我们之前主要考虑的都是在一个固定线性空间里面做一些操作。这一章我们引进一个新的概念：线性映射。这个概念可以把前面的东西都串联在一起。

1 线性映射和矩阵

我们考虑一个映射$f : R^n \to R^m$ (映射就是一个规则，把R^n中的每一个元素映到R^m中的一个元素。)。这个映射称之为线性映射，如果 和R^n中的线性空间的运算是交换的。用数学的语言来描述就是：

$$f(c_1v_1 + c_2v_2 + \ldots + c_sv_s) = c_1f(v_1) + \ldots + c_sf(v_s). \tag{1}$$

因为上面这个重要的性质，对于线性映射来说，我们只需要给定它对一组基的映射规则即可，其他元素的映射规则可以用上面的 关系来决定。

重点： 给定一个线性映射, 我们可以得到一个矩阵。 我们需要选取R^n中的一组基$(v_1, \ldots, v_n)$, 并且选取R^m中的一组基$(w_1, \ldots, w_m)$. 那么$f(v_i)$是R^m中的一个向量，所以可以用$(w_1, \ldots, w_m)$的线性组合来描述：

$$f(v_i) = a_{1i}w_1 + a_{2i}w_2 + \ldots + a_{mi}w_m = \sum_{j=1}^{m} a_{ji}w_j \tag{2}$$

这些系数就得了一个$m \times n$的矩阵：

$$(A)_{ji} = a_{ji}, \quad j = 1, \ldots, m, \quad i = 1, \ldots, n. \tag{3}$$

规则就是把上述对$f(v_i)$的展开系数(2)当作矩阵A的第i列。

结论： 给定一个线性映射$f : R^n \to R^m$， 并且选取两个线性空间的一组基，我们可以得到一个$m \times n$的矩阵。

对于线性映射和它对应的矩阵，我们还有下面一些重要的性质

1. 对于R^n中的一个向量$\overrightarrow{x}$, 用我们给定的基，我们有关于x的线性组合：

$$\overrightarrow{x} = x_1v_1 + \ldots + x_nv_n \tag{4}$$

那么线性映射作用在x上的结果是：

$$f(\overrightarrow{x}) = f(x_1v_1 + \ldots + x_nv_n) = \sum_{i=1}^{n} f(v_i)x_i = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ji}x_iw_j \tag{5}$$

所以$f(x)$在(v, w)这组基下面的坐标就是

$$Ax \tag{6}$$

这里x是一个$n \times 1$的矩阵$x = [x_1, \ldots, x_n]^T$.

2. 如果我们在两个线性空间里面换一组基，那么我们有

$$v_i^{'} = \sum_{j=1}^{n} P_{ji}v_j, \quad w_i^{'} = \sum_{j=1}^{m} Q_{ji}w_j \tag{7}$$

这里P是一个$n \times n$的可逆矩阵，Q是一个$m \times m$的可逆矩阵。用矩阵乘法的表示，就是$V^{'} = VP$，$W^{'} = WQ$. 那么我们有

$$f(v_i^{'}) = f(\sum_{j=1}^{n} P_{ji} v_j) = \sum_{j=1}^{n} P_{ji} f(v_j) = \sum_{j=1}^{n}\sum_{k=1}^{m} P_{ji} a_{kj} w_k = \sum_{j=1}^{n}\sum_{k=1}^{m}\sum_{l=1}^{m} P_{ji} a_{kj} Q_{lk}^{-1} w_l^{'} \tag{8}$$

我们得到新的矩阵为

$$A^{'} = Q^{-1}AP \tag{9}$$

我们现在考虑在新的基下面对一个向量x的变换。 在原来的基下面，$\overrightarrow{x}$的坐标是$[x_1, \ldots, x_n]^T$：

$$\overrightarrow{x} = x_1 v_1 + \ldots + x_n v_n = \sum_{i=1}^{n} x_i v_i = Vx = V^{'}P^{-1}x = V^{'}P^{-1}x \tag{10}$$

所以在新的基下面$\overrightarrow{x}$的坐标矩阵为$x^{'} = P^{-1}x$. 那么线性映射的作用为

$$f(\overrightarrow{x}) = A^{'}x^{'} = Q^{-1}APP^{-1}x = Q^{-1}Ax \tag{11}$$

3. $ker(f)$ 定义为R^n中的向量，并且这些向量在f映射下变成R^m中的0向量，如果我们选定两个线性空间的基，那么这个空间就是下面这个关于x的方程的解:

$$Ax = 0 \tag{12}$$

这个空间其实就是我们之前研究的矩阵A的零空间。

4. $im(f)$定义为在f作用下在R^m里面的像。选了一组基之后，就是Ax, 这里x可以取任意的R^n里面的向量。 这其实是我们之前研究的列向量子空间！

5. 矩阵乘法。 我们现在考虑两个线性映射：$R^n \to^f R^k \to^g R^m$, 那么复合映射gof定义了一个从R^n到R^m的映射。 在三个线性空间里面选择基，如果f对应的矩阵为A（$k \times n$矩阵）, g 对应的矩阵为B（$m \times k$ 矩阵）。那么gof对应的矩阵$(m \times n)$是

$$BA \tag{13}$$

6. 逆映射和逆矩阵：考虑线性空间R^n到自己的线性映射，如果一个线性映射f的像都不一样，那么这个映射是可逆的， 记为f^{-1}. f^{-1}所对应的矩阵为

$$A^{-1} \tag{14}$$

2 基的变换和矩阵分解

我们之前关于矩阵分解的问题其实都可以归结于不同基的选取。 给定一个矩阵，可以看成一个线性映射在标准基下的表示。我们可以通过不同基的选取使得线性变换对应的矩阵取比较好的形式:

1. 高斯消去法或者LU分解是改变R^m中的基，对应于左乘一个可逆矩阵。如果我们同时做行变换和列变换，相当于左乘（行变换）和右乘（列变换）一个可逆矩阵。这样我们最后可以得到一个标准型，叫做相抵标准型:
$$\begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix} = PAQ \tag{15}$$

2. 方块可逆矩阵的QR分解也可以看成一个换基的操作, 来使得我们的矩阵变成一个正交矩阵:
$$Q = AR^{-1} \tag{16}$$

3. 我们可以考虑一个从R^n到它自己的线性映射，这样我们换基的话就是左乘和右乘相同的可逆矩阵 （也就是$P = Q$), 根据我们的换基变换公式，我们有
$$A^{'} = P^{-1}AP \tag{17}$$
这其实是相似变换。如果矩阵可以对角化，我们可以得到一个对角矩阵。如果矩阵不可以对角化，那么我们有所谓的若当标准型。这种标准型称之为相似标准型。

4. 奇异值分解可以写成
$$\Sigma = U^T AV \tag{18}$$
也可以看成是换基的操作，这里的换基操作是通过正交矩阵来实现的。

3 线性映射构成的线性空间

给定两个线性空间R^n和R^m, 考虑它们之间所有的线性映射，那么这些线性映射构成实数域上的线性空间。

证明：给定两个线性映射f_1和f_2, 它们的和也是线性映射:
$$(f_1 + f_2)(x) = f_1(x) + f_2(x). \tag{19}$$
一个实数乘上一个线性映射还是一个线性映射:
$$(cf)(x) = c(f(x)). \tag{20}$$
所以所有的线性映射在加法和数乘的运算下是封闭的，所以它们构成一个线性空间。这个线性空间的维数是$m \times n$.

同样的给定一个线性空间R^n，考虑所有到自己的线性映射。 这些映射构成一个线性空间。 这个线性空间的维数是n^2.

考虑线性映射$f : V \to R$. 所有这些线性映射构成的线性空间V^*称之为线性空间V的对偶空间。 给定R^n中的一组基$(e_1, \ldots, e_n)$, 我们可以构造出对偶空间的一组基$(e_1^*, \ldots, e_n^*)$, 这些线性映射的作用是
$$e_i^*(e^j) = \delta_{ij} \tag{21}$$

δ_{ij}的意义是$\delta_{ij} = 1, i = j$, 其他都等于0. 给定一组基以后，任何的线性映射$f: R^n \to R$ 都可以写成下列的形式:

$$f = f_1 e_1^* + \ldots + f_n e_n^* \tag{22}$$

我们现在考虑V换一组基下对偶空间的变化。换基有下面的公式：

$$e_i^{'} = \sum_{j=1} e_j P_{ji} \tag{23}$$

用矩阵的表达方式，我们有$E^{'} = EP$, 所以我们有$E = E^{'} P^{-1}$. 对于一个向量$\overrightarrow{x}$, 我们有下列的展开:

$$\overrightarrow{x} = x^1 e_1 + x^2 e_2 + \ldots + x^n e_n = Ex = E^{'}(P^{-1}x) \tag{24}$$

所以在新的下面$\overrightarrow{x}$的展开系数是

$$x^{'i} = \sum_{j=1}^{n} (P^{-1})^i{}_j x^j \tag{25}$$

我们要求对偶空间的基$e_i^{'*}$还满足同样的公式：

$$e_i^{'*}(e_j^{'}) = \delta_{ij} \tag{26}$$

这样我们就有

$$e_i^{'*} = \sum_{j=1} P_{ij}^{-1} e_j^* \tag{27}$$

任何一个线性映射有下面的展开：

$$f = f_1 e_i^{'*} + \ldots + f_n e_n^{'*} \tag{28}$$

在新的坐标下，我们有

$$f_i^{'} = \sum_{j=1}^{n} f_j P^j{}_i \tag{29}$$

这样任何一个线性映射$f: R^n \to R$ 的作用就是（选定线性空间的一组基和它的对偶基）

$$f(\overrightarrow{x}) = f_i x^i \tag{30}$$

并且这个作用形式是不依赖于基的选取的。

4 构造新的线性空间

我们讨论一下线性变换的一些性质(R^n到R^n的线性映射). 我们讨论了两个线性映射子空间$ker(f)$和$Im(f)$. 我们再定义一个子空间。 一个子空间V_1称之为线性映射的不变子空间， 如果对于任何$v \in V_1$, 我们有$f(v) \in V_1$. 也就是说f把V_1中的向量映射到V_1里面。 注意，不变子空间不是唯一的。

商空间：给定一个线性空间V和它的一个线性子空间U,我们可以定义一个商空间 $W=V/U$。商空间的的一个元素是V里面的一个等价类 $[v_1]$. v_1和v_2同属于$[v_1]$,如果它们之间差别一个U里面的元素：

$$v_1 = v_2 + u \tag{31}$$

这里$u \in U$.

我们讨论一下一个线性空间的直和. 一个线性空间V可以写成两个线性子空间 $V = V_1 \oplus V_2$的直和，如果我们有下列的条件:

1. V中任何一个向量都可以写成$v = v_1 + v_2$, 这里$v_1 \in V_1$, 并且$v_2 \in V_2$.

2. $V_1 \cap V_2 = \{0\}$. V_1 和 V_2只共有零向量。

类似的，我们可以把V分成很多子空间的直和: $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$.

通过已有的线性空间，我们可以构造新的线性空间。一个重要的观点是：选定一组基，我们的线性空间可以由这组基的线性组合来生成。我们主要有下面的两种构造。

1. 直积 $V = V_1 \times V_2$. 如果 V_1 的基是$(e_1, \ldots, e_n)$, V_2的基是 $(\eta_1, \ldots, \eta_m)$, 那么V的基是$(e_i, 0)$ 和 $(0, \eta_i)$. 这个新的线性空间的维数是 $n + m$.

2. 张量积 $V = V_1 \otimes V_2$. 如果 V_1 的基是$(e_1, \ldots, e_n)$, V_2的基是 $(\eta_1, \ldots, \eta_m)$. 那么 V的基是 $e_i \otimes \eta_j$. 这个线性空间的维数是nm.

我们下面重点考察由一个线性空间V生成的各种线性空间。

1. 我们可以考虑$V^l = V \otimes V \otimes V \otimes \ldots \otimes V$. 如果我们选择$V$的一组基，那么我们新的线性空间的基就是：$e_{i_1} \otimes e_{i_2} \ldots \otimes e_{i_l}$. 那么这个空间的一个向量就可以用这组基来展开，展开系数为

   $$V^{i_1 i_2 \ldots i_l} \tag{32}$$

   这个通常称之为$(0, l)$张量。

2. 我们有V的对偶空间 V^* 和 构造的基e_i^*, 那么我们可以构造$V^{*k} = V^* \otimes V^* \otimes V^* \otimes \ldots \otimes V^*$. 那么我们新的线性空间的基就是：$e^*_{i_1} \otimes e^*_{i_2} \ldots \otimes e^*_{i_k}$. 这个线性空间的任何一个元素的展开系数是

   $$V_{i_1 i_2 \ldots i_k} \tag{33}$$

   这个通常称之为$(k, 0)$张量。

3. 我们还可以考虑一个线性空间的对称积：$Sym^2 V$. 这个线性空间的基是$\frac{1}{2}(e_i \otimes e_j + e_i \otimes e_j)$. 这个线性空间的任何一个元素的展开系数是

   $$g^{ij} \tag{34}$$

   这里系数关于i和j是对称的。类似的，我们可以定义$Sym^2 V^*$.

4. 我们还可以考虑一个线性空间的反对称积：$\Lambda^2 V$. 这个线性空间的基是$\frac{1}{2}(e_i \otimes e_j - e_i \otimes e_j)$. 这个线性空间的任何一个元素的展开系数是

$$f^{ij} \tag{35}$$

这里系数关于i和j是反对称的。类似的，我们可以定义$\Lambda^2 V^*$.

5. 更一般的话，我们可以构造张量$(V^*)^k \otimes V^l$张量。 这些张量的基是$e^*_{i_1} \otimes e^*_{i_2} \otimes \dots e^*_{i_k} \otimes e_{j_1} \otimes e_{j_2} \otimes \dots \otimes e_{j_l}$. 这样一个 (k,l)张量的分量是

$$V^{j_1 j_2 \dots j_l}{}_{i_1 i_2 \dots i_k} \tag{36}$$

定义这些张量最重要的性质就是分量随着V的基的变化而变换。

$$e_i^{'} = \sum_{j=1}^{n} e_j P_{ji} \tag{37}$$

对于一个(k,l)张量，每一个上指标都乘上一个P矩阵，每一个下指标都乘上一个P^{-1} 矩阵。 上指标称之为协变指标，下指标称之为逆变指标。通常我们可以把矩阵P_{ij}写成$P^i{}_j$, 逆矩阵P^{-1}_{ij}写成$(P^{-1})_i{}^j$. 这样 张量的变换就可以省略掉求和，而用Einstein求和法则。

$$V^{j_1^{'} j_2^{'} \dots j_l^{'}}{}_{i_1^{'} i_2^{'} \dots i_k^{'}} = (P^{-1})^{j_1^{'}}{}_{j_1} \dots (P^{-1})^{j_l^{'}}{}_{j_l} V^{j_1 j_2 \dots j_l}{}_{i_1 i_2 \dots i_k} (P)^{i_1}{}_{i_1^{'}} \dots (P)^{i_k}{}_{i_k^{'}} \tag{38}$$

5 内积空间

给定一个线性空间 R^n，我们之前考虑了这个线性空间的一个内积。在标准基下面，这个内积定义为

$$v \cdot w = v^T w = w^T v. \tag{39}$$

更一般的，我们可以有一个和基选取无关的内积的定义。 内积是一个下面的双线性函数

$$g : V \times V \to R \tag{40}$$

这个双线性函数满足下面的条件：

1. 这个函数对每一个变量都是线性的：

$$g(au+bv,w) = ag(u,w) + bg(v,w), \quad g(w,av+bu) = ag(w,u) + bg(w,v). \tag{41}$$

2. 这个函数是对称的：$g(u,v) = g(v,u)$.

3. $g(v,v) \geq 0$. 并且 $g(v,v) = 0$当且仅当$v = 0$. 如果我们选定V中的一组基，因为g是多线性的，我们只需要知道下列实对称矩阵：

$$g(e_i, e_j) = g_{ij} \tag{42}$$

这个矩阵还需要满足下列条件

$$x^T g x \geq 0 \tag{43}$$

这里x是一个向量在我们取的基下面的坐标，并且等于0的时候当且仅当$x = 0$. 所以我们发现取了一组基之后，一个内积就是我们之前定义的正定二次型！实际上$g \in V^* \otimes V^*$ (检查一下g在换基下面的变换规则。).

内积空间的一个用处是可以把对偶空间V^*和线性空间V联系起来。我们构造的对应是这样的。

$$v \to g(v, \cdot) \tag{44}$$

这里$g(v, \cdot)$是个V上的线性映射： $g(v, \cdot) : V \to R$, 作用是$g(v, w)$, 这里v是固定的，w是任意的。

我们可以考虑对偶空间V^*的对偶空间V^{**}. V^{**} 和我们原来的线性空间V有一个自然的对应。这个对应的构造是这样的： 给定V里面的一个元素v, 我们需要找到V^* 上面的一个线性映射v^{**}. 这个线性映射需要作用在每一个V^*里面的元素：

$$v^{**}(f) = f(v). \tag{45}$$

这样定义的映射v^{**}是线性的。