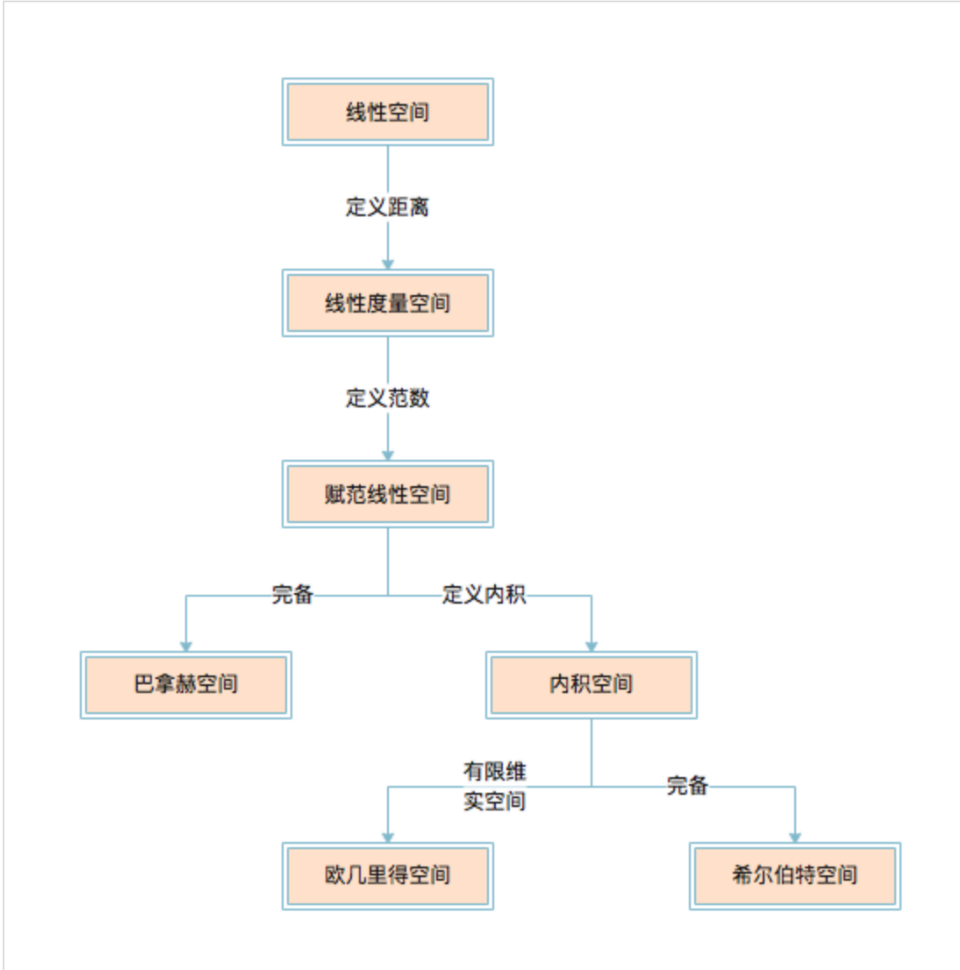

线性空间

线性空间（linear space）是定义了数乘和加法的空间。所以我们可以找到一组基底（basis），然后通过这一组基底的线性组合来得到空间中的所有点。

举个例子，二次函数空间，基底就可以定义为 $\{1, x, x^2\}$ ，这个空间中任意的函数 $f(x)$ 都可以表示为 $f(x) = \alpha_1 \cdot 1 + \alpha_2 \cdot x + \alpha_3 \cdot x^2$ ，所以我们可以说 $f(x)$ 在这组基底下的坐标是 $\{\alpha_1, \alpha_2, \alpha_3\}$ 。（请自行想象三维空间中，三个维度的坐标不是 $\{x, y, z\}$ 。）

当然基底也可以定义为 $\{1, x+1, (x+1)^2\}$ ，但是计算上会增大难度。所以我们可以看出来基底只要是线性不相关的就行了。那么基底中有 n 个元素的话，我们就叫这个空间是 n 维的。

一般来讲，选用的基底都是正交基（orthogonal basis），也就是基底中的任意两个元素的内积为0。

线性度量空间

线性度量空间（metric linear space）是在线性空间中定义了**距离（metric）**的空间。

距离的定义必须满足如下三个条件：

1. 非负性：$d(x, y) \geq 0; d(x, y) = 0 \Leftrightarrow x = y$ 。
2. 对称性：$d(x, y) = d(y, x)$ 。
3. 三角不等式：$d(x, z) + d(z, y) \geq d(x, y)$ 。

线性赋范空间

线性赋范空间（normed linear space）是定义了**范数（norm）**的线性度量空间。

范数的定义必须满足：

1. 非负性：$\|x\| \geq 0$ 。
2. 齐次性：$\|\alpha x\| = |\alpha| \|x\|$ 。
3. 三角不等式：$\|x\| + \|y\| \geq \|x + y\|$ 。

由范数可以导出距离（定义 $d(x, y) = \|x - y\|$ ），但是不可以由距离导出范数。

巴拿赫空间

巴拿赫空间（Banach space）是完备的赋范线性空间。

完备性（completeness）：任一柯西序列（Cauchy sequence）都收敛（convergence）。

内积线性空间

内积线性空间（inner product linear space）是定义了**内积（inner product）**的赋范线性空间。

其中内积也叫**标量积（scalar product）**或**点积（dot product）**。

这里要注意，内积的定义跟范数其实没关系，不过内积可以导出范数，所以一般内积空间都有范数。

内积的定义必须满足：

1. 对称性：$\langle x, y \rangle = \langle y, x \rangle$ 。
2. 线性性：$\langle x, y \rangle + \langle x, z \rangle = \langle x, y + z \rangle$ 、$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ 。注意：数乘只对第一变元有效。
3. 正定性：$\langle x, x \rangle \geq 0$ 。

由内积可以导出范数（定义 $\|x\|^2 = \langle x, x \rangle$ ），但是范数不可以导出内积。

欧几里得空间

欧几里得空间（Euclidean space）是有限维的实内积线性空间。

希尔伯特空间

希尔伯特空间（Hilbert space）是完备的内积线性空间。

两个例子

1. 泰勒级数展开（Taylor series）：将一个函数用 $\{x^i\}_0^\infty$ 作为基底表示的一个空间。
2. 傅里叶级数展开（Fourier series）：将一个函数用 $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots\}$ 作为基底表示的一个空间。

核函数

终于到主题了，写的手都酸了……

一般的欧式空间中，我们可以定义一个 $n \times n$ 矩阵的特征值和特征向量。

$$Ax = \lambda x$$

考虑一个矩阵的列空间，当这个矩阵可以进行特征值分解的时候，其特征向量就构成了这个 n 维空间的一组基底。

现在我们把这个概念推广到函数空间。

我们把每个函数 $f(x)$ 看作一个无穷维的向量，然后定义一个函数空间中无穷维的矩阵 $K(x, y)$，如果它满足：

1. 正定性：$\forall f \rightarrow \iint f(x)K(x,y)f(y)\,\mathrm{d}x\,\mathrm{d}y \geq 0$。
2. 对称性：$K(x,y) = K(y,x)$。

我们就把它称作**核函数（kernel function）**。

和特征值与特征向量的概念相似，存在特征值 λ 和特征函数 $\psi(x)$。满足：

$$\int K(x,y)\psi(x)\mathrm{d}x = \lambda\psi(y)$$

对于不同的特征值 λ_1、λ_2，对应不同的特征函数 $\psi_1(x)$、$\psi_2(x)$，很容易得到：

$$\begin{aligned}\int \lambda_1\psi_1(x)\psi_2(x)\mathrm{d}x &= \iint K(y,x)\psi_1(y)\mathrm{d}y\,\psi_2(x)\mathrm{d}x \\ &= \iint K(y,x)\psi_2(x)\mathrm{d}x\,\psi_1(y)\mathrm{d}y \\ &= \int \lambda_2\psi_2(y)\psi_1(y)\mathrm{d}y \\ &= \int \lambda_2\psi_2(x)\psi_1(x)\mathrm{d}x\end{aligned}$$

因此，

$$\langle\psi_1,\psi_2\rangle = \int \psi_1(x)\psi_2(x)\mathrm{d}x = 0$$

所以我们找到了一个可以生成这个空间的矩阵 K，一组无穷多个特征值 $\{\lambda_i\}_{i=1}^{\infty}$，和一组无穷多个元素的正交基 $\{\psi_i\}_{i=1}^{\infty}$。

再生核希尔伯特空间

如果我们把 $\{\sqrt{\lambda_i}\psi_i\}_{i=1}^{\infty}$ 当成一组正交基来生成一个希尔伯特空间 $\mathcal{H}$ 。则该空间中的所有函数都能表示为这组正交基的线性组合。

$$f = \sum_{i=1}^{\infty} f_i \sqrt{\lambda_i}\psi_i$$

于是我们就可以把函数 f ，看作 $\mathcal{H}$ 中的一个向量 $f = (f_1, f_2, \cdots)_{\mathcal{H}}^{\mathsf{T}}$ 。
对于另外一个函数 $g = (g_1, g_2, \cdots)_{\mathcal{H}}^{\mathsf{T}}$ ，我们有:

$$\langle f, g \rangle_{\mathcal{H}} = \sum_{i=1}^{\infty} f_i g_i$$

有了这个内积，我们就可以把核函数看成一种内积形式了，即:

$$K(x, \cdot) = \sum_{i=0}^{\infty} \lambda_i \psi_i(x)\psi_i(\cdot)$$

如果把 ψ_i 当成一个算子来看的话，我们就取函数名的一个形式：$K(x, \cdot) = \sum_{i=0}^{\infty} \lambda_i \psi_i(x)\psi_i$ 。
所以我们就可以把 K 当作一个向量来看了。

$$K(x, \cdot) = (\sqrt{\lambda_1}\psi_1(x), \sqrt{\lambda_2}\psi_2(x), \cdots)_{\mathcal{H}}^{\mathsf{T}}$$

因此,

$$\langle K(x, \cdot), K(y, \cdot) \rangle_{\mathcal{H}} = \sum_{i=0}^{\infty} \lambda_i \psi_i(x)\psi_i(y) = K(x, y)$$

这个性质就叫再生性（reproducing），这个 $\mathcal{H}$ 就叫做再生核希尔伯特空间（reproducing kernel Hilbert space，RKHS）。

回到我们最初的问题，怎么把一个点映射到一个特征空间上呢?

定义一个映射:

$$\Phi(x) = K(x, \cdot) = (\sqrt{\lambda_1}\psi_1(x), \sqrt{\lambda_2}\psi_2(x), \cdots)_{\mathcal{H}}^{\mathsf{T}}$$

则

$$\langle \Phi(x), \Phi(y) \rangle_{\mathcal{H}} = \langle K(x, \cdot), K(y, \cdot) \rangle_{\mathcal{H}} = K(x, y)$$