{x},x)$。令 $R^{-1}(x)=\{\vec{x}:R(\vec{x},x)\}$。称 R 是有限的,如果对任意 $x\in X$,$R^{-1}(x)$是有限集。有了以上概念,下面定义卷积核。

定义 3.3.2 设 $x,y\in X$,$\vec{x}=x_1,x_2,\cdots,x_D\in R^{-1}(x)$,

(下转第 178 页)

由表 2 给出。

表 2　形式背景

	a	b	c	d	e
1		×	×	×	
2		×		×	×
3					×
4	×		×		×
5			×		
6		×		×	×

从表 2 可以看出，$a^*=4$，$b^*=126$，$c^*=145$，$d^*=126$，$e^*=2346$。

根据定理 1 及定义 1，我们得到

$EX(K_2)=\{U,2346,126,145,26,4,1,\phi\}$

$IN(K_2)=\{A,bcd,bde,ace,bd,c,e,\phi\}$

$L(K_2)=\{(U,\phi),(2346,e),(126,bd),(145,c),(26,bde),(4,ace),(1,bcd),(\phi,A)\}$

因为 $a^*=c^*\cap e^*$，所以 $a^*\notin CM_{EX}$。这样

$CM_{EX}=\{b^*,c^*,d^*,e^*\}=\{2346,126,145\}$

应用定理 4、定理 6 及定理 7，可确定形式背景 K_2 中的属性可分类为：

${}_e=\{c,e\}$，${}_k=\{b,d\}$，${}_t=\{a\}$

上面的结果说明，对任意的 $D\in RED(K_2)$，$c,e\in D$，$a\notin D$。容易验证 $D_1=\{b,c,e\}$ 和 $D_2\{c,d,e\}$ 是概念格 $L(K_2)$ 的两个约简。

结论　在形式概念分析中，属性的分类对知识约简具有重要的意义。目前，关于概念格属性分类和约简问题的研究已取得一些成果，这些结果大多是应用序结构及辨识矩阵方法来处理。本文研究了形式背景的外延(内涵)闭系统中不可约元的性质，揭示了不可约元与属性特征的关系，从而给出了一种属性特征分类识别方法。这种方法不仅操作简单，而且为概念格约简及知识发现的进一步研究提供了新的途径。

参考文献

1 Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival I ed. Ordered Sets. Reidel, Dordrecht, 1982. 445～470

2 Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations. New York: Springer-Verlag, 1999

3 Düntsch I, Gediga G. Algebraic aspects of attribute dependencies in information systems. Fundamenta Informaticae, 1997, 29: 119～133

4 Düntsch I, Gediga G. Approximation operators in qualitative data analysis. In: de Swart H, Orlowska E, Schmidt G, et al. eds. Theory and Application of Relational Structures as Knowledge Instruments. Springer, Heidelberg, 2003. 216～233

5 Pagliani P. From concept lattices to approximation spaces: Algebraic structures of some spaces of partial objects. Fundamenta Informaticae, 1993, 18(1): 1～25

6 Yao Y Y. Concept lattices in rough set theory. Proceedings of 23rd International Meeting of the North American Fuzzy Information Processing Society, 2004

7 Diday E, Emilion R. Maximal and stochastic Galois lattices. Discrete Applied Mathematics, 2003, 127: 271～284

8 Bělohlaávek R. Concept lattices and order in fuzzy logic. Annals of Pure and Applied Logic, 2004, 128: 277～298

9 张文修，梁怡，吴伟志. 信息系统与知识发现. 北京：科学出版社，2003

10 张文修，吴伟志，梁吉业，等. 粗糙集理论与方法. 北京：科学出版社，2001

(上接第 174 页)

$\vec{y}=y_1,y_2,\cdots,y_D\in R^{-1}(y)$。又设 K_i 是 $X_i\times X_i$ 上的核函数，$1\leqslant i\leqslant D$，$K_i(x_i,y_i)$ 给出 x_i 与 y_i 的相似性度量。x 与 y 的相似性度量由下面的广义卷积给出：

$$K(x,y)=\sum_{\vec{x}\in R^{-1}(x),\vec{y}\in R^{-1}(y)}\prod_{i=1}^{D}K_i(x_i,y_i)$$

上式在 $S\times S$ 上有定义，其中 $S=\{x: R^{-1}(x)\text{非空}\}\subset X$。$K$ 在 $X\times X$ 上的零扩张称为核函数 $K_1,K_2,\cdots,K_D$ 的 R-卷积，记作 $K_1*K_2*\cdots*K_D$。

定理 3.3.1[12]　设 $K_1,K_2,\cdots,K_D$ 分别是 $X_1\times X_1$，$X_2\times X_2$，$\cdots$，$X_D\times X_D$ 上的核函数，R 是 $X_1\times\cdots\times X_D\times X$ 上的有限关系，则 $K_1*K_2*\cdots*K_D$ 是 $X\times X$ 上的核函数。

例 3.7　设 Ω 是任意非空集合，映射 $f:\Omega\rightarrow X$。若 $k: X\times X\rightarrow\mathbb{R}$ 是核函数，则 $\tilde{k}:\Omega\times\Omega\rightarrow\mathbb{R}$

$\tilde{k}(\omega,\omega')=k(f(\omega),f(\omega'))$

是核函数。我们经常用到的 $k(x-x')=k(\|x-x'\|)$ 就属于这种类型，是由低维空间上核函数构造高维空间上核函数的常用方法。

卷积核是构造核函数，尤其是复杂结构上核函数的重要方法。很多构造核函数的方法，均可纳入卷积核这个框架。例如，设，

$x=(x_1,\cdots,x_d)$，$x'=(x'_1,\cdots,x'_d)$

• 核函数的张量积：

$\tilde{k}(x,x')=k(x_1,x'_1)\cdots\cdots+k(x_d,x'_d)$；

• 核函数的直和：

$\tilde{k}(x,x')=k(x_1,x'_1)+\cdots\cdots+k(x_d,x'_d)$；

• 核函数的投影：

$\tilde{k}(x,x')=k((x_1,0,\cdots,0),(x'_1,0,\cdots,0))$。

灵活地定义关系 R，能得到很多重要核函数。进一步的讨论见文[12]及那里的参考文献。

结束语　本文首先对核函数的基本性质做了深入分析和概括，然后对 3 类重要核函数，即平移不变核函数、旋转不变核函数和卷积核，提出了简单实用的判别准则，并据此验证和构造了很多重要核函数。这些工作，为核函数的设计奠定了理论基础。下一步的工作是，把问题领域的知识与核函数的设计结合起来，即把问题领域的知识嵌入核函数之中，这对提升支持向量机的性能非常重要。

参考文献

1 Burges C. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 1998, 2: 121～167

2 Vapnik V. Statistical learning theory. New York: John Wiley & Sons, Inc, 1998

3 Cristianini N, Shawe-Taylor J. An introduction to support Vector Machines: and other kernel-based learning methods. New York: Cambridge University Press, 1999

4 Hsu C-W, Chang C-C, Lin C-J. A practical guide to support vector classification. http://www.csie.ntu.edu.tw/~cjlin/papers.html, 2003

5 Poggio. On optimal nonlinear associative recall. Biological Cybernetics, 1975, 19: 201～209

6 Cortes C, Vapnik V. Support vector networks. Machine Learning. 1995, 20: 273～297

7 Ressel B C. Harmonic analysis on semigroups. Beijing: Springer-Verlag, World Publishing Corporation, 1984

8 Scholkopf S W. Generalization bounds for convex combinations of kernel functions: [Technical Report]. NeuroCOLT2. Technical Report Series NC2-TR-1998-022, August 1998

9 Cristianini N, Shawe-Taylor J, Kandola. On Kernel Target Alignment. In: Proceedings of the Neural Information Processing Systems, NIPS'01, MIT Press, 2002

10 Lin Hsuan-Tien, Lin Chih-Jen. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. http://www.csie.ntu.edu.tw/~cjlin/papers.html, 2003

11 Steinwart I. On the influence of the kernel on the consistency of support vector machines. The Journal of Machine Learning Research, 2002, 2: 67～93

12 Haussler D. Convolutional kernels on discrete structures: [technical report]. UCSC-CRL-99-10, Santa Cruz: Computer Science Department, University of California, 1999

13 Feller W. An introduction to probability theory and its applications. Vol 2. New York: Wiley, 1971