矩阵初等变换的一些探讨与教学体会

姜德烁

百色学院 数理科学与统计学院，广西 百色 533000

摘　　要： 在日常课程教学的基础上，结合具体实例对矩阵的初等变换在逆矩阵求解、向量组理论、线性方程组与特征向量求解、矩阵对角化及求方阵幂等线性代数问题中的应用进行了较为详细的分析与阐述，给出了教学中的一些心得体会，以便初涉该课程的读者对此有一个较好的了解与认识，藉此更好地学习和掌握线性代数这门课程。从课程思政角度探讨了初等变换在信息编码、保密计算等现代科学问题中的应用，以便为课堂教学提供一些可供参考的素材。

关 键 词： 初等变换；逆矩阵；向量组的秩；线性方程组；特征向量；信息编码；保密计算

Some Discussion and Teaching Experience of Elementary Transformation of Matrix

Jiang Desuo

School of Mathematical Science and Statistics, Baise University, Baise, Guangxi 533000

Abstract: On the basis of daily course teaching, combined with concrete examples, the application of elementary transformation of matrix in inverse matrix solving, vector group theory, linear equations and eigenvector solving, matrix diagonalization and linear algebra problems such as idempotent square matrix are analyzed and expounded in detail, and some teaching experiences are given. In order to have a better understanding and understanding of this course readers, so as to better learn and master the course of linear algebra. In order to provide some reference materials for classroom teaching, this paper discusses the application of elementary transformation in information coding, confidential computing and other modern scientific problems from the perspective of curriculum ideology and politics.

Keywords: elementary transformation; inverse matrix; the rank of a vector set; linear equations; eigenvectors; information coding; security calculation

矩阵的初等变换是线性代数课程中一个重要的内容。虽然概念本身较为简单，但应用却非常广泛，涉及到线性代数课程中的大多数知识点。对该部分内容做一个较为系统的梳理与总结，并融入到平时的教学中，相信会起到很好的促进作用，同时也能让学生深刻认识到知识间的密切联系，并极大激发其学习的兴趣和动力。本文，我们从逆矩阵求解、向量组理论、线性方程组与特征向量求解、矩阵对角化等方面对初等变换的应用进行了较为详细的归纳探讨，给出了教学中的一些心得体会。最后，考察了初等变换在信息编码、保密计算等现代科学问题中的应用。

一、相关概念

设 A 为 $m\times n$ 型矩阵。本文中，我们采用如下表述方式：r_i 表示矩阵的行，即第 i 行；c_j 表示矩阵的行，即第 j 行。同时用 E 表示单位矩阵。

（一）初等变换

矩阵的初等变换有初等行变换和初等列变换之分。初等行变换有如下三种情形[1,2]：

（1）交换矩阵中两行（如第 i 行和第 j 行交换，则记为 $r_i \leftrightarrow r_j$）；

（2）矩阵的某一行乘以一个不为0的常数 k（如 kr_i 表示第 i 行中的每个元素都乘以 k）；

（3）矩阵中的某一行乘以一个不为0的常数加到矩阵的另一行（如 r_j+kr_i 表示第 i 行中的每个元素都乘以 k 加到第 j 行的对应元素上）。

类似地，初等列变换也有三种情形。

（二）行阶梯形矩阵与行最简形矩阵

若非零矩阵 A 满足条件：

（1）非零行在零行的上面；

（2）非零行的首元（即该行第一个不为0的元素）所在列在上一行（如果存在的话）首元所在列的右面，

则称矩阵 A 为行阶梯形矩阵[1]。

进一步，若行阶梯形矩阵还满足如下条件：

（1）非零行首元为1；

（2）首元所在列的其它元均为0，

则称该矩阵为行最简形矩阵[1]。

基金项目：百色学院2023年校级教学改革工程项目（2023JG70）。

二、在线性代数课程中的应用

（一）逆矩阵求解

若矩阵 A 可逆，则可利用初等行变换求出其逆矩阵。首先，作一新的矩阵 $(A\vdots E)$，对其作初等行变换，使得左边的矩阵 A 化为单位矩阵 E，则右边的那块即为 A 的逆矩阵[2]。如下图所示

$(A\vdots E)\xrightarrow{\text{初等行变换}}(E\vdots A^{-1})$。

例1 求矩阵 $A=\begin{pmatrix}1&1&0\\-1&0&1\\-1&2&-1\end{pmatrix}$ 的逆矩阵。

解 对矩阵 $(A\vdots E)$ 施行初等行变换，使得左边矩阵 A 化为单位矩阵 E，如下

$$(A\vdots E)=\begin{pmatrix}1&1&0&\vdots&1&0&0\\-1&0&1&\vdots&0&1&0\\-1&2&-1&\vdots&0&0&1\end{pmatrix}\overset{\substack{r_2+r_1\\r_3+r_1}}{\sim}\begin{pmatrix}1&1&0&\vdots&1&0&0\\0&1&1&\vdots&1&1&0\\0&3&-1&\vdots&1&0&1\end{pmatrix}$$

$$\overset{r_3-3r_2}{\sim}\begin{pmatrix}1&1&0&\vdots&1&0&0\\0&1&1&\vdots&1&1&0\\0&0&-4&\vdots&-2&-3&1\end{pmatrix}\overset{r_3\times(-\frac{1}{4})}{\sim}\begin{pmatrix}1&1&0&\vdots&1&0&0\\0&1&1&\vdots&1&1&0\\0&0&1&\vdots&1/2&3/4&-1/4\end{pmatrix}$$

$$\overset{r_2-r_3}{\sim}\begin{pmatrix}1&1&0&\vdots&1&0&0\\0&1&0&\vdots&1/2&1/4&1/4\\0&0&1&\vdots&1/2&3/4&-1/4\end{pmatrix}$$

$$\overset{r_1-r_2}{\sim}\begin{pmatrix}1&0&0&\vdots&1/2&-1/4&-1/4\\0&1&0&\vdots&1/2&1/4&1/4\\0&0&1&\vdots&1/2&3/4&-1/4\end{pmatrix}$$

故 A 的逆矩阵为

$$A^{-1}=\begin{pmatrix}1/2&-1/4&-1/4\\1/2&1/4&1/4\\1/2&3/4&-1/4\end{pmatrix}。$$

与其它方法的比较 当矩阵的阶数不太高时，也可以利用伴随矩阵来求其逆矩阵（当矩阵可逆时）。但若矩阵的阶数比较高（大于或等于3时），则伴随矩阵中各代数余子式的阶数也比较高，计算起来其实是相当麻烦的（如 n 阶矩阵，它的每一元素都对应一代数余子式，因此共有 n^2 个，而每一个代数余子式中行列式的阶数都为 $n-1$，全部计算出来需花费相当多时间），因此这种方法也就不太实用，而如果利用初等变换的方法，则可以很容易地求出矩阵的逆矩阵。

（二）在向量组理论中的应用

利用初等行变换，可以求矩阵的秩，进而可以求向量组的秩。由于等价矩阵的秩是相等的，我们可以先利用初等行变换把矩阵化为与其等价的行阶梯形矩阵（或行最简形矩阵），则这两个矩阵的秩是相等的。由于行阶梯形矩阵的秩即为其中非零行的行数[1,2]，故原矩阵的秩也为非零行的行数。如，

$$A=\begin{pmatrix}1&1&0\\-1&0&2\\2&1&3\end{pmatrix}\overset{\substack{r_2+r_1\\r_3-2r_1}}{\sim}\begin{pmatrix}1&1&0\\0&1&2\\0&-1&3\end{pmatrix}\overset{r_3+r_2}{\sim}\begin{pmatrix}1&1&0\\0&1&2\\0&0&5\end{pmatrix}。$$

上例中，矩阵 A 经过两次初等行变换之后化为行阶梯形矩阵，其中非零行的行数为3，故原矩阵的秩也为3。

设有 n 维向量组 $\vec{\alpha}_1$，$\vec{\alpha}_2$，…，$\vec{\alpha}_s$，把每个向量的分量按列写出来，则作成一个 $n\times s$ 型矩阵。对该矩阵作初等行变换，化为行阶梯形矩阵，其中非零行的行数即为该矩阵的秩。由于矩阵的秩也为其列向量组的秩，由此得到向量组 $\vec{\alpha}_1$，$\vec{\alpha}_2$，…，$\vec{\alpha}_s$ 的秩。

例2 已知向量 $\vec{\alpha}_1=(1,3,2,0)$，$\vec{\alpha}_2=(7,0,14,3)$，$\vec{\alpha}_3=(2,-1,0,1)$，$\vec{\alpha}_4=(5,1,6,2)$，$\vec{\alpha}_5=(2,-1,4,1)$，求这个向量组的秩和一个极大无关组，并把其余的向量用这个极大无关组表示出来。

解 设矩阵

$$A=(\vec{\alpha}_1,\vec{\alpha}_2,\vec{\alpha}_3,\vec{\alpha}_4,\vec{\alpha}_5)=\begin{pmatrix}1&7&2&5&2\\3&0&-1&1&-1\\2&14&0&6&4\\0&3&1&2&1\end{pmatrix}.$$

对 A 施行初等行变换，将其化为行最简形矩阵：

$$A\xrightarrow{\text{初等行变换}}\begin{pmatrix}1&0&0&2/3&-1/3\\0&1&0&1/3&1/3\\0&0&1&1&0\\0&0&0&0&0\end{pmatrix}。$$

由于行最简形矩阵中非零行的行数为3，故矩阵 A 的秩也为3，进而向量组 $\vec{\alpha}_1$，$\vec{\alpha}_2$，…，$\vec{\alpha}_5$ 的秩为3。

另外，观察行最简形矩阵，它的第1列、第2列、第3列对应的3个向量是线性无关的，因而矩阵 A 的前3列对应的三个向量 $\vec{\alpha}_1,\vec{\alpha}_2,\vec{\alpha}_3$ 也是线性无关的（初等行变换保持列向量组的线性相关性）。注意到向量组 $\vec{\alpha}_1$，$\vec{\alpha}_2$，…，$\vec{\alpha}_5$ 的秩为3，故这3个向量为该向量组 A 的一个极大无关组。

最后，观察行最简形矩阵中的第四列和第五列，可以得到 $\vec{\alpha}_4,\vec{\alpha}_5$ 的表示式：

$$\vec{\alpha}_4=\frac{2}{3}\vec{\alpha}_1+\frac{1}{3}\vec{\alpha}_2+\vec{\alpha}_3，\quad \alpha_5=-\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2。$$

（三）解线性方程组

对线性方程组，可利用初等变换把其转化为一个同解方程组，进而求出它的通解。首先，写出方程组的系数矩阵或增广矩阵（如果是齐次线性方程组，则写出系数矩阵即可；如果是非齐次线性方程组，则写出其增广矩阵）。对系数矩阵或增广矩阵作初等行变换，化为行最简形矩阵。得到的行最简形矩阵对应一个方程组，该方程组与原来的方程组等价。考察该等价方程组即得原方程组的解[2]。这是求线性方程组的一种有效且常用的方法。如下例

例3 求齐次线性方程组 $\begin{cases}x_1+x_2+x_3+4x_4-3x_5=0\\x_1-x_2+3x_3-2x_4-x_5=0\\2x_1+x_2+3x_3+5x_4-5x_5=0\end{cases}$ 的通解.

解 对系数矩阵作初等行变换，化为行最简形：

$$A=\begin{pmatrix}1&1&1&4&-3\\1&-1&3&-2&-1\\2&1&3&5&-5\end{pmatrix}\xrightarrow{\text{初等行变换}}\begin{pmatrix}1&0&2&1&-2\\0&1&-1&3&-1\\0&0&0&0&0\end{pmatrix}.$$

该最简形矩阵对应一个方程组，为

$$\begin{cases}1\cdot x_1+0\cdot x_2+2x_3+x_4-2x_5=0,\\0\cdot x_1+1\cdot x_2-x_3+3x_4-x_5=0.\end{cases}$$

移项得

$$\begin{cases}x_1=-2x_3-x_4+2x_5,\\x_2=x_3-3x_4+x_5.\end{cases}$$

由于原方程组含有5个未知量 x_1、x_2　x_3　x_4　x_5，为求出该方程组的解，再补充3个方程，得到如下方程组

$$\begin{cases} x_1 = -2x_3 - x_4 + 2x_5, \\ x_2 = x_3 - 3x_4 + x_5, \\ x_3 = x_3, \\ x_4 = x_4, \\ x_5 = x_5. \end{cases}$$

将其写成向量形式，并以 c_1、c_2 c_3 代换右边的未知量 x_3、x_4 x_5，即得原方程组的通解

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = c_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$，（c_1, c_2, c_3 为任意常数）。

（四）求特征向量

由于特征向量的求解离不开解线性方程组，因此，初等变换自然也起着重要的作用。

设 A 为一 n 阶方阵。由特征方程 $|A-\lambda E|=0$ 可求出其特征值[2]，设为 λ_1 、λ_2 、…、λ_n。对特征值 λ_i $(i=1,2,\cdots,n)$，解线性方程组

$$(A-\lambda_i E)\vec{x}=\vec{0}\text{，} \qquad (*)$$

即可求出属于 λ_i 的所有特征向量。这里线性方程组（*）的求解，需要对系数矩阵作初等变换，化为行最简形矩阵。

（五）在矩阵对角化和求方阵幂中的应用

若矩阵 A 对称，则可以对角化[1]，即存在可逆矩阵 P，使得

$$P^{-1}AP=\Lambda \text{（对角矩阵）。}$$

于是，$A=P\Lambda P^{-1}$，从而 $A^n=P\Lambda^n P^{-1}$。这里的对角矩阵 Λ，其主对角线上的元素即为矩阵 A 的特征值，可由特征多项式 $|A-\lambda E|=0$ 求出。而可逆矩阵 P，则通过 A 的特征向量求得。不妨设 n 阶方阵 A 的特征值分别为 λ_1 、λ_2 、…、λ_n，属于每个特征值的线性无关的特征向量分别为 $\vec{p}_1$ 、$\vec{p}_2$ 、…、$\vec{p}_n$，则 $P=(\vec{p}_1,\vec{p}_2,\cdots,\vec{p}_n)$。利用初等行变换，可以求出 P 的逆矩阵 P^{-1}。于是可以求得 A^n。注意到特征向量的求解离不开解线性方程组，故初等变换在这里也起着重要的作用。

另外，初等变换在不等式证明、求多项式最大公因式等方面也有着奇妙的作用，可参阅文献 [3-5] 等。

由上可以看出，初等变换这一概念很好地融合了线性代数课程中的一些重要的知识点。在课程后期或课程结束以后，以课外作业或小论文的形式让学生对此内容进行归纳总结，有助于学生更好地理解并掌握线性代数这门课程。

三、在现代科学问题中的应用

矩阵的初等变换不但在处理线性代数的一些理论问题时起着重要作用，而且在信息编码、故障定位、逻辑电路、保密计算等生产实践领域也是有着广泛的应用。我们对此做了一些介绍，相关内容可作为课程思政的参考素材。

（一）信息编码

在密码学中，利用矩阵理论对信息进行加密和解密是一种重要的、安全性较高的方法，由希尔在1929年首先提出。它的思想[6]是：

首先，26个英文字母与26个阿拉伯数字一一对应（如A对应1，B对应2，C对应3，依此类推），另外用0表示空格。

加密时，将信息中单词的字母从左到右每 n 个字符分为一组，不足 n 个字符的用空格补上（如 academic，转化为编码为“1、3、1、4、5、13、9、3”，每3个字符分为一组，作成3个列向量 $\vec{b}_1=\begin{pmatrix}1\\3\\1\end{pmatrix}$、$\vec{b}_2=\begin{pmatrix}4\\5\\13\end{pmatrix}$、$\vec{b}_3=\begin{pmatrix}9\\3\\0\end{pmatrix}$，进而得到一个编码矩阵，不妨记为 B，即 $B=\begin{pmatrix}1&4&9\\3&5&3\\1&13&0\end{pmatrix}$）。然后，确定一个可逆矩阵 A 作为加密矩阵，用加密矩阵左乘编码矩阵，得到密文矩阵 C（$C=AB$），这就是要发送的矩阵。

解码时，则是上述过程的逆过程。由矩阵等式 $C=AB$ 的两端左乘 A 的逆矩阵 A^{-1}，可求得译码矩阵 $B=A^{-1}C$。将 B 对应回英文字符即可得到发送的信息。在此过程中，一个重要的数学问题是要求加密矩阵的逆矩阵，通常的方法是利用初等行变换。

（二）保密计算

空间两直线位置关系的判定是保密计算中的一个重要问题。而判定两直线的相关位置，一种是向量方法，即通过与两直线相关的一些向量的相互关系来判定两直线的位置关系（这里假定给出的直线方程为标准方程。具体可参阅文献[7]）；另一种则是代数上的，即借助两直线所对应矩阵的秩[8]（假定给出的直线方程为一般方程）。要求矩阵的秩，一个有效且常用的方法就是对矩阵进行初等变换。

设空间两直线 L_1 、L_2 的（一般）方程分别为

$$L_1=\begin{cases} l_{11}: a_{11}x_1+a_{12}x_2+a_{13}x_3=a_{14} \\ l_{12}: a_{21}x_1+a_{22}x_2+a_{23}x_3=a_{24} \end{cases}\text{，}L_2=\begin{cases} l_{21}: b_{11}x_1+b_{12}x_2+b_{13}x_3=b_{14} \\ l_{22}: b_{21}x_1+b_{22}x_2+b_{23}x_3=b_{24} \end{cases}\text{。}$$

由方程可以看出，L_1 实际上是两平面 l_{11} 与 l_{12} 的交线，L_2 则是 l_{21} 与 l_{22} 这两个平面的交线。利用平面方程的参数构造矩阵 A_{12} 及增广矩阵 $\tilde{A}_{12}$：

$$A_{12}=\begin{pmatrix} a_{11} & a_{21} & b_{11} & b_{21} \\ a_{12} & a_{22} & b_{12} & b_{22} \\ a_{13} & a_{23} & b_{13} & b_{23} \end{pmatrix}\text{，}\quad \tilde{A}_{12}=\begin{pmatrix} a_{11} & a_{21} & b_{11} & b_{21} \\ a_{12} & a_{22} & b_{12} & b_{22} \\ a_{13} & a_{23} & b_{13} & b_{23} \\ a_{14} & a_{24} & b_{14} & b_{24} \end{pmatrix}\text{。}$$

则两直线的位置关系可由这两矩阵秩的关系来确定。若 $R(A_{12})=R(\tilde{A}_{12})=3$，则 L_1 与 L_2 相交于一点；若 $R(A_{12})=R(\tilde{A}_{12})=2$，则 L_1 与 L_2 重合；若 $R(A_{12})=2$ 而 $R(\tilde{A}_{12})=3$，则 L_1 与 L_2 平行；若 $R(A_{12})=3$ 而 $R(\tilde{A}_{12})=4$，则 L_1 与 L_2 异面。

类似的问题还有很多，因篇幅原因不再一一介绍，可参阅文献 [9-10] 等。

参考文献

[1] 同济大学数学科学学院. 工程数学线性代数［M］. 北京：高等教育出版社，2023.
[2] 周勇. 线性代数［M］. 北京：北京大学出版社，2019.
[3] 鲁翠仙. Frobenius 不等式的证明方法［J］. 齐齐哈尔大学学报，2016, 32(1): 93-94.
[4] 黄述亮. 关于矩阵秩的几个重要不等式［J］. 辽东学院学报（自然科学版），2021, 28(1): 61-65.
[5] 孙树东. 初等变换求多项式的最大公因式法［J］. 数学学习与研究，2015 (15): 89-90.
[6] 熊允发. 矩阵在信息编码中的应用［J］. 中国人民公安大学学报（自然科学版），2017, (1): 75-78.
[7] 吕林根，许子道. 解析几何［M］. 北京：高等教育出版社，2023.
[8] 杜润萌，刘旭红，李顺东，魏琼. 矩阵与增广矩阵秩相等问题的保密计算及应用［J］. 密码学报，2019, 6(2): 205-218.
[9] 王冬，刘强，李善治. 基于矩阵初等变换的量子可逆逻辑电路双向综合算法［J］. 计算机科学，2014, 41(9): 18-23.
[10] 李小强，王军，张贺，倪明. 一种改进的故障定位矩阵算法［J］. 中国科技论文，2017, 12(23): 2685-2689.