17.1 线性变换

定义

1. (17.1.1 行向量) 设$n \geq 1$是一个整数，我们将$\mathbb{R}^n$中的元素称为n**维行向量**。我们一般将一个n维行向量记为$x = (x_1, x_2, \ldots, x_n)$，也可以记为$(x_i)_{1 \leq i \leq n}$，其中每一个分量$x_1, x_2, \ldots, x_n$都是实数。

   如果$(x_i)_{1 \leq i \leq n}$和$(y_i)_{1 \leq i \leq n}$都是n维行向量，那么我们定义他们的**向量和**为：

   $$(x_i)_{1 \leq i \leq n} + (y_i)_{1 \leq i \leq n} := (x_i + y_i)_{1 \leq i \leq n}$$

   另外，如果$c \in \mathbb{R}$是任意一个标量，那么我们定义**标量积**$c(x_i)_{1 \leq i \leq n}$为：

   $$c(x_i)_{1 \leq i \leq n} := (cx_i)_{1 \leq i \leq n}$$

   如果两个向量的维数不同，那么我们不定义他们之间的加法运算（例如$(1,2,3)+(4,5)$是无定义的）。另外我们称$\mathbb{R}^n$中的向量$(0, \ldots, 0)$为**零向量**，并记为0（严格来说应该记为$0_{\mathbb{R}^n}$，但是没有必要可以标注）。最后，我们将$(-1)x$简写为$-x$。

   （注：都是代数耳熟能详的内容了，本节差不多就是一个回顾）

2. (17.1.3 转置) 如果$(x_i)_{1 \leq i \leq n}$是一个n维行向量，那么我们定义它的**转置**$(x_i)_{1 \leq i \leq n}^T$为：

   $$(x_i)_{1 \leq i \leq n}^T = (x_1, x_2, \ldots, x_n)^T := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

   并且我们将形如$(x_i)_{1 \leq i \leq n}^T$的对象称为$n$**维列向量**。

   （注：这个定义纯粹是为了合并矩阵与向量的乘法）

3. (17.1.5 标准基行向量) 我们将$\mathbb{R}^n$中的n个特殊行向量$e_1, \ldots, e_n$称为**标准基行向量**。其中对每一个$1 \leq j \leq n$，e_j是第j个分量为1其余分量均为0的n维行向量。

   （注：因此$\mathbb{R}^n$中的每一个向量$x = (x_1, x_2, \ldots, x_n)$都可以表示为形如$\sum\limits_{j=1}^{n} x_j e_j$的标准基行向量的**线性组合**；类似地也可以定义**标准基列向量**；基向量可以有很多可能，但是这里不讨论这些，这里给出的是最简单的一种）

4. (17.1.6 线性变换) **线性变换**$T : \mathbb{R}^n \to \mathbb{R}^m$是一类满足特殊公理的从一个欧几里得空间$\mathbb{R}^n$到另一个欧几里得空间$\mathbb{R}$函数，具体需要满足：

   > 1. (可加性) 对于任意的$x, x' \in \mathbb{R}^n$，都有$T(x + x') = T(x) + T(x')$。
   > 2. (齐次性) 对于任意的$x \in \mathbb{R}^n$和任意的$c \in \mathbb{R}$，都有$T(cx) = cT(x)$。

   （注：本节中给出了几个线性变换的例子，看看就行）

5. (17.1.7 膨胀算子？) 对于任意的n，定义为$T_1 x := 5x$的**膨胀算子**$T_1 : \mathbb{R}^n \to \mathbb{R}^n$是一个线性变换。

6. (17.1.8 旋转算子？) **旋转算子**$T_1 : \mathbb{R}^n \to \mathbb{R}^n$的定义是将$\mathbb{R}^2$中的每一个向量都沿顺时针方向旋转$\pi/2$弧度，这个算子也是一个线性变换。

7. (17.1.9 三个例子？) 定义为$T_3(x, y, z) := (x, y)$的**射影算子**$T_3 : \mathbb{R}^3 \to \mathbb{R}^2$是一个线性变换；定义为$T_2(x, y) := (x, y, 0)$的**包含算子**$T_4 : \mathbb{R}^2 \to \mathbb{R}^3$是一个线性变换；最后，对于任意的$n$，定义为$I_n x := x$的**恒等算子**$I_n : \mathbb{R}^n \to \mathbb{R}^n$是一个线性变换。

8. (17.1.10 矩阵) $m\times n$**矩阵**是具有如下形式的对象A：

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

   也可以简写为：

$$A=(a_{ij})_{1\leq i\leq m;1\leq j\leq n}$$

   其中对每一个$1\leq i\leq m$与$1\leq j\leq n$，a_{ij}都是一个实数。因此n维行向量就是一个$1\times n$矩阵，n维列向量就是一个$n\times 1$矩阵。
9. (17.1.11 矩阵的乘积) 给定一个$m\times n$矩阵A和一个$n\times p$矩阵B，我们可以把**矩阵乘积**AB定义为下面这个$m\times p$矩阵：

$$(a_{ij})_{1\leq i\leq m;1\leq j\leq n}(b_{jk})_{1\leq j\leq n;1\leq k\leq p}:=\left(\sum_{j=1}^{n}a_{ij}b_{jk}\right)_{1\leq i\leq m;1\leq k\leq p}$$

   特别地，如果$x^T=(x_i)^T_{1\leq i\leq n}$是一个$n$维列向量，并且$A=(a_{ij})_{1\leq i\leq m;1\leq j\leq n}$是一个$m\times n$矩阵，那么$Ax^T$就是一个$m$维列向量：

$$Ax^T=\left(\sum_{j=1}^{n}a_{ij}x_j\right)^T_{1\leq i\leq m}$$

   借此我们可以将矩阵和线性变换联系起来。如果A是一个$m\times n$矩阵，那么我们可以把变换$L_A:\mathbb{R}^n\to\mathbb{R}^m$定义为：

$$(L_Ax)^T:=Ax^T$$

   (注：这揭示了每一个矩阵都对应者一个线性变换)

命题

1. (17.1.2 $\mathbb{R}^n$是一个向量空间) 设x,y,z都是$\mathbb{R}^n$中的向量，并设c,d是实数。那么我们有加法交换律：$x+y=y+x$；加法结合律：$(x+y)+z=x+(y+z)$；加法恒等性：$x+0=0+x=x$；加法逆元性：$x+(-x)=(-x)+x=0$；乘法结合律：$(cd)x=c(dx)$；分配律：$c(x+y)=cx+cy$和$(c+d)x=cx+dx$；乘法恒等性：$1x=x$。
2. (17.1.13) 设$T:\mathbb{R}^n\to\mathbb{R}^m$是一个线性变换，那么恰好存在一个$m\times n$矩阵$A$使得$T=L_A$。

   (注：这揭示了每一个线性变换都对应者一个矩阵，这样便建立了线性变换与矩阵之间的一一对应关系；如果$T=L_A$，那么A有时被称为T的**矩阵表示**，并记有$A=[T]$（但是本书不用这个记号）)
3. (17.1.16) 设A是一个$m\times n$矩阵，B是一个$n\times p$矩阵，那么$L_AL_B=L_{AB}$。

课后习题

17.1.1 证明引理17.1.2

> 我们设$x=(x_i)_{1\leq i\leq n}$，$y=(y_i)_{1\leq i\leq n}$与$z=(z_i)_{1\leq i\leq n}$。然后逐条证明：
>
> ---
>
> 1. 加法交换律：$x+y=y+x$。
>
> 根据定义我们有：
>
> $$x+y=(x_i+y_i)_{1\leq i\leq n}=(y_i+x_i)_{1\leq i\leq n}=y+x$$
>
> 于是结论得证。
>
> ---
>
> 2. 加法结合律：$(x+y)+z=x+(y+z)$。
>
> 根据定义我们有：
>
> $$\begin{aligned}(x+y)+z&=(x_i+y_i)_{1\leq i\leq n}+(z_i)_{1\leq i\leq n}\\&=((x_i+y_i)+z_i)_{1\leq i\leq n}\\&=(x_i+(y_i+z_i))_{1\leq i\leq n}\\&=(x_i)_{1\leq i\leq n}+(y_i+z_i)_{1\leq i\leq n}=x+(y+z)\end{aligned}$$
>
> 于是结论得证。
>
> ---
>
> 3. 加法恒等性：$x+0=0+x=x$。
>
> 根据定义我们有：
>
> $$x+0=(x_i+0)_{1\leq i\leq n}=(x_i)_{1\leq i\leq n}=x$$
>
> 利用加法交换律即有$x+0=0+x$，于是结论得证。
>
> ---
>
> 4. 加法逆元性：$x+(-x)=(-x)+x=0$。
>
> 根据定义我们有：
>
> $$x+(-x)=(x_i+(-x_i))_{1\leq i\leq n}=(0)_{1\leq i\leq n}=0$$
>
> 利用加法交换律即有$x+(-x)=(-x)+x$，于是结论得证。
>
> ---
>
> 5. 乘法结合律：$(cd)x=c(dx)$；
>
> 根据定义我们有：
>
> $$\begin{aligned}(cd)x&=((cd)x_i)_{1\leq i\leq n}\\&=(c(dx_i))_{1\leq i\leq n}\\&=c(dx_i)_{1\leq i\leq n}=c(dx)\end{aligned}$$
>
> 于是结论得证。
>
> ---
>
> 6. 分配律：$c(x+y)=cx+cy$和$(c+d)x=cx+dx$；
>
> 根据定义我们有：

$$
\begin{aligned}
c(x+y) &= c(x_i+y_i)_{1\le i\le n} \\
&= (c(x_i+y_i))_{1\le i\le n} \\
&= (cx_i+cy_i)_{1\le i\le n} \\
&= (cx_i)_{1\le i\le n}+(cy_i)_{1\le i\le n} = cx+cy
\end{aligned}
$$

$$
\begin{aligned}
(c+d)x &= ((c+d)x_i)_{1\le i\le n} \\
&= (cx_i+dx_i)_{1\le i\le n} \\
&= (cx_i)_{1\le i\le n}+(dx_i)_{1\le i\le n} = cx+dx
\end{aligned}
$$

于是结论得证。

---

7. 乘法恒等性：$1x=x$。

根据定义我们有：

$$
\begin{aligned}
1x &= (1\cdot x_i)_{1\le i\le n} \\
&= (x_i)_{1\le i\le n} = x
\end{aligned}
$$

于是结论得证。

17.1.2 设$T:\mathbb{R}^n\to\mathbb{R}^m$是一个线性变换，并且$S:\mathbb{R}^p\to\mathbb{R}^n$也是一个线性变换。定义$T$和$S$的复合$TS:\mathbb{R}^p\to\mathbb{R}^m$为$TS(x):=T(S(x))$。试证明：这两个变换的复合$TS$也是一个线性变换（提示：通过使用大量的括号，小心地展开$TS(x+y)$和$TS(cx)$）

根据线性变换的定义，对任意的$x,y\in\mathbb{R}^p$与任意的$c\in\mathbb{R}$，应当有：

$$
TS(x+y)=T(S(x+y))=T(S(x)+S(y))=T(S(x))+T(S(y))=TS(x)+TS(y)
$$
$$
TS(cx)=T(S(cx))=T(cS(x))=cT(S(x))=cTS(x)
$$

于是根据线性变换的定义（定义17.1.6）我们有TS也是一个线性变换，结论得证。

17.1.3 证明引理17.1.16

设$A=(a_{ij})_{1\le i\le m;1\le j\le n}$与$B=(b_{jk})_{1\le j\le n;1\le k\le p}$。考虑任意的$p$维行向量$x=(x_i)_{1\le k\le p}$，根据定义我们有：

$$
\begin{aligned}
(L_B(x))^T &= Bx^T \\
&= \left(\sum_{k=1}^{p} b_{jk}x_k\right)^T_{1\le j\le n} \iff L_B(x)=\left(\sum_{k=1}^{p} b_{jk}x_k\right)_{1\le j\le n}
\end{aligned}
$$

$$
\begin{aligned}
(L_A(L_B(x)))^T &= AL_B(x)^T \\
&= \left(\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}x_k\right)\right)^T_{1\le i\le m} \iff L_AL_B(x)=\left(\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}x_k\right)\right)_{1\le i\le m}
\end{aligned}
$$

同样，根据定义有：

$$
\begin{aligned}
(L_{AB}(x))^T &= (AB)x^T \\
&= \left(\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)x_k\right)^T_{1\le i\le m} \iff L_{AB}(x)=\left(\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)x_k\right)_{1\le i\le m}
\end{aligned}
$$

注意到根据有限级数的富比尼定理（定理7.1.14），对任意的$1\le i\le m$都应该有：

$$
\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)x_k=\sum_{k=1}^{p}\sum_{j=1}^{n} a_{ij}b_{jk}x_k=\sum_{j=1}^{n}\sum_{k=1}^{p} a_{ij}b_{jk}x_k=\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}x_k\right)
$$

> 因此即有$L_A L_B(x) = L_{AB}(x)$对任意的$x \in \mathbb{R}^p$都成立，从而引理17.1.16得证。

17.1.4 设$T: \mathbb{R}^n \to \mathbb{R}^m$是一个线性变换。证明：存在一个数$M > 0$，使得对于所有的$x \in \mathbb{R}^n$都有$\|Tx\| \leq M\|x\|$（提示：根据引理17.1.13，用矩阵$A$来表述$T$。然后让$M$等于$A$的所有元素的绝对值之和；多使用三角不等式，它要比处理平方根之类的事情要容易）进而推导出从$\mathbb{R}^n$到$\mathbb{R}^m$的每一个线性变换都是连续的

> 我们先证明第一个结论，我们知道$\|Tx\| \leq M\|x\|$当且仅当$\|Tx\|^2 \leq M^2\|x\|^2$，于是我们只需要讨论平方的情况即可（这要方便一点）。
>
> 根据引理17.1.13，我们知道存在一个$m \times n$矩阵$A = (a_{ij})_{1 \leq i \leq m; 1 \leq j \leq n}$使得$T = L_A$，从而对任意的$x = (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n$有：
>
> $$Tx = L_A x = \left(\sum_{j=1}^{n} a_{ij} x_j\right)_{1 \leq i \leq m}$$
>
> 于是有：
>
> $$\|Tx\|^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right)^2$$
>
> 注意到：
>
> $$\left(\sum_{j=1}^{n} a_{ij} x_j\right)^2 \leq \left(\sum_{j=1}^{n} |a_{ij} x_j|\right)^2 \leq n\sum_{j=1}^{n} a_{ij}^2 x_j^2$$
>
> （利用出租车度量与欧几里得度量的大小关系（例12.1.7））
>
> 于是利用有限级数的富比尼定理我们有：
>
> $$\begin{aligned}\|Tx\|^2 &\leq n\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 x_j^2 \\ &= n\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}^2\right) x_j^2\end{aligned}$$
>
> 于是我们令$M := \sqrt{n\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2}$，则此时有：
>
> $$\begin{aligned}M^2\|x\|^2 &= n\sum_{j=1}^{n}\left(\sum_{k=1}^{n}\sum_{i=1}^{m} a_{ik}^2\right) x_j^2 \\ &\geq n\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}^2\right) x_j^2 \\ &\geq \|Tx\|^2\end{aligned}$$
>
> 于是第一个结论得证。
>
> 然后我们证明第二个结论，对任意的线性变换$T: \mathbb{R}^n \to \mathbb{R}^m$，根据上面的结论我们知道存在一个对应的$M > 0$。然后考虑任意的$\varepsilon > 0$与任意的$x_0 \in \mathbb{R}^n$，我们取$\delta := \frac{\varepsilon}{M}$，于是对任意的$x \in \mathbb{R}^n$满足$d_{l^2}(x, x_0) < \delta$，我们有：

> $$
> \begin{aligned}
> d_{l^2}(Tx, Tx_0) &= \|Tx - Tx_0\| \\
> &= \|T(x - x_0)\| \\
> &\leq M\|x - x_0\| \\
> &< M\delta = \varepsilon
> \end{aligned}
> $$
>
> 于是即T在x_0处连续，由于x_0任意因此也即T是$\mathbb{R}^n$上的连续函数，第二个结论得证。