第一章 核函数

§1 多项式空间和多项式核函数

定义 1.1 （**核或正定核**） 设 X 是 R^n 中的一个子集，称定义在 $X\times X$ 上的函数 $\mathrm{K}(x,z)$ 是核函数，如果存在一个从 X 到 Hilbert 空间 H 的映射 Φ

$$\Phi: x \mapsto \Phi(x) \in H \tag{1.1}$$

使得对任意的 $x,z\in X$，

$$\mathrm{K}(x,z) = (\Phi(x)\cdot\Phi(z)) \tag{1.2}$$

都成立。其中 $(\cdot)$ 表示 Hilbert 空间 H 中的内积。

定义 1.2 （**d 阶多项式**）设 $x=([x]_1,[x]_2,\cdots,[x]_n)^T\in R^n$，则称乘积 $[x]_{j_1}[x]_{j_2}\ldots[x]_{j_d}$ 为 x 的一个 d 阶多项式，其中 $j_1,j_2,\ldots,j_d\in\{1,2,\ldots,n\}$。

1. 有序齐次多项式空间

考虑 2 维空间中（$x\in R^n$）的模式 $x=([x]_1,[x]_2)^T$，其所有的 2 阶单项式为

$$[x]_1^2，[x]_2^2，[x]_1[x]_2，[x]_2[x]_1 \tag{1.3}$$

注意，在表达式(1.3)中，我们把 $[x]_1[x]_2$ 和 $[x]_2[x]_1$ 看成两个不同的单项式，所以称式（1.3）中的单项式为有序单项式。这 4 个有序单项式张成的是一个 4 维特征空间，称为 2 阶有序齐次多项式空间，记为 H。相应地可建立从原空间 R^2 到多项式空间 H 的非线性映射

$$C_2: x=([x]_1,[x]_2)^T \mapsto C_2(x) = ([x]_1^2,[x]_2^2,[x]_1[x]_2,[x]_2[x]_1)^T \in H \tag{1.4}$$

同理，从 R^n 到 d 阶有序齐次多项式空间 H 的映射可表示为

$$C_d: x=([x]_1,[x]_2,\ldots,[x]_n)^T \mapsto C_d(x) = ([x]_{j_1}[x]_{j_2}\ldots[x]_{j_d} \mid j_1,j_2,\ldots,j_d\in\{1,2,\ldots,n\})^T \in H \tag{1.5}$$

这样的有序单项式 $[x]_{j_1}[x]_{j_2}\ldots[x]_{j_d}$ 的个数为 n^d，即多项式空间 H 的维数 $n_H=n^d$。如果在 H 中进行内积运算 $C_d(x)\cdot C_d(z)$，当 n 和 d 都不太小时，多项式空间 H 的维数 $n_H=n^d$ 会相当大。如当 $n=200$，$d=5$ 时，维数可达到上亿维。显然，在多项式空间 H 中直接进行内积运算将会引起“维数灾难”问题，那么，如何处理这个问题呢？

我们先来考查 $n=d=2$ 的情况，计算多项式空间 H 中两个向量的内积

$$(C_2(x)\cdot C_2(z)) = [x]_1^2[z]_1^2 + [x]_2^2[z]_2^2 + [x]_1[x]_2[z]_1[z]_2 + [x]_2[x]_1[z]_2[z]_1 = (x\cdot z)^2 \tag{1.6}$$

若定义函数

$$\mathrm{K}(x,z)=(x\cdot z)^2 \tag{1.7}$$

则有

$$(C_2(x)\cdot C_2(z))=\mathrm{K}(x,z) \tag{1.8}$$

即 4 维多项式空间 H 上的向量内积可以转化为原始 2 维空间上的向量内积的平方。对于一般的从 R^n 到 d 阶有序多项式空间 H 的映射（1.5）也有类似的结论。

定理 1.1 考虑由式（1.5）定义的从 R^n 到多项式空间 H 的映射 $C_d(x)$，则在空间 H 上的内积 $(C_d(x)\cdot C_d(z))$ 可表为

$$(C_d(x)\cdot C_d(z))=\mathrm{K}(x,z) \tag{1.9}$$

其中

$$\mathrm{K}(x,z)=(x\cdot z)^d \tag{1.10}$$

证明：直接计算可得

$$\begin{aligned}(C_d(x)\cdot C_d(z))&=\sum_{j_1=1}^{n}\cdots\sum_{j_d=1}^{n}[x]_{j_1}\ldots[x]_{j_d}\cdot[z]_{j_1}\ldots[z]_{j_d}\\&=\sum_{j_1=1}^{n}[x]_{j_1}\cdot[z]_{j_1}\ldots\sum_{j_d=1}^{n}[x]_{j_d}\cdot[z]_{j_d}\\&=(\sum_{j=1}^{n}[x]_j\cdot[z]_j)^d=(x\cdot z)^d\end{aligned} \tag{1.11}$$

上述定理表明，我们并不需要在高维的多项式空间 H 中直接做内积运算 $(C_d(x)\cdot C_d(z))$，而利用式（1.10)给出的输入空间 R^n 上的二元函数 $\mathrm{K}(x,z)$ 来计算高维多项式空间中的内积。

2. 有序多项式空间

在式（1.5）定义的映射中，多项式空间 H 的分量由所有的 d 阶有序单项式组成。如果把该多项式空间的分量扩充为所有不超过 d 阶的有序单项式，便得到从 R^n 到有序多项式空间的映射 $\tilde{C}_d$

$$\tilde{C}_d: x=([x]_1,[x]_2,\ldots,[x]_n)^T\mapsto\tilde{C}_d(x)=([x]_{j_1}\ldots[x]_{j_d},\sqrt{d}[x]_{j_1}\ldots[x]_{j_d},\ldots,\sqrt{d}[x]_1,\ldots,\sqrt{d}[x]_n,1\,|\,j_1,j_2,\ldots,j_d\in\{1,2,\ldots,n\})^T \tag{1.12}$$

对于这个映射，我们有如下的定理：

定理 1.2　考虑有式（1.12）定义的从 R^n 到多项式空间 H 的映射 $\tilde{C}_d$ ，则空间 H 上的内积 $(\tilde{C}_d(x)\cdot\tilde{C}_d(z))$ 可表为空间 R^n 上的内积 $(x\cdot z)$ 的函数 $((x\cdot z)+1)^d$，即若定义两个变量 x 和 z 的函数

$$\mathrm{K}(x,z)=((x\cdot z)+1)^d \tag{1.13}$$

则有

$$(\tilde{C}_d(x)\cdot\tilde{C}_d(z))=\mathrm{K}(x,z) \tag{1.14}$$

上述有序多项式空间的一个简单的例子是

$$\tilde{C}_2: x=([x]_1,[x]_2)^T \mapsto \tilde{C}_2(x)=([x]_1^2,[x]_2^2,[x]_1[x]_2,[x]_2[x]_1,\sqrt{2}[x]_1,\sqrt{2}[x]_2,1)^T \tag{1.15}$$

3. 无序多项式空间

如果我们把式（1.4）中的 $[x]_1[x]_2$ 和 $[x]_2[x]_1$ 看作相同的单项式，那么我们就可以把从 R^2 到 4 维多项式空间 H 的映射（1.4）简化为从 R^2 到 3 维多项式空间的映射

$$([x]_1[x]_2)^T \mapsto ([x]_1^2,[x]_2^2,[x]_1[x]_2)^T \tag{1.16}$$

将映射（1.16）调整为

$$\Phi_2(x)=\Phi_2([x]_1,[x]_2)=([x]_1^2,[x]_2^2,\sqrt{2}[x]_1[x]_2) \tag{1.17}$$

则相应的多项式空间称为 2 阶无序多项式空间，并且有

$$(\Phi_2(x)\cdot\Phi_2(z))=(x\cdot z)^2 \tag{1.18}$$

对式（1.5）所示的变换 $C_d(x)$ 按下述方式操作：把 $C_d(x)$ 中次序不同但因子相同的各分量合并为一个分量，并在该分量前增加一个系数，这个系数取为相应次序不同但因子相同的分量在 $C_d(x)$ 中出现次数的平方根。这样得到的从 R^n 到 d 阶无序多项式空间的变换 $\Phi_d(x)$ 仍满足关系式

$$(\Phi_d(x)\cdot\Phi_d(z))=\mathrm{K}(x,z) \tag{1.19}$$

其中

$$\mathrm{K}(x,z)=(x\cdot z)^d \tag{1.20}$$

根据定义 1.1，我们称（1.13）和（1.20）分别为 d 阶多项式核函数和 d 阶齐次多项式核函数。

比较式（1.4）定义的变换 $C_2(x)$ 和式（1.17）定义的 $\Phi_2(x)$ 可以发现，它们所映射到的多项式空间是不同的。前者是一个 4 维多项式空间，后者为一个 3 维多项式空间。但是内

积是相同的，它们都可以表示为内积的函数$\mathrm{K}(x,z)=(x\cdot z)^2$。这说明：多项式空间不是由核函数唯一确定的。

§2　Mercer 核

1. 半正定矩阵的特征展开

给定向量集合$X=\{x_1,x_2,\ldots,x_l\}$，其中$x_i\in R^n, i=1,2,\ldots,l$　。设$\mathrm{K}(x,z)$是$X\times X$上的对称函数，我们定义

$$G_{ij}=K(x_i,x_j), i,j=1,2,\cdots,l \tag{1.21}$$

则称$G=(G_{ij})$是$\mathrm{K}(x,z)$关于X的 Gram 矩阵。我们首先要研究的问题是：当 Gram 矩阵G满足什么条件时，函数$K(\cdot,\cdot)$是一个核函数。

定义 1.2 （**矩阵算子**）定义在R^l上的矩阵算子G：对$u=(u_1,u_2,\ldots,u_l)^T\in R^l$，$Gu$的分量由下式确定

$$[Gu]_i=\sum_{j=1}^{l}K(x_i,x_j)u_j, i=1,2,\ldots,l \tag{1.22}$$

定义 1.3 （**特征值和特征向量**）考虑定义 1.2 给出的矩阵算子G。称$\lambda\in R$为它的特征值，并称v为相应的特征向量，如果

$$Gv=\lambda v \qquad \text{且 } v\neq 0 \tag{1.23}$$

定义　1.4（**半正定性**）　考虑定义 1.2 给出的矩阵算子G。称它是半正定的，如果对$\forall u=(u_1,u_2,\ldots,u_l)^T\in R^l$，有

$$u^TGu=\sum_{i,j=1}^{l}K(x_i,x_j)u_iu_j\geq 0 \tag{1.24}$$

引理 1.1　　若定义 1.2 给出的矩阵算子G是半正定的，则存在着l个非负特征值λ_t和互相正交的单位特征向量v_t，使得

$$K(x_i,x_j)=\sum_{t=1}^{l}\lambda_t v_{ti}v_{tj}\,,\quad i,\ j=1,2,\ldots,n \tag{1.25}$$

证明：　由于G是对称的，所以存在着正交矩阵$V=(v_1,v_2,\cdots,v_l)$和对角矩阵$\Lambda=diag(\lambda_1,\lambda_2,\cdots,\lambda_l)$，使得

$$G=V\Lambda V^T \tag{1.26}$$

这里 $v_t=(v_{t1},v_{t2},\cdots,v_{tl})^T$ 是矩阵 G 的第 t 个特征向量，它对应的特征值是 λ_t。因为 G 是半正定的，所以所有特征值均为非负数。于是由（1.26）推知

$$K(x_i,x_j)=\sum_{t=1}^{l} v_{ti}\lambda_t v_{tj}=\sum_{t=1}^{l}\lambda_t v_{ti}v_{tj} \tag{1.27}$$

引理 1.2 若引理 1.1 的结论成立，则存在着从 X 到 R^l 的映射 Φ，使得

$$K(x_i,x_j)=(\Phi(x_i)\cdot\Phi(x_j)),i,j=1,2,\cdots,l \tag{1.28}$$

其中 $(\cdot)$ 是特征空间 R^l 的内积。因而 $K(\cdot,\cdot)$ 是一个核函数。

证明： 定义映射

$$\Phi: x_i \mapsto \Phi(x_i)=(\sqrt{\lambda_1}v_{1i},\sqrt{\lambda_2}v_{2i},\cdots,\sqrt{\lambda_l}v_{li})^T\in R^l \tag{1.29}$$

直接验证可知引理 1.2 成立。

引理 1.3 若引理 1.2 的结论成立，则矩阵 G 是半正定的。

证明： 设 G 不是半正定的，则一定存在着与一个负特征值 λ_s 相对应的单位特征向量 v_s。定义 R^l 中的向量 z

$$z=[\Phi(x_1),\Phi(x_2),\cdots,\Phi(x_l)]v_s \tag{1.30}$$

则有

$$0\le\|z\|^2=v_S^T[\Phi(x_1),\cdots,\Phi(x_l)]^T[\Phi(x_1),\cdots,\Phi(x_l)]v_S=v_S^T K v_S=\lambda_S\|v_S\|^2=\lambda_S \tag{1.31}$$

显然，这与 λ_s 是负特征值相矛盾。因此 K 必须是半正定的。

定理 1.3 设 X 是有限集合 $X=\{x_1,x_2,\ldots,x_l\}$，$K(x,z)$ 是定义在 $X\times X$ 上的对称函数。则由定义 1.2 给出的矩阵算子 G 半正定，等价于 $K(\cdot,\cdot)$ 可表示为

$$K(x_i,x_j)=\sum_{t=1}^{l}\lambda_t v_{ti}v_{tj} \tag{1.32}$$

其中 $\lambda_t\ge 0$ 是矩阵

$$G=(K(x_i,x_j))_{i,j=1}^{l} \tag{1.33}$$

的特征值，$v_t=(v_{t1},v_{t2},\cdots,v_{tl})^T$ 为对应于 λ_t 的特征向量，也等价于 $K(x,z)$ 是一个核函数，即 $K(x_i,x_j)=(\Phi(x_i)\cdot\Phi(x_j))$，其中映射 Φ 由式（1.29）定义。

2. 半正定积分算子的特征展开

设输入集合为R^n中的紧集X，并设$K(x,z)$是$X\times X$的连续对称函数。我们要研究的问题是，当$K(x,z)$满足什么条件时，它是一个核函数。

定义 1.5 （**积分算子T_K**） 定义积分算子T_K为按下式确定的在$L_2(x)$上的积分算子

$$T_K f = T_K f(\cdot) = \int_X \mathrm{K}(\cdot,z) f(z)dz, \forall f \in L_2(x) \tag{1.34}$$

定义 1.6 （**特征值和特征函数**） 考虑定义 1.5 给出的积分算子T_K，称λ为它的特征值，φ为相应的特征函数，如果

$$T_K\varphi = \lambda\varphi \tag{1.35}$$

定义 1.7 （**半正定性**）考虑定义 1.5 给出的积分算子T_K。称它是半正定的，如果对$\forall f \in L_2(x)$，有

$$\int_{X\times X} K(x,z) f(x) f(z)dxdz \ge 0 \tag{1.36}$$

引理 1.4 若定义 1.5 给出的积分算子T_K是半正定的，则存在着可数个非负特征值λ_t和相应的互相正交的单位特征函数$\varphi_t(x)$，使得$K(\cdot,\cdot)$可表示为$X\times X$上的一致收敛的级数

$$K(x,z) = \sum_{t=1}^{\infty} \lambda_t \varphi_t(x)\varphi_t(z) \tag{1.37}$$

引理 1.5 若引理 1.4 的结论成立，则存在着$X \in R^n$到 Hilbert 空间l_2的映射Φ，使得

$$K(x,z) = (\Phi(x)\cdot\Phi(z)), \qquad x,z \in X \tag{1.38}$$

其中$(\cdot)$是l_2上的内积。因而$K(\cdot,\cdot)$是一个核函数。

证明： 定义映射

$$\Phi\text{：} x \mapsto \Phi(x) = (\sqrt{\lambda_1}\varphi_1(x), \sqrt{\lambda_2}\varphi_2(x), \cdots)^T \tag{1.39}$$

则可验证引理 1.5 成立。

引理 1.6 若引理 1.5 的结论成立则积分算子T_k是半正定的。

定理 1.4（Mercer 定理） 令X是R^n上的一个紧集，$K(x,z)$是$X\times X$上的连续实值对称函数。则由定义 1.5 给出的积分算子T_k半正定

$$\int_{X\times X} K(x,z)f(x)f(z)dxdz \ge 0 \quad , \quad \forall f \in L_2(x) \tag{1.40}$$

等价于 $K(\cdot\,,\cdot)$ 可表示为 $X\times X$ 的一致收敛序列

$$K(x,z)=\sum_{i-1}^{\infty}\lambda_t\varphi_t(x)\cdot\varphi_t(z) \tag{1.41}$$

其中 $\lambda_t>0$ 是 T_k 的特征值，$\varphi_t\in L_2(x)$ 是对应 λ_t 的特征函数。它也等价于 $K(x,z)$ 是一个核函数

$$K(x,z)=(\Phi(x)\cdot\Phi(z)) \tag{1.42}$$

其中映射 Φ 由式（1.39）定义，而 $(\cdot)$ 是 Hilbert 空间 l_2 上的内积。

定义 1.8 （**Mercer 核**） 称函数 $K(x,z)$ 为 Mercer 核，如果 $K(x,z)$ 是定义在 $X\times X$ 上的连续对称函数，其中 X 是 R^n 的紧集，且由定义 1.5 给出的积分算子是半正定的。

定理 1.5 设 X 为 R^n 上的紧集，$\mathrm{K}(x,z)$ 是 $X\times X$ 上的连续对称函数，则积分算子 T_K 半正定的充要条件是 $\mathrm{K}(x,z)$ 关于任意的 $x_1,x_2,\cdots,x_l\in X$ 的 Gram 矩阵半正定。

§3 正定核

定理 1.6 设 X 是 R^n 的子集。若 $\mathrm{K}(x,z)$ 是定义在 $X\times X$ 上的正定核，则对 $\forall x_1,x_2,\cdots,x_l\in X$，函数 $\mathrm{K}(x,z)$ 关于 $x_1,x_2,\cdots,x_l$ 的 Gram 矩阵都是半正定的。

证明： $\mathrm{K}(x,z)$ 是定义在 $X\times X$ 上的正定核，因此存在着从 X 到 Hilbert 空间 H 的映射 Φ，使得

$$K(x,z)=(\Phi(x)\cdot\Phi(z)) \tag{1.43}$$

任取 $x_1,x_2,\cdots,x_l\in X$，构造 $\mathrm{K}(\cdot,\cdot)$ 关于 $x_1,x_2,\cdots,x_l$ 的 Gram 矩阵 $(\mathrm{K}_{ij})_{i,j=1}^{l}=(K(x_i,x_j))_{i.j=1}^{l}$。显然，根据由式（1.43）可以断言，对 $\forall C_1,C_2,\cdots,C_l\in R$，我们有

$$\sum_{i,j}C_iC_j\mathrm{K}(x_i,x_j)=\sum_{i,j}C_iC_j\Phi(x_i)\cdot\Phi(x_j)=(\sum_i C_i\Phi(x_i)\cdot\sum_j C_j\Phi(x_j))=\left\|\sum_i C_j\Phi(x_i)\right\|^2\ge 0 \tag{1.44}$$

这表明 $\mathrm{K}(x,z)$ 关于 $x_1,x_2,\cdots,x_l$ 的 Gram 矩阵是半正定的。

引理 1.7 若集合 S 由所有的下列元素组成

$$f(\cdot)=\sum_{i=1}^{l}\alpha_i \mathrm{K}(\cdot,x_i) \tag{1.45}$$

其中l为任意的正整数，$\alpha_1,\alpha_2,\cdots,\alpha_l \in R^n$, $x_1,x_2,\cdots,x_l \in X$,则 S 为一个向量空间。

证明： 由于集合 S 中的元素对于加法和数乘封闭，所以 S 构成一个向量空间。

引理 1.8 若对 S 中的两元素

$$f(\cdot)=\sum_{i=1}^{l}\alpha_i \mathrm{K}(\cdot,x_i) \text{ 和 } g(\cdot)=\sum_{j=1}^{l}\beta_j \mathrm{K}(\cdot,x_j) \tag{1.46}$$

定义运算$*$

$$f*g=\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\beta_j \mathrm{K}(x_i,x_j) \tag{1.47}$$

并由此定义在$S\times S$上的函数$\tilde{\mathrm{K}}(f,g)=f*g$，则该函数关于$\forall f_1,f_2,\cdots,f_l\in S$的 Gram 矩阵都是半正定的。

证明： 由$f*f=\sum_{i,j=1}^{l}\alpha_i\alpha_j \mathrm{K}(x_i,x_j)\ge 0$知：

若任意选取$f_1,f_2,\cdots,f_l\in S$，记函数$\tilde{\mathrm{K}}$相应的 Gram 矩阵为$(f_i*f_j)_{i.j=1}^{l}$。显然它是对称矩阵。由（1.47）可知对$\forall C_1,C_2,\cdots,C_l\in R$有：

$$\sum_{i,j=1}^{l}C_iC_j(f_i*f_j)=(\sum_{i=1}^{l}C_i f_i)*(\sum_{i=1}^{l}C_j f_j)\ge 0 \tag{1.48}$$

这表明 Gram 矩阵$(f_i*f_j)_{i.j=1}^{l}$是半正定的。

引理 1.9 在引理 1.8 中定义的运算$*$具有如下性质：对于$\forall f,g\in S$，有

$$|f*g|^2\le (f*f)\cdot(g*g) \tag{1.49}$$

证明： 任取$f,g\in S$,则$\tilde{\mathrm{K}}$关于f,g的 Gram 矩阵为

$$\begin{bmatrix}\tilde{\mathrm{K}}(f,f) & \tilde{\mathrm{K}}(f,g)\\ \tilde{\mathrm{K}}(g,f) & \tilde{\mathrm{K}}(g,g)\end{bmatrix} \tag{1.50}$$

因为$\tilde{\mathrm{K}}(f,g)=\tilde{\mathrm{K}}(g,f)$，所以由引理 1.8 可知：矩阵（1.50）是半正定的，其行列式非负。

由此可知 $$\tilde{\mathrm{K}}(f,f)\cdot\tilde{\mathrm{K}}(g,g)-\tilde{\mathrm{K}}(f,g)\cdot\tilde{\mathrm{K}}(g,f)\ge 0$$

$$\Rightarrow \left|\tilde{\mathrm{K}}(f,g)\right|^2 \le \tilde{\mathrm{K}}(f,f)\cdot\tilde{\mathrm{K}}(g,g)$$
$$\Rightarrow \left|f*g\right|^2 \le (f*f)\cdot(g*g) \tag{1.51}$$

引理 1.10 引理 1.8 中定义的运算$*$是 S 上的内积运算，因而可记为

$$f*g=(f\cdot g) \tag{1.52}$$

证明： 直接验证可知该运算具有内积运算应满足的如下性质：对 $\forall f,g,h\in S$ 和 $c,d\in R$ 有

$$f*f\ge 0 \tag{1.53}$$

$$f=0\Rightarrow f*f=0 \tag{1.54}$$

$$(cf+dg)*h=c(f*h)+d(g*h) \tag{1.55}$$

$$f*g=g*f \tag{1.56}$$

只需证明：若 $f*f=0$，则有 $f=0$。事实上，若

$$f(\cdot)=\sum_{i=1}^{l}\alpha_i\mathrm{K}(\cdot,x_i) \tag{1.57}$$

则按运算规则（1.47）知，对 $\forall x\in X$，有

$$\mathrm{K}(\cdot,x)*f=f(x) \tag{1.58}$$

由于

$$\left|\mathrm{K}(\cdot,x)*f\right|^2\le(\mathrm{K}(\cdot,x)\cdot\mathrm{K}(\cdot,x))\cdot(f*f)=(\mathrm{K}(x,x)\cdot(f*f)) \tag{1.59}$$

所以

$$\left|f(x)\right|^2=\mathrm{K}(\cdot,x)\cdot(f*f) \tag{1.60}$$

此式意味着当 $f*f=0$ 时，对 $\forall x$，都有 $\left|f(x)\right|=0$，即 $f(x)$ 为零元素。

引理 1.11 若 H 是引理 1.7 中的集合 S 在引理 1.8 中定义的内积运算意义下的闭包，则 H 是一个 Hilbert 空间。

定理 1.7 设 $\mathrm{K}(x,z)$ 是定义在 $X\times X$ 上的对称函数。若对 $\forall x_1,x_2,\cdots,x_l\in X$，函数 $\mathrm{K}(x,z)$ 关于 $x_1,x_2,\cdots,x_l$ 的 Gram 矩阵都是半正定的，则 $\mathrm{K}(x,z)$ 是一个正定核。

证明： 定义映射

$$\Phi: x\to\mathrm{K}(\cdot,x) \tag{1.61}$$

由引理 1.7 和 1.11 知，该映射是从 X 到某一 Hilbert 空间的映射。由式（1.58）可得到

$$\mathrm{K}(\cdot,x)*\mathrm{K}(\cdot,z)=\mathrm{K}(x,z) \tag{1.62}$$

由引理 1.10 知引理 1.8 中定义的运算$*$是内积运算。利用式（1.61）可得到

$$\mathrm{K}(x,z)=(\Phi(x)\cdot\Phi(z)) \tag{1.63}$$

由定义 1.1 知$\mathrm{K}(x,z)$是正定核。

定义 1.12 （正定核的等价定义） 设 X 是 R^n 的子集。称定义在 $X\times X$ 上的对称函数 $\mathrm{K}(x,z)$ 为一个正定核，如果对 $\forall x_1,x_2,\cdots,x_l\in X$， $\mathrm{K}(\cdot,\cdot)$ 相对于 $x_1,x_2,\cdots,x_l$ 的 Gram 矩阵都是半正定的。

定义 1.13 （再生核的 Hilbert 空间） 令 X 是一个非空的集合，H 是一个由函数 f: $X\to R$ 组成的，内积由式（1.47）定义以及范数由 $\|f\|\overset{\Delta}{=}\sqrt{f\cdot f}$ 定义的 Hilbert 空间。称 H 是一个再生核 Hilbert 空间（简称 RKHS）,如果存在 K: $X\times X\to R$ 满足如下性质

（1） K 具有再生性，即对 $\forall f\in H$，有

$$(f\cdot\mathrm{K}(x,\cdot))=f(x) \tag{1.64}$$

特别地

$$(\mathrm{K}(x,\cdot)\cdot\mathrm{K}(\cdot,z))=\mathrm{K}(x,z) \tag{1.65}$$

（2） K 张成空间 H，即

$$H=\overline{span\{\mathrm{K}(x,\cdot)\mid x\in X\}} \tag{1.66}$$

其中 $\overline{A}$ 表示集合 A 的闭包。

若函数 K 是 Mercer 核，则对 $\forall c\in R^m$，有

$$\sum_{i,j}^{l}C_iC_j\mathrm{K}(x_i,x_j)=\sum_{i,j}^{l}C_iC_j\Phi(x_i)\Phi(x_j)=\left\|\sum_i C_i\Phi(x_i)\right\|^2\geq 0$$

因此，K 一定是一个正定核。因为 Mercer 是正定的，所以它是再生核。

§4 核函数的构造

根据正定核的等价定义，我们可以从简单的核来构造复杂的核。

定理 1.8 设 $\mathrm{K}_3(\theta,\theta')$ 是 $R^l\times R^l$ 上的核。若 $\theta(x)$ 是从 $X\subset R^n$ 到 R^l 的映射，则 $\mathrm{K}(x,z)=\mathrm{K}_3(\theta(x),\theta(z))$ 是 $R^n\times R^n$ 上的核。特别地，若 $n\times n$ 矩阵 B 是半正定的，则 $\mathrm{K}(x,z)=x^TBz$ 是 $R^n\times R^n$ 的核。

证明： 任取$x_1, x_2, \cdots, x_l \in X$，则$\mathrm{K}(x,z)=\mathrm{K}_3(\theta(x),\theta(z))$相应的 Gram 矩阵为

$$(\mathrm{K}(x_i, x_j))_{i,j=1}^{l} = (\mathrm{K}_3(\theta(x_i),\theta(x_j)))_{i,j=1}^{l} \tag{1.67}$$

记$\theta(x_t)=\theta_t$， $t=1,2,\ldots,l$，则有

$$(\mathrm{K}(x_i, x_j))_{i,j=1}^{l} = (\mathrm{K}_3(\theta_i,\theta_j))_{i,j=1}^{l} \tag{1.68}$$

由$\mathrm{K}_3(\theta,\theta')$是$R^l \times R^l$上的正定核可知：上式右端矩阵是半正定的。从而左端矩阵半正定。所以$\mathrm{K}(x,z)=\mathrm{K}_3(\theta(x),\theta(z))$是正定核。

当 B 为半正定矩阵时，它可分解为

$$B = V^T \Lambda V \tag{1.69}$$

定义$R^l \times R^l$上的核$\mathrm{K}_3(\theta,\theta')=(\theta,\theta')$，令$\theta(x)=\sqrt{\Lambda}Vx$，则有

$$\mathrm{K}(x,z)=\mathrm{K}_3(\theta(x),\theta(z))=\theta(x)^T\theta(z)=x^TV^T\sqrt{\Lambda}\sqrt{\Lambda}Vz=x^TBz \geq 0 \tag{1.70}$$

从而$\mathrm{K}(x,z)=x^TBz$是正定核。

定理 1.9 若 $f(\cdot)$是定义在$X \subset R^n$上的实值函数，则$\mathrm{K}(x,z)=f(x)\cdot f(z)$是正定核。

证明： 只需把双线性形式重写如下

$$\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j \mathrm{K}(x_i,x_j)=\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j f(x_i)\ f(x_j)$$

$$=\sum_{i=1}^{l}\alpha_i f(x_i)\cdot\sum_{j=1}^{l}\alpha_j f(x_j)$$

$$=(\sum_{i=1}^{l}(\alpha_i f(x_i))^2 \geq 0 \tag{1.71}$$

定理 1.10 设K_1和K_2是$X \times X$上的核，$X \subset R^n$。设常数$a \geq 0$，则下面的函数均是核：

（1） $$\mathrm{K}(x,z)=K_1(x,z)+K_2(x,z) \tag{1.72}$$

（2） $$\mathrm{K}(x,z)=aK_1(x,z) \tag{1.73}$$

（3） $$\mathrm{K}(x,z)=K_1(x,z)\cdot K_2(x,z) \tag{1.74}$$

证明： 对给定的一个有限集合$\{x_1, x_2, \cdots, x_l\}$，令$K_1$和$K_2$分别是$K_1$和$K_2$相对于

这个集合的 Gram 矩阵。

（1） 对$\forall \alpha \in R^l$,有

$$\alpha^T(\mathrm{K}_1+\mathrm{K}_2)\alpha = \alpha^T \mathrm{K}_1 \alpha + \alpha^T \mathrm{K}_2 \alpha \geq 0 \tag{1.75}$$

所以$\mathrm{K}_1+\mathrm{K}_2$是半正定的，因而$\mathrm{K}_1+\mathrm{K}_2$是核函数。

（2）$\alpha^T a \mathrm{K}_1 \alpha = a\alpha^T \mathrm{K}_1 \alpha \geq 0 \Rightarrow a\mathrm{K}_1$是核函数。

（3） 设K为$\mathrm{K}(x,z)=K_1(x,z)\cdot K_2(x,z)$对应于$\{x_1,x_2,\cdots,x_l\}$的 Gram 矩阵，则 K的元素是K_1和K_2对应元素的乘积

$$\mathrm{K}=\mathrm{K}_1 \circ K_2 \tag{1.76}$$

现证明K是半正定矩阵。令$K_1=C^T C$, $K_2=D^T D$，则

$$\begin{aligned} x^T(\mathrm{K}_1 \circ K_2)x &= tr[(diagx)K_1(diagx)K_2^T] \\ &= tr[(diagx)C^T C(diagx)D^T D] \\ &= tr[D(diagx)C^T C(diagx)D^T] \\ &= tr[C(diagx)D^T]^T[C(diagx)D^T] \\ &\geq 0 \end{aligned} \tag{1.77}$$

定理 1.11 设$\mathrm{K}(x,z)$是$X\times X$上的核。又设$p(x)$是系数全为正数的多项式。则下面的函数均是核。

（1） $$\mathrm{K}(x,z)=p(K_1(x,z)) \tag{1.78}$$

（2） $$\mathrm{K}(x,z)=\exp(K_1(x,z)) \tag{1.79}$$

（3） $$\mathrm{K}(x,z)=\exp(-\frac{\|x-z\|^2}{2}) \tag{1.80}$$

证明： （1）记系数全为正数的q阶多项式为$p(x)=a_q x^q+\cdots+a_1 x+a_0$，则有

$$\mathrm{K}(x,z)=p(K_1(x,z))=a_q[K_1(x,z)]^q+\cdots+a_1 K_1(x,z)+a_0 \tag{1.81}$$

由定理 1.10 知结论成立。

（2）由于指数函数可以用多项式无限逼近，所以$\exp(K_1(x,z))$是核函数的极限。再注意到核函数是闭集，便知结论成立。

（3） 由于

$$\exp(-\frac{\|x-z\|^2}{\sigma^2}) = \exp(-\frac{\|x\|^2}{\sigma^2}) \cdot \exp(-\frac{\|z\|^2}{\sigma^2}) \cdot \exp(\frac{2(x \cdot z)}{\sigma^2}) \tag{1.82}$$

由定理 1.9 知，上式右端前两个因子构成一个正定核。由定理 1.8 知，$(x \cdot z)$ 是一个正定核。从结论（2）可知：第三个因子也是一个正定核。由定理 1.10 的结论（3），便知结论成立。