

doi:10.13582/j.cnki.1674-5884.2017.11.005

线性变换与特征向量的几何化教学探索

陈志彬, 张爱平, 王学斌

(湖南工业大学 理学院应用数学系, 湖南 株洲 412000)

摘要:在低维空间中用几何图形描述线性变换具有的特性及特征向量在线性变换中具有的不变性,引导学生将研究低维空间的方法向高维空间推广,获得高维空间中研究特征向量数形结合的方法,让学生的思维由形象思维过渡到抽象思维,加深学生对代数中抽象概念的理解,在教学中开展研究性教学,探索线性代数中几何化教学的途径。

关键词:线性变换;特征根;特征向量;几何化

中图分类号:O151.2

文献标志码:A

文章编号:1674-5884(2017)11-0020-05

线性代数是工科院校的基础理论课,是为学生学习专业课和将来从事专业工作打下必要基础的课程,也是培养学生思维能力和分析问题解决问题能力的一门学科,它属于数学中的代数系列。在其内容描述上,线性代数通常有三种模式:抽象模式、代数模式和几何模式。若模式不同,则使用的语言也不同,抽象模式使用形式语言,代数模式使用代数语言,几何模式使用几何语言。三种语言的发展由三种思维形式主导,综合几何思维形式用于几何模式,解析算法思维形式用于代数模式,解析结构思维形式用于抽象模式。因此,这三种模式语言及其转换让学生觉得线性代数内容具有高度抽象性,使得抽象的内容与个人的知识结构联系不上,在短时间内较难适应线性代数课程的教学。教学实践表明:线性代数课程的教学难点集中在概念的抽象上,这些抽象的概念多数具有直观的几何背景,不足的是传统的线性代数教材^[1,2]中几乎没有提供直观几何背景知识的感性材料,将代数与几何很好地联系起来,做到数形结合。因此,这使得初学者在认识上难以将新知识与已学过的知识发生联系,有效地产生知识迁移,反而因基本概念难理解常常产生畏惧心理,学习热情受挫,丧失兴趣,给学习线性代数造成障碍。

为了解决教学中的这些问题,教学的最有效办法便是由低维空间到高维空间,将几何化教学融入“线性代数”课程的教学。其方法首先是从代数到几何,结合线性代数教材中已有的教学内容,挖掘抽象概念中的几何原型,赋予抽象概念的几何意义,将运算、变换的过程转化成对应几何图形的变化,建立初学者对线性代数的感性认识。其次,由几何到代数,利用代数的方法针对性地处理一些复杂的几何问题,让学生感受代数处理几何问题的方便与简洁。这种由形象到抽象、再由抽象到代数的“形象与抽象”相结合的教学形式,有助于学生加深对知识的理解,增强几何直观分析问题的能力,促进思维的发展,有效化解线性代数课程知识体系中具有的抽象性,减少恐惧心理,提高学习兴趣。

1 线性变换几何化

线性变换是研究线性空间中元素之间的最基本联系,是线性空间上的一种自映射。线性空间上的一个自映射被称为它的一个变换或算子,但一个变换不一定就是一个线性变换,只有这个变换在线性空间上对线性加法运算封闭时才是线性变换。为便于理解,下面先给出相关概念的定义:

定义 1.1^[3] 设 V 是数域 K 上的线性空间, T 是 V 的自映射,使对于 V 中的任意向量 x 都有 V 中的

收稿日期:20170909

基金项目:湖南省普通高等教育教学改革研究项目(228)

作者简介:陈志彬(1965-),男,湖南邵阳人,教授,硕士,主要从事泛函微分方程定性理论与稳定性理论研究。

唯一向量 y 与之对应,则称 T 是 V 的一个变换或算子,通常记为 $Tx = y$,称 y 为 x 在 T 下的象, x 是 y 的原象。

定义 1.2^[3] 如果数域 K 上的线性空间 V 的一个变换或算子 T 具有性质

$$T(kx + ly) = kT(x) + lT(y),$$

其中 $x, y \in V, k, l \in K$,则称 T 是 V 的一个线性变换或线性算子。

定义 1.3^[3] 设 T 是线性空间 V 的线性变换, V 中所有向量的象构成的集合,称为 T 的值域,用 $R(T)$ 表示,即:

$$R(T) = \{Tx \mid x \in V\};$$

V 中所有被 T 变为零向量的原象构成的集合,称为 T 的核,用 $N(T)$ 表示,即

$$N(T) = \{x \mid Tx = 0, x \in V\}.$$

这个知识点呈现给学生的问题是要确定线性变换 T ,如果有确定它的方法,引导学生探究线性变换 T 的结构,这无疑会有助于学生对线性变换概念的理解,把抽象的定义用代数算式表达。

下面以有限维线性空间上的线性变换为例探究线性变换的结构。首先,将向量用基组向量的坐标表示出来,再通过坐标把线性变换用矩阵表示,由矩阵具有的几何背景将线性变换几何化,让学生直观地理解线性变换。在有限维线性空间上,任意向量都可以用它的基向量组作唯一线性表示,根据线性变换 T 确定基向量组的象,则能获得 V 中任意向量在 T 下的象,即获得线性变换矩阵运算形式的等式。

设 T 是 n 维线性空间 V^n 的线性变换,向量 $x \in V^n$ 且 e_1, e_2, \dots, e_n 是 V^n 的一个基,则有 $x = x_1e_1 + x_2e_2 + \dots + x_n e_n = (e_1, e_2, \dots, e_n)[x_1, x_2, \dots, x_n]^T$,

$$Tx = x_1(Te_1) + x_2(Te_2) + \dots + x_n(Te_n), \quad (2.2)$$

其中 $[x_1, x_2, \dots, x_n]^T$ 为向量 x 关于基向量组 e_1, e_2, \dots, e_n 的坐标列向量,同时由(2.2)式得, V^n 中的任意向量 x 的象 Tx 由基象向量组 Te_1, Te_2, \dots, Te_n 唯一确定,再根据线性变换 T 是 V^n 上的自映射得基象向量仍属于 V^n ,于是基象向量组被基向量组 e_1, e_2, \dots, e_n 唯一线性表示,故令

$$\begin{cases} Te_1 = a_{11}e_1 + a_{21}e_2 + \dots + a_{n1}e_n \\ Te_2 = a_{12}e_1 + a_{22}e_2 + \dots + a_{n2}e_n \\ \vdots \\ Te_n = a_{1n}e_1 + a_{2n}e_2 + \dots + a_{nn}e_n, \end{cases} \quad (2.3)$$

其中 $[a_{1i}, a_{2i}, \dots, a_{ni}]^T \in R^n (i = 1, 2, \dots, n)$ 为基象 Te_i 关于基向量组 e_1, e_2, \dots, e_n 的坐标列向量,通过矩阵运算形式,将(2.3)式简单表示为

$$(Te_1, Te_2, \dots, Te_n) = (e_1, e_2, \dots, e_n)A, \quad (2.4)$$

其中 $A = (a_{ij})_{n \times n}$ 为 n 阶方阵,教材中^[1]将矩阵 A 称为线性变换 T 关于基向量组 e_1, e_2, \dots, e_n 的矩阵。结合(2.1), (2.2)和(2.4)式,以基向量组、方阵 A 和 x 的坐标列向量的矩阵运算形式,得到象 Tx 的代数等式如下:

$$Tx = (Te_1, Te_2, \dots, Te_n)[x_1, x_2, \dots, x_n]^T = (e_1, e_2, \dots, e_n)A[x_1, x_2, \dots, x_n]^T. \quad (2.5)$$

在(2.5)式中矩阵 A 的列向量是基象 $Te_i (i = 1, 2, \dots, n)$ 关于基向量组 e_1, e_2, \dots, e_n 的坐标列向量, $A[x_1, x_2, \dots, x_n]^T$ 是向量 x 的象 Tx 关于基 e_1, e_2, \dots, e_n 向量组的坐标列向量。(2.5)式是线性变换 T 最关键的等式,当基向量组被确定时,它将定义 1.1 抽象的线性变换 T 表成了一个具体的以矩阵运算形式相乘的代数等式。由(2.5)式可知:在理解线性变换 Tx 的代数等式时,应该结合选定的基向量组理解矩阵 A 的变化,方阵 A 由 V^n 中选定的基向量组唯一确定,若基向量组改变,则方阵 A 亦改变,但同一个线性变换在不同基向量组下的矩阵存在着一种相似关系,这是一种等价关系,因此,一个线性变换对应着一类相似矩阵。在线性变换(2.5)式中,还隐含着 R^n 到 R^n 的一种坐标之间的线性对应关系,即原象的坐标对应着象的坐标,它是 R^n 上的线性变换,用如下的代数等式表示:

$$\begin{aligned} T^* : [x_1, x_2, \dots, x_n]^T &\mapsto A[x_1, x_2, \dots, x_n]^T, \\ T^* x &= Ax, \end{aligned} \quad (2.6)$$

在教材中常以矩阵(2.6)式的形式作为线性变换教学的主要内容,事实上它是伴随一个具体的线性变换在选定基向量组 e_1, e_2, \dots, e_n 时,在 R^n 上原象的坐标与象的坐标相对应的变换;反之,若知道形如(2.6)式在 R^n 上的一个坐标对应关系式 T^* ,则在已知 V^n 基向量组 e_1, e_2, \dots, e_n 的条件下,亦可以反推得到 V^n 上的线性变换 T 。

例1 设 $x = [x_1, x_2, x_3]^T \in R^3$ 且线性变换 $Tx = [x_1 + x_2 - 3x_3, 3x_1 - x_2 - 3x_3, 0]^T$, 求下列基向量组矩阵运算形式的线性变换 T

1) 已知基向量组 $e_1 = [1, 0, 0]^T, e_2 = [0, 1, 0]^T, e_3 = [0, 0, 1]^T$;

2) 已知基向量组 $\tilde{e}_1 = [1, 0, 0]^T, \tilde{e}_2 = [1, 1, 0]^T, \tilde{e}_3 = [1, 1, 1]^T$ 。

解 1) 由 $Tx = [x_1 + x_2 - 3x_3, 3x_1 - x_2 - 3x_3, 0]^T$ 得基向量组的象

$$Te_1 = [1, 3, 0]^T = e_1 + 3e_2, Te_2 = [1, -1, 0]^T = e_1 - e_2, Te_3 = [-3, -3, 0]^T = -3e_1 - 3e_2$$

$$(Te_1, Te_2, Te_3) = (e_1, e_2, e_3) \begin{bmatrix} 1 & 1 & -3 \\ 3 & -1 & -3 \\ 0 & 0 & 0 \end{bmatrix}, \text{其中矩阵 } A = \begin{bmatrix} 1 & 1 & -3 \\ 3 & -1 & -3 \\ 0 & 0 & 0 \end{bmatrix}.$$

对于任意给定的向量 $x = [x_1, x_2, x_3]^T \in R^3$, 用矩阵运算形式得到线性变换 T :

$$Tx = (e_1, e_2, e_3)A[x_1, x_2, x_3]^T \quad (2.7)$$

2) 同问题1)的方法推得基向量组的象

$$(T\tilde{e}_1, T\tilde{e}_2, T\tilde{e}_3) = (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3) \begin{bmatrix} 1 & 2 & -1 \\ 3 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \text{其中矩阵 } \tilde{A} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\begin{aligned} Tx &= T[(x_1 - x_2)\tilde{e}_1 + (x_2 - x_3)\tilde{e}_2 + x_3\tilde{e}_3] = (T\tilde{e}_1, T\tilde{e}_2, T\tilde{e}_3) \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3)\tilde{A} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned} \quad (2.8)$$

针对实例中的问题1)与问题2),根据图形变化及代数等式的表示引导学生从如下两方面理解线性变换的概念:

一方面,(2.7)与(2.8)式表明:同一线性变换 T 通过矩阵运算表示成了两个不同的代数等式。因此,在线性空间中对于同一个线性变换,当选取的基向量组不同时,得到的代数等式一般也不同,但线性变换的两个基向量组的矩阵 A 与 \tilde{A} 彼此相似,所求关系式的过程演示如下:

$$Tx = (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3)\tilde{A} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (e_1, e_2, e_3)A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ 且 } (\tilde{e}_1, \tilde{e}_2, \tilde{e}_3) = (e_1, e_2, e_3) \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\text{令 } C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \text{ 得矩阵 } A \text{ 与 } \tilde{A} \text{ 的关系式 } A = C\tilde{A}C^{-1} \text{ 或 } \tilde{A} = C^{-1}AC.$$

推广到一般有结论:同一线性变换不同基向量组的矩阵彼此都相似。

另一方面,将两个线性变换 $T_1^*x = Ax$ 与 $T_2^*x = \tilde{A}C^{-1}x$ 置于空间直角坐标系中,用图形描述原象与象的变化得知:其象坐标的向量 T_1^*x 与 T_2^*x 由原象的坐标向量 $x = [x_1, x_2, x_3]^T$ 绕坐标原点在空间经过旋转且伸缩得到,且都落在 xoy 直角坐标平面上,所得到的象空间为二维平面向量空间,维数等于矩阵 A 或 $\tilde{A}C^{-1}$ 的秩;结合线性空间的性质若用代数等式表示,则象 T_1^*x 与 T_2^*x 相应用矩阵 A 和矩阵 $\tilde{A}C^{-1}$ 的列向量作线性表示,象空间的结构由这两个矩阵的列向量确定,矩阵的列向量极大线性无关组的个数等于象空间的维数,与矩阵的秩相等。

以上实例针对低维的三维线性空间,对象空间具有的结构特点进行了分析,在教学中,如果将低维空间具有类似的这些特点,逐步地推广到高维空间,这将有助于学生理解定义 1.2 与定义 1.3。因此,在任意有限维向量空间中如下描述两个集合 $R(T)$ 及 $N(T)$ 的代数与空间的几何意义,学生也就不难理解了:

如果线性变换 T 的基向量组矩阵为 A , 则 $R(T)$ 中象的坐标列向量由原象的坐标列向量绕坐标原点在空间旋转经过伸缩得到,并落在由矩阵 A 的列向量组所张成的多面体中;象集合 $R(T) = \{Tx | x \in V\}$ 对线性运算是封闭的,其维数等于矩阵 A 的秩,是 V 的线性子空间;若 $R(T)$ 中的向量 Tx 用坐标的代数等式表示,则它的坐标列向量是矩阵 A 的列向量组的线性组合;线性变换 T 的核 $N(T) = \{x | Tx = 0, x \in V\}$ 对线性运算同样也是封闭的,是 V 的线性子空间,其维数等于齐次线性方程组 $Ax = 0$ 解空间 $\{x | Ax = 0, x \in K\}$ 的维数,且关于维数有一个重要结论:原象空间 V 的维数等于象子空间 $R(T)$ 维数与核子空间 $N(T)$ 维数之和。

2 特征向量几何化

线性空间 V 中有一些非零向量 x 经过线性变换 T , 得到的象向量 Tx 与原象向量是线性相关的,这些向量与零向量构成的集合是 V 的子空间,具有某些共同的特征,若用几何化的方法将这些特征呈现给学生,抽象的代数等式则转变成为向量几何图形,线性变换对应着几何图形中原象向量与象向量的变化。

下面首先定义具有此种特征的向量,尔后通过实例说明数形结合,再在空间中描述这些向量的特性。这将有利于促进学生思维的发展,掌握好线性变换中的不变量。

定义 2.1^[3] 设 T 是数域 K 上的线性空间 V^n 的线性变换,且对 K 中某一数 λ , 存在非零向量 $x \in V^n$, 使得

$$Tx = \lambda x \quad (2.9)$$

成立,则称数 λ 为 T 的特征值, x 为 T 的属于 λ 的特征向量。

由线性变换的定义及(2.9)式表明:如果 x 是 T 的属于特征值 λ 的特征向量,则任意数 $k \neq 0$ 与 x 的乘积 kx 也是属于特征值 λ 的线性变换 T 的特征向量,即等式 $T(kx) = \lambda(kx)$ 成立;如果矩阵 A 是线性变换 T 关于基向量组 e_1, e_2, \dots, e_n 的矩阵,当属于特征值 λ 的特征向量 x 用坐标形式表成 $x = (e_1, e_2, \dots, e_n)[x_1, x_2, \dots, x_n]^T$ 时,由(2.9)式易得向量 x 的坐标等式

$$(\lambda E - A)[x_1, x_2, \dots, x_n]^T = 0, \quad (2.10)$$

其中 E 是 n 阶单位矩阵,(2.10)式是一个齐次线性方程组。根据特征向量 $x \neq 0$, 则有结论:

齐次线性方程组(2.10)式有非零解,当且仅当行列式 $\det(\lambda_0 E - A) = 0$ 。

行列式 $\det(\lambda_0 E - A)$ 是 λ 的一个 n 次多项式,在教材中被定义为矩阵 A 的特征多项式,方程 $\det(\lambda E - A) = 0$ 被定义为矩阵 A 的特征方程, x 的坐标向量 $[x_1, x_2, \dots, x_n]^T$ 称为矩阵 A 属于 λ_0 的特征向量。

定义 2.2^[3] 设 T 是线性空间 V^n 的线性变换, λ 是 T 的一个特征值,则 V^n 的子空间

$$V_\lambda = \{x | Tx = \lambda x, x \in V^n\} \quad (2.11)$$

称为 T 的属于 λ 的特征子空间。

例 2 设 T 是线性空间 V^3 的线性变换,已知 T 关于基向量组 e_1, e_2, e_3 的矩阵 A 是

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},$$

求 T 的特征值、特征向量和特征子空间。

解 易求出矩阵 A 的特征多项式

$$\det(\lambda E - A) = (\lambda + 1)^2(\lambda - 5),$$

于是,得到线性变换 T 的特征值 $\lambda_1 = -1$ (二重特征值), $\lambda_2 = 5$.

当特征值 $\lambda_1 = -1$ 时,齐次线性方程组 $(\lambda_1 E - A)[x_1, x_2, x_3]^T = 0$ 的一个基础解系为 $[1, 0, -1]^T$, $[0, 1, -1]^T$, 这两个向量是矩阵 A 的特征向量,矩阵 A 属于 λ_1 的全体特征向量表示为 $k[1, 0, -1]^T + \bar{k}[0, 1, -1]^T$ ($k, \bar{k} \in R$ 且不同时为零)。根据 T 与 A 的特征向量之关系,线性变换 T 属于 λ_1 的两个线性无关的特征向量为

$$y = (e_1, e_2, e_3)[1, 0, -1]^T, \tilde{y} = (e_1, e_2, e_3)[0, 1, -1]^T。$$

线性变换 T 属于 λ_1 的全体特征向量表示为

$$ky + \bar{k}\tilde{y} \quad (k, \bar{k} \in R \text{ 且不同时为零})。 \quad (2.12)$$

λ_1 的特征子空间是一个二维向量空间,可表示为

$$V_{\lambda_1} = \{x | Tx = \lambda_1 x, x \in V^n\}。 \quad (2.13)$$

若将(2.12)式用 V^3 的基向量组 e_1, e_2, e_3 作线性表示,则由(2.11)式可得特征子空间

$$V_{\lambda_1} = \{k(e_1 - e_3) + \bar{k}(e_2 - e_3) | k, \bar{k} \in R\}, \quad (2.14)$$

当 $\lambda_2 = 5$ 时,齐次线性方程组 $(\lambda_2 E - A)[x_1, x_2, x_3]^T = 0$ 的一个基础解系为 $[1, 1, 1]^T$, 矩阵 A 属于 λ_2 的全体特征向量表示为 $k[1, 1, 1]^T$ ($k \in R$ 且不为零)。同理,可得线性变换 T 属于 λ_2 的全体特征向量

$$\bar{y} = k(e_1, e_2, e_3)[1, 1, 1]^T \quad (k \in R \text{ 且不为零}) \quad (2.15)$$

λ_2 的特征子空间为

$$V_{\lambda_2} = \{k\bar{y} | k \in R\} = \{k(e_1 + e_2 + e_3) | k \in R\}。 \quad (2.16)$$

例2中(2.12)与(2.13)两式表明:从几何图形观察,特征子空间 V_{λ_1} 由 λ_1 的全体特征向量与零向量构成,特征向量都落在向量 y 与 \tilde{y} 所张成的二维平面上,而落在 y 与 \tilde{y} 所张成平面上的非零向量都是 T 的属于特征值 λ_1 的特征向量;属于特征子空间 V_{λ_1} 中的向量 x 与象向量 Tx 共线,象向量 Tx 由原象向量 x 反向拉伸得到,向量的模长彼此相等。

以上的实例,详尽地探讨了三维线性空间中线性变换 T 及其基向量组矩阵 A 的特征值与特征向量的问题,以直观感性的几何变化建立学生对低维空间中特征向量的理解与认识,让学生正确理解 T 的特征向量与矩阵 A 的特征向量之间既有密切联系但又有区别。 T 的特征向量是线性变换 T 象空间中的量,而 A 的特征向量则是 T 的原象向量关于基组的坐标向量在坐标空间中的量,而在各自的空间中,两者的几何位置是一一对应的。因此,由定义2.1表述的特征向量所具有的代数与几何意义,在三维空间与高维空间中是类似的,用几何语言作如下描述:

n 维向量空间中的线性变换 T ,在确定基向量组后,它对应着一个方阵,同一线性变换下不同基向量组的矩阵彼此相似;线性变换的过程就是把原象空间中的任意一个向量与对应的坐标向量在各自的空间中进行旋转、伸缩,得到方向或长度多数不同的新向量;如果线性变换与它对应的矩阵分别对某些原象及原象的坐标向量只进行同向(特征值为正)或反向(特征值为负)或变成零向量(特征值为零)的伸缩变换,而无旋转变化,则这些向量相应的就是线性变换的特征向量及它所对应矩阵的特征向量,对于非零特征值的特征向量,其伸缩比等于特征值的绝对值;特征向量的不变性表现为原象向量与象向量共线即是线性相关的,特征向量所在的直线及特征子空间在线性变换下保持不变。

参考文献:

- [1] 周勇,朱砾. 线性代数[M]. 上海:复旦大学出版社,2010.
- [2] 陈志杰. 高等代数与几何[M]. 北京:高等教育出版社,2008.
- [3] 程云鹏,张凯院,徐仲. 矩阵论[M]. 西安:西北工业大学出版社,2001(2):29-35.
- [4] 李尚志. 从问题出发引入线性代数概念[J]. 高等数学研究,2006(6):15-17.

(责任校对 莫秀珍)