善用初等变换展开线性代数理论教学

欧阳毅 [1]，　张神星 [2]

（1. 中国科学技术大学数学科学学院, 合肥 230026；2. 合肥工业大学数学学院, 合肥 230601）

[**摘　要**] 初等变换是线性代数的基本变换, 在线性代数课程中常常被用来计算, 例如求解线性方程组、计算方阵的行列式、矩阵的求逆以及更一般的矩阵方程 $\mathsf{AX} = \mathsf{B}$ 的求解、计算整数矩阵和域上多项式矩阵的 Smith 标准形、以及计算对称阵的相合标准形等. 本文说明如何灵活利用初等变换, 给出线性代数课程中一些重要理论结果的系统而又简洁的证明.

[**关键词**] 线性代数；初等变换；矩阵运算

[**中图分类号**] O151.2

Theoretical Application of Elementary Transformations in Linear Algebra Teaching

OUYANG Yi[1], ZHANG Shenxing[2]

（1. School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, China;
2. School of Mathematics, Hefei University of Technology, Hefei 230601, China）

Abstract: Elementary transformations are basic operations in linear algebra, where they are often used for computation, for example, to solve linear equations, to compute the determinant, to compute the inverse of a matrix or solve general matrix equations, to determine Smith normal forms of matrices over integers or polynomials with coefficients in a field and to determine canonical forms of symmetric matrices. In this paper, we illustrate applications of elementary transformations in the theoretical side by applying them to give systematic and concise proofs of several important results in linear algebra.

Key words: linear algebra; elementary transformation; matrix operation

0 引　言

线性代数学习中, 初等变换是矩阵运算的基本操作方法, 用中国科学技术大学线性代数教学的行语来说, 就是“打洞”方法. 它处于线性代数各类计算问题求解的核心位置. 它起源于求解线性方程组的消元法, 并被用来求方阵的行列式、可逆矩阵的逆以及更一般的矩阵方程 $\mathsf{AX} = \mathsf{B}$ 的解、向量组中的极大线性无关组、整矩阵和域上多项式矩阵的 Smith 标准形、对称阵的相合标准形等 [1, 2, 3, 4]. 在理论推导中, 可以通过运用初等变换来对矩阵进行降阶, 从而使用数学归纳法来证明. 在线性代数教学中, 我们发现如果能够更灵活地使用初等变换, 特别地, 灵活运用初等变换的不变性质, 不少线性代数核心结果的证明会比经典教材里的证明更简单.

本文基于这一观点, 阐述如何系统而简洁地证明线性代数下述重要理论结果：线性方程组求解的基本定理, 行列式函数的积性, 克莱姆法则, 可逆的等价条件, 矩阵相抵标准形定理, 矩阵秩三种定义的等价

[基金项目] 国家自然科学基金（12001510,12371013）；基础学科拔尖计划 2.0 研究课题（20211035）；中国科学技术大学十四五教材项目（2021xghjc16）

性, 整系数或域上多项式矩阵行列式因子的唯一性和 Smith 标准形的存在唯一性定理等. 通过本文介绍的方法, 教师可以在线性代数的教学中引入初等变换的思考方式, 让学生更容易理解和应用线性代数的理论. 这种"化繁为简"的思考方式可以帮助学生将抽象的问题转化为具体的行、列变换, 从而更加直观地理解和解决问题.

近几年, 本文作者在教学中采用了这种讲授方式, 并取得了比较积极的反响和良好的效果. 学生们通过实际操作和练习, 逐渐掌握了初等变换的技巧和应用方法. 他们发现, 通过初等变换, 原本复杂的线性代数问题可以被简化为一系列简单的步骤, 从而更容易理解和解决. 综上所述, 通过引入初等变换的教学方法, 可以让学生更容易理解和应用线性代数的理论. 这不仅提高了教师的教学效果, 也提高了学生的学习效果. 因此, 鼓励更多的教师在线性代数的教学中尝试使用初等变换的教学方法, 以提升教学质量和学生的学习成果.

1 初等变换的核心结果

本文中, 记 I_n 为 n 阶单位矩阵, 记 O 为零矩阵, e_j 为第 j 个分量等于 1 而其它分量均为 0 的列向量（它们的维度由上下文确定）. 除去文章最后一节, 所有涉及的矩阵都是定义在某个域 F 上的矩阵.

1.1 初等变换

先回顾一下初等变换的定义.

定义 1 ([1])**.** 矩阵的初等变换是指如下三类行或者列变换：

(i) 第一类初等变换：互换矩阵的两行或者两列；

(ii) 第二类初等变换：在矩阵的某行或者某列乘以一个非零常数；

(iii) 第三类初等变换：矩阵的某行或者列的常数倍加到另外一行或者列.

下面的事实十分简单, 但在后面的论述中常常会遇到.

命题 1. 初等变换是可逆变换, 它的逆变换是同类型的初等变换. 更确切地说,

(i) 第一类初等变换的逆就是它自己；

(ii) 第二类初等变换如果乘以的常数是 λ, 那么它的逆变换就是乘以 λ^{-1} (即除以 λ)；

(iii) 第三类初等变换如果加上的常数倍是 a, 那么它的逆就是减去 a 倍.

注 1. 因为初等变换是可逆变换, 所以当我们证明初等变换前后两个集合相等或者两个量相等时, 只需要证明两集合间的包含关系, 或者两个量之间具有 $\leqslant$ 或整除的关系, 而另外一个方向的结果就是显然的. 这个方法在后面的证明中常常用到, 不再做特别说明.

本文要初等矩阵的概念.

定义 2 ([1])**.** 单位矩阵 I_n 做初等变换得到的方阵, 称为初等矩阵.

(i) 初等矩阵 P_{ij} 是互换 I_n 的第 i、j 行得到的矩阵, 它也是互换 I_n 的第 i、j 列得到的矩阵. 称 P_{ij} 为第一类初等矩阵.

(ii) 初等矩阵 $\mathsf{D}_i(\lambda)$ 是将 I_n 的第 i 行乘以 λ 得到的矩阵, 它也是将 I_n 第 i 列乘以 λ 得到的矩阵. 称 $\mathsf{D}_i(\lambda)$ 为第二类初等矩阵.

(iii) 初等矩阵 $\mathsf{T}_{ij}(a)$ 是将 I_n 的第 j 行的 a 倍加到第 i 行得到的矩阵, 它也是将 I_n 的第 i 列的 a 倍加到第 j 列得到的矩阵. 称 $\mathsf{T}_{ij}(a)$ 为第三类初等矩阵.

下面两个定理是初等变换理论的核心定理.

定理 1. 对 $m \times n$ 阶矩阵 A 做初等行变换相当于对 A 左乘对应的初等矩阵 P, 做初等列变换相当于对 A 右乘对应的初等矩阵 P, 其中 P 是对单位阵 I_m 做同样的初等行变换（或者对 I_n 做同样的初等列变换）得到的方阵.

将命题 1 通过定理 1 翻译成矩阵语言, 就有

$$\mathsf{P}_{ij}^2 = \mathsf{D}_i(\lambda)\mathsf{D}_i(\lambda^{-1}) = \mathsf{D}_i(\lambda^{-1})\mathsf{D}_i(\lambda) = \mathsf{T}_{ij}(a)\mathsf{T}_{ij}(-a) = \mathsf{T}_{ij}(-a)\mathsf{T}_{ij}(a) = \mathsf{I}_n.$$

也就是说：

推论 1. 初等矩阵都是可逆矩阵, 具体地说, 即

$$\mathsf{P}_{ij}^{-1} = \mathsf{P}_{ij}, \quad \mathsf{D}_i(\lambda)^{-1} = \mathsf{D}_i(\lambda^{-1}), \quad \mathsf{T}_{ij}(a)^{-1} = \mathsf{T}_{ij}(-a).$$

定义 3 ([1])**.** 若存在 $r \geqslant 1$ 及 $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$, 使得矩阵 $\mathsf{U} = (u_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ 满足条件：

$$u_{ij_i} \neq 0 \ (1 \leqslant i \leqslant r) \text{ 且 } u_{ij} = 0 \ (i > r \text{ 或者} j < j_i),$$

则称 U 是阶梯形矩阵, 称 u_{ij_i} $(1 \leqslant i \leqslant r)$ 为转角元, 称 r 为台阶数.

如果阶梯型矩阵 U 的所有转角元等于 1 而它们所在列的其它元均为 0, 即它的第 j_i 个列向量是向量 e_i, 我们称 U 是特殊阶梯形矩阵.

根据定义, 显然有 $1 \leqslant r \leqslant \min\{m, n\}$.

引理 1. 台阶数 $r = n$ 的 n 阶阶梯形方阵是主对角元全为非零元的上三角阵. 台阶数 $r = n$ 的 n 阶特殊阶梯形方阵只有一个, 即单位阵 I_n.

证. 这是由于 $1 \leqslant j_1 < \cdots < j_n \leqslant n$, 故 $j_i = i$ 对于 $1 \leqslant i \leqslant n$ 成立. □

下面的定理是初等变换的基本操作：

定理 2. 任意非零矩阵 A 均可通过有限次初等行变换化为特殊阶梯形矩阵, 且 j_1 等于最小的列 j, 使得 A 在该列中有非零元.

翻译成矩阵语言, 这个定理就是说

定理 3. 设 A 是 $m \times n$ 矩阵, 则存在 m 阶初等矩阵 $\mathsf{P}_1, \ldots, \mathsf{P}_s$ 使得 $\mathsf{P}_1 \cdots \mathsf{P}_s\mathsf{A}$ 是特殊阶梯形矩阵.

推论 2. 任意非零矩阵 A 可以通过有限次初等行、列变换化为矩阵 $\begin{pmatrix} \mathsf{I}_r & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix}$, 即存在在 m 阶初等矩阵 $\mathsf{P}_1, \ldots, \mathsf{P}_s$ 和 n 阶初等矩阵 $\mathsf{Q}_1, \ldots, \mathsf{Q}_t$ 使得

$$\mathsf{P}_1 \cdots \mathsf{P}_s\mathsf{A}\mathsf{Q}_1 \cdots \mathsf{Q}_t = \begin{pmatrix} \mathsf{I}_r & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix}.$$

1.2 行列式的积性和矩阵的可逆

假设本小节涉及的矩阵都是 n 阶方阵.

引理 2. 若 A 是 n 阶方阵, P 是 n 阶初等矩阵, 则

$$\det(\mathsf{PA}) = \det(\mathsf{P})\det(\mathsf{A}).$$

由此可知, 若 $\mathsf{P}_1,\dots,\mathsf{P}_s$ 是初等阵, 则

$$\det(\mathsf{P}_1\cdots\mathsf{P}_s\mathsf{A}) = \det(\mathsf{P}_1)\cdots\det(\mathsf{P}_s)\det(\mathsf{A}).$$

证. 在 $\mathsf{P}=\mathsf{P}_{ij}$ 情形证明 $\det(\mathsf{PA})=\det(\mathsf{P})\det(\mathsf{A})$, 其余情形类似. 一方面, $\mathsf{P}_{ij}\mathsf{A}$ 是将方阵 A 互换第 i、j 行得到的矩阵, 根据行列式的基本性质, $\det(\mathsf{P}_{ij}\mathsf{A}) = -\det(\mathsf{A})$. 另一方面, 根据行列式的定义, $\det(\mathsf{P}_{ij})=-1$. 故 $\det(\mathsf{P}_{ij}\mathsf{A})=\det(\mathsf{P}_{ij})\det(\mathsf{A})$. □

定理 4. 设 A 与 B 是同阶方阵, 则 $\det(\mathsf{AB})=\det(\mathsf{A})\det(\mathsf{B})$.

证. 由定理 3, 设 $\mathsf{A}=\mathsf{P}_1\cdots\mathsf{P}_s\mathsf{U}$, 其中 P_1 是初等阵而 U 是特殊阶梯形矩阵.

如果 U 的最后一行全为 0, 那么 $\det(\mathsf{U})=0$. 故由引理 2 得

$$\det(\mathsf{A}) = \det(\mathsf{P}_1)\cdots\det(\mathsf{P}_s)\det(\mathsf{U}) = 0.$$

此时 UB 的最后一行也全是 0, 故 $\det(\mathsf{UB})=0$. 再由引理 2 即得

$$\det(\mathsf{AB}) = \det(\mathsf{P}_1\cdots\mathsf{P}_s\mathsf{UB}) = \det(\mathsf{P}_1)\cdots\det(\mathsf{P}_s)\det(\mathsf{UB}) = 0.$$

所以在这种情况, 有 $\det(\mathsf{AB})=\det(\mathsf{A})\det(\mathsf{B})$.

如果 U 最后一行不全为 0, 那么 $m=n=r$, 故 $\mathsf{U}=\mathsf{I}_m$. 此时 $\mathsf{A}=\mathsf{P}_1\cdots\mathsf{P}_s$, 其中 P_i 为初等阵. 故由引理 2 得

$$\det(\mathsf{AB}) = \det(\mathsf{P}_1\cdots\mathsf{P}_s\mathsf{B}) = \det(\mathsf{P}_1)\cdots\det(\mathsf{P}_s)\det(\mathsf{B}) = \det(\mathsf{A})\det(\mathsf{B}).$$ □

定理 5. 下列条件等价：

(i) 方阵 A 可逆, 即存在 B 使得 $\mathsf{AB}=\mathsf{BA}=\mathsf{I}$.

(ii) A 的行列式 $\det(\mathsf{A})\neq 0$.

(iii) A 是有限个初等矩阵的乘积.

证. (iii)⇒(i)⇒(ii) 是显然的. 下面证明 (ii)⇒(iii).

设 $\mathsf{A}=\mathsf{P}_1\cdots\mathsf{P}_s\mathsf{U}$, 其中 P_i 是初等阵而 U 是特殊阶梯形矩阵. 由 $\det(\mathsf{A})\neq 0$ 知 $\det(\mathsf{U})\neq 0$, 从而 U 的最后一行不全为 0, 故 $r=n, \mathsf{U}=\mathsf{I}_n$. 此时 $\mathsf{A}=\mathsf{P}_1\cdots\mathsf{P}_s$ 是有限多个初等矩阵的乘积. □

注 2. 上述定理说明了求逆算法的正确性：若 A 可逆, 总可以将 (A,I) 经过初等变换变为特殊阶梯形矩阵 (I,B), 此时 B 只能是 A^{-1}.

定理 6 (克莱姆法则)**.** 设 A 是 n 阶可逆方阵, $\mathsf{b}=(b_1,\dots,b_n)^{\mathrm{T}}$, A_i 是将 A 的第 i 列用 b 代替后得到的矩阵. 则线性方程组 $\mathsf{AX}=\mathsf{b}$ 有唯一解

$$\left(\frac{\det(\mathsf{A}_1)}{\det(\mathsf{A})},\frac{\det(\mathsf{A}_2)}{\det(\mathsf{A})},\dots,\frac{\det(\mathsf{A}_n)}{\det(\mathsf{A})}\right)^{\mathrm{T}}.$$

证. 对增广矩阵 (A,b) 做有限次初等行变换后变为矩阵 $(\mathsf{I}_n,\tilde{\mathsf{b}})$. 对于线性方程组 $\mathsf{I}_n\mathsf{X}=\tilde{\mathsf{b}}$, 克莱姆法则显然成立. 但每次初等行变换前后的 $\det(\mathsf{A}_i)$ 和 $\det(\mathsf{A})$ 都相差同一个非零常数倍, 故它们的比值 $\dfrac{\det(\mathsf{A}_i)}{\det(\mathsf{A})}$ 不发生变化. □

1.3 秩的三种定义的等价性

定理 7. 设矩阵 A 经过初等行变换变为矩阵 B.

(i) 矩阵的行向量空间保持不变, 从而 A 与 B 的行秩相等.

(ii) 矩阵列向量组间的线性关系保持不变, 从而 A 与 B 的列秩相等.

(iii) 对任意 $1 \leqslant k \leqslant \min\{m, n\}$, B 的 k 阶子式都是 A 的 k 阶子式的线性组合, 从而 A 与 B 的秩相等.

证. (i) 记 $\mathsf{A} = (\alpha_1, \ldots, \alpha_m)^{\mathrm{T}}$, $\mathsf{B} = (\beta_1, \ldots, \beta_m)^{\mathrm{T}}$. 若初等行变换是

(a) 互换两行, 则 $\{\alpha_1, \ldots, \alpha_m\} = \{\beta_1, \ldots, \beta_m\}$;

(b) 第 i 行乘以 λ 倍, 则 $\beta_i = \lambda\alpha_i$, $\alpha_i = \lambda^{-1}\beta_i$, 而 $\beta_{i'} = \alpha_{i'}$ 若 $i' \neq i$;

(c) 第 j 行的 a 倍加到第 i 行, 则 $\beta_i = \alpha_i + a\alpha_j$, $\alpha_i = \beta_i - a\beta_j$, 而 $\beta_{i'} = \alpha_{i'}$ 若 $i' \neq i$.

所有情况都说明它们的行向量空间是同一个空间.

(ii) 记 $\mathsf{A} = (a_{ij}) = (\alpha'_1, \ldots, \alpha'_n)$, $\mathsf{B} = (b_{ij}) = (\beta'_1, \ldots, \beta'_n)$. 设 $\lambda_1\alpha'_1 + \cdots + \lambda_n\alpha'_n = 0$, 等价地说, 即对 $1 \leqslant i \leqslant m$, 均有 $\sum_{j=1}^n \lambda_j a_{ij} = 0$. 根据三类初等行变换的具体情况, 对于固定的 i 和 $1 \leqslant j \leqslant n$, 同时有 b_{ij} 等于 a_{ij} 或 $a_{i'j}$ 或 λa_{ij} 或 $a_{ij} + a a_{i'j}$, 故 $\sum_{j=1}^n \lambda_j b_{ij} = 0$ 对所有 $1 \leqslant i \leqslant m$ 成立.

(iii) 设我们讨论的 B 的子式为 $\mathsf{B}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}$.

(a) 若初等变换涉及到的行都在或者都不在集合 $\{i_1, \ldots, i_k\}$ 中, 则

$$\mathsf{B}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = c\mathsf{A}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix},$$

其中 $c = \pm 1$ 或者 λ.

(b) 初等变换为互换第 i 行与第 j 行, 其中 $i \in \{i_1, \ldots, i_k\}$ 而 $j \notin \{i_1, \ldots, i_k\}$. 将集合 $\{i_1, \ldots, i_k, j\} \backslash \{i\}$ 中元素排序为 $i'_1 < \cdots < i'_k$, 则

$$\mathsf{B}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = \pm\mathsf{A}\begin{pmatrix} i'_1 & \cdots & i'_k \\ j_1 & \cdots & j_k \end{pmatrix}.$$

(c) 初等变换将第 j 行的 a 倍加到第 i 行, 其中 $i \in \{i_1, \ldots, i_k\}$ 而 $j \notin \{i_1, \ldots, i_k\}$. 如同 (ii) 一样排序集合 $\{i_1, \ldots, i_k, j\} \backslash \{i\}$ 中的元素, 则

$$\mathsf{B}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = \mathsf{A}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} \pm a\mathsf{A}\begin{pmatrix} i'_1 & \cdots & i'_k \\ j_1 & \cdots & j_k \end{pmatrix}. \qquad \square$$

同理, 对于初等列变换则有

定理 8. 设矩阵 A 经过初等列变换变为矩阵 B.

(i) 矩阵的行向量组间的线性关系保持不变, 从而 A 与 B 的行秩相等.

(ii) 矩阵的列向量空间保持不变, 从而 A 与 B 的列秩相等.

(iii) 对任意 $1 \leqslant k \leqslant \min\{m,n\}$, B 的 k 阶子式都是 A 的 k 阶子式的线性组合, 从而 A 与 B 的秩相等.

推论 3. 矩阵的列秩、行秩和秩三者相等. 特别地, 当矩阵通过初等变换化为阶梯形矩阵时, 台阶数等于矩阵的秩是唯一确定的.

证. 根据定理 7, 只需考虑 A 是 (特殊) 阶梯形矩阵的情形, 此时这是显然的.

更显然地, 根据定理 7 和定理 8, 只需考虑 $\mathsf{A} = \begin{pmatrix} \mathsf{I}_r & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix}$ 的情形. □

推论 4. 同型矩阵 A 与 B 相抵当且仅当它们的秩相同, 故每个秩 r 的矩阵都相抵于 $\begin{pmatrix} \mathsf{I}_r & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix}$, 且不相抵于 $\begin{pmatrix} \mathsf{I}_s & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix}$, 其中 $s \neq r$.

推论 5. 设 A 经过初等行变换变为特殊阶梯形矩阵 U.

(i) U 的非零行向量构成 A 的行向量空间的一组基.

(ii) A 的第 $j_1, \dots, j_r$ 个列向量构成 A 的列向量空间的一组基, 即它们构成 A 的列向量组的一个极大线性无关组.

证. (i) 这是由于 A 的行向量空间等于 U 的行向量空间, 维数 $\mathrm{rank}(\mathsf{A})$ 等于 U 的非零行向量的个数, 故 U 的这些非零行向量一定构成这个空间的一组基.

(ii) 由于 U 的第 $j_1, \dots, j_r$ 列列向量集合 $\{\mathsf{e}_1, \dots, \mathsf{e}_r\}$ 是 U 的列向量空间的一组基, 而初等行变换不改变列向量间的线性关系, 故 A 的第 $j_1, \dots, j_r$ 列列向量集合组成 A 的列向量空间的一组基. □

1.4 线性方程组的求解

设 $\mathsf{A} = (a_{ij})_{m\times n}, \mathsf{b} = (b_1, \dots, b_m)^{\mathrm{T}}, \mathsf{X} = (x_1, \dots, x_n)$. 考虑一般线性方程组 $\mathsf{AX} = \mathsf{b}$ 的求解. 我们不妨假设存在 $a_{i1} \neq 0$, 否则变量 x_1 没有出现在线性方程组里面, 其取值可以自由选择.

熟知需要对增广矩阵 (A, b) 做初等行变换. 这里的关键在于：

引理 3. 初等行变换不改变线性方程组的解集合.

证. 由于初等变换的可逆性, 只需要证明变换前方程组的解都是变换后方程组的解, 而这是显然的. □

定理 9. 设增广矩阵 (A, b) 经初等行变换变为特殊阶梯形矩阵 $(\mathsf{U}, \tilde{\mathsf{b}})$, 其中 $\mathsf{U} = (u_{ij})$ 作为阶梯形矩阵转角元所在的列为 $1 = j_1 < \cdots < j_r$. 则 $\mathsf{AX} = \mathsf{b}$ 的解集合等于 $\mathsf{UX} = \tilde{\mathsf{b}}$ 的解集合. 更确切地说,

(i) 非齐次线性方程组 $\mathsf{AX} = \mathsf{b}$ 有解当且仅当 $\mathrm{rank}(\mathsf{A}, \mathsf{b}) = \mathrm{rank}(\mathsf{A}) = r$, 换言之, 当且仅当 $\tilde{\mathsf{b}} = (\tilde{b}_1, \dots, \tilde{b}_r, 0, \dots)^{\mathrm{T}}$.

(ii) 如果 $\mathrm{rank}(\mathsf{A}, \mathsf{b}) = \mathrm{rank}(\mathsf{A}) = r$, 设 X_0 是 $\mathsf{AX} = b$ 的一个解, V_{A} 是齐次线性方程组 $\mathsf{AX} = 0$ 的解空间, 那么 $\mathsf{AX} = \mathsf{b}$ 的解集合等于 $\{\mathsf{X}_0 + \mathsf{Y} \mid \mathsf{Y} \in V_{\mathsf{A}}\}$.

(iii) 如果 $\mathrm{rank}(\mathsf{A}, \mathsf{b}) = \mathrm{rank}(\mathsf{A}) = r$, 那么 $\mathsf{AX} = \mathsf{b}$ 的一组特解 $\mathsf{X}_0 = (x_j)_{1\leqslant j\leqslant n}$ 可以取为：$x_{j_r} = \tilde{b}_r, x_{j_k} = \tilde{b}_k - \tilde{b}_{k+1} (1 \leqslant k < r)$ 而对于其它的 j 则 $x_j = 0$.

(iv) 记 $j_{r+1} = n + 1$. 对于 $j_k < j < j_{k+1}$, 令 $\alpha_j = \mathsf{e}_j - \sum_{i\leqslant k} u_{ij}\mathsf{e}_{j_i} \in F^{n\times 1}$. 则 V_{A} 是以 $\{\alpha_j \mid j \neq j_1, \dots, j_r\}$ 为基的 $n - r$ 维线性空间.

注 3. 此处我们将 A 变为特殊阶梯形矩阵 U, 这让定理的结果可以写得更漂亮. 当然在实际计算中, U 为阶梯形矩阵就已经足够求解.

证. (i) 一方面, 若 $\mathsf{UX} = \tilde{b}$ 有解, 则必有 $\tilde{b}_i = 0$ 对于 $i > r$ 成立. 反过来由 (iii) 即得.

(ii) 一方面若 $\mathsf{Y} \in V_{\mathsf{A}}$, 则 $\mathsf{A}(\mathsf{X}_0 + \mathsf{Y}) = b$. 另一方面, 若 $\mathsf{AX} = \mathsf{AX}_0 = b$, 则 $\mathsf{A}(\mathsf{X} - \mathsf{X}_0) = 0$, 即 $\mathsf{X} - \mathsf{X}_0 \in V_{\mathsf{A}}$ 且 $\mathsf{X} = \mathsf{X}_0 + (\mathsf{X} - \mathsf{X}_0)$.

(iii) 直接验算这样得到的 X_0 满足 $\mathsf{UX}_0 = \tilde{b}$, 故 $\mathsf{AX}_0 = b$.

(iv) 直接验算 $\mathsf{U}\alpha_j = 0$, 故 $\mathsf{A}\alpha_j = 0$ 从而 $\alpha_j \in V_{\mathsf{A}}$. 现在设 $\alpha = \sum_{j=1}^n c_j \mathsf{e}_j \in V_{\mathsf{A}}$, 则

$$\beta = \alpha - \sum_{j \notin \{j_1, \dots, j_r\}} c_j \alpha_j = \sum_{k=1}^{r} d_{j_k} \mathsf{e}_{j_k} \in V_{\mathsf{A}}$$

满足等式 $\mathsf{U}\beta = 0$. 将其展开就得到 $d_{j_k} = 0$ 对于 $1 \leqslant k \leqslant r$ 成立, 故 $\beta = 0$. 故 V_{A} 是由 $\{\alpha_j \mid j \neq j_1, \dots, j_r\}$ 生成的线性空间. 另一方面, $\{\alpha_j \mid j \neq j_1, \dots, j_r\}$ 的线性无关性通过考虑它们的 j 分量 $(j \notin \{j_1, \dots, j_r\})$ 立得. □

1.5 整系数和域上多项式系数矩阵的初等变换

本节设 $R = \mathbb{Z}$ 或者 $F[x]$, 其中 F 是域. 环 R 与域有一个主要区别：R 的非零元不一定是乘法可逆元而域的非零元都是乘法可逆元. 事实上 R 的乘法可逆元集合 $R^\times$ 即

$$R^\times = \begin{cases} \{\pm 1\}, & \text{若 } R = \mathbb{Z}; \\ F^\times, & \text{若 } R = F[x]. \end{cases}$$

R 上矩阵的第一类初等变换和第三类初等变换与 F 上的初等变换没有什么不同. 但对于第二类初等变换, 对矩阵的某行或者某列同乘以 λ 倍. 要使它为可逆变换, 则需要 $\lambda \in R^\times$. 这样, 在对 R 上的矩阵做初等变换时, 不能任意除一个非零数得到 1. 需要用到 R 上的带余除法.

定义 4. 设 $a \in R$. 称

$$h(a) = \begin{cases} |a|, & \text{若 } R = \mathbb{Z}; \\ \deg(a), & \text{若 } R = F[x] \end{cases}$$

为 a 的高度. 若 $a > 0$ (当 $R = \mathbb{Z}$ 时) 或者 a 首一 (当 $R = F[x]$ 时), 称 a 恒正.

引理 4 (带余除法)**.** 若 $a, b \in R$ 且 $a \neq 0$, 则存在唯一的 q 和 $r \in R$, 使得 $b = qa + r$ 且 $h(r) < h(a)$.

引理 5. 设 $(a, *, \dots, *)$ 与 $(b, *, \dots, *)$ 是矩阵 A 的两行且 $a \neq 0$. 则经过有限次初等变换后可以将这两条行向量变为 $(d, *, \dots, *)$ 与 $(0, *, \dots, *)$, 其余行不变, 这里 d 是 a 和 b 的最大公因子.

证. 这个引理实际上就是欧几里得算法的实现. 不妨设 $(a, *, \dots, *)$ 是 A 的第 1 行, $(b, *, \dots, *)$ 是第 2 行. 设 $b = qb + r$, 则第 1 行的 $-q$ 倍加到第 2 行, 它就变为 $(r, *, \dots, *)$. 若 $r \neq 0$, 令 $a = q_1 r + r_1$. 则第 2 行的 $-q_1$ 倍加到第 1 行, 它就变为 $(r_1, *, \dots, *)$. 继续这个过程, 我们就将 (a, b) 变为 $(r_n, 0)$ 或者 $(0, r_n)$. 再做一次第二类变换 (使 r_n 恒正) 和最多一次第一类变换 (将 $(0, r_n)$ 换为 $(r_n, 0)$), (a, b) 就变为 $(d, 0)$, 其中 d 是 a 和 b 的最大公因子. □

定理 10. 设 A 是 R 上的矩阵, B 是 A 经过初等变换得到的矩阵. 则对任意 $1 \leqslant k \leqslant \min\{m, n\}$, B 的 k 阶子式都是 A 的 k 阶子式的线性组合. 从而

(i) A 的秩等于 B 的秩.

(ii) 对于 $1 \leqslant k \leqslant \operatorname{rank}(\mathsf{A})$, A 的 k 阶行列式因子 $D_k(\mathsf{A})$, 即 A 的所有 k 阶子式的最大公因子, 等于 B 的 k 阶行列式因子 $D_k(\mathsf{B})$.

证. 只需证明 $D_k(\mathsf{A}) = D_k(\mathsf{B})$, 其余同定理 7(iii) 的证明. 由于 B 的 k 阶子式都是 A 的 k 阶子式的线性组合, 故 $D_k(\mathsf{A}) \mid D_k(\mathsf{B})$. 再由初等变换的可逆性, $D_k(\mathsf{B}) \mid D_k(\mathsf{A})$. 故两者相等. □

定理 11 (Smith 标准形)**.** *R 上任意非零阵 A 经过有限次初等变换可以变为矩阵 $\mathsf{S} = \begin{pmatrix} \mathsf{D} & 0 \\ 0 & 0 \end{pmatrix}$ 的形式, 其中 $\mathsf{D} = \operatorname{diag}(d_1, \ldots, d_r)$, d_k 恒正且 $d_1 \mid d_2 \mid \cdots \mid d_r$, 而且这样形式的 S 是唯一的：$r = \operatorname{rank}(\mathsf{A})$ 和 $d_k = d_k(\mathsf{A}) = D_k(\mathsf{A})/D_{k-1}(\mathsf{A})$ 由 A 决定.*

证. (i) 首先证明存在性.

(a) 由于 $\mathsf{A} \neq 0$, 设 $a_{ij} \neq 0$, 互换第 i 行与第 1 行, 再互换第 j 列与第 1 列, 最后对第 1 行乘以 a_{11} 的符号或者首项系数的逆, 这样得到的新矩阵 A 就满足 a_{11} 恒正.

(b) 由于高度不能无限递降, 可以假设经过有限次初等变换后 $h(a_{11})$ 达到最小. 我们断言此时 $a_{11} \mid a_{i1}$. 否则, 若 $a_{11} \nmid a_{i1}$, 根据引理 5, 我们可以通过初等行变换将 a_{11} 替换为 $\gcd(a_{11}, a_{i1})$, 它恒正且高度比 a_{11} 小, 与 $h(a_{11})$ 的最小性矛盾. 同理我们有 $a_{11} \mid a_{1j}$. 这样可以通过第三类初等变换使得 $a_{i1} = a_{1j} = 0 (i > 1, j > 1)$ 而 a_{11} 不变.

接下来断言此时 $a_{11} \mid a_{ij}$ 对所有 i, j 成立. 否则, 设 $a_{11} \nmid a_{ij}$. 将第 j 列加到第 1 列, 那么新的第 1 行是 $(a_{11}, \ldots)$ 而第 i 行是 $(a_{ij}, \ldots)$. 根据引理 5, 我们可以通过初等行变换将 a_{11} 替换为 $\gcd(a_{11}, a_{ij})$, 它恒正且高度比 a_{11} 小, 与 $h(a_{11})$ 的最小性矛盾. 故经过有限次初等变换后, 可以得到新的 A 为 $\begin{pmatrix} d_1 & 0 \\ 0 & \mathsf{B} \end{pmatrix}$ 的形式, 其中 d_1 恒正且 $\mathsf{B} = (b_{ij})$ 中每个元素均被 d_1 整除.

(c) 根据归纳, B 经有限次初等变换后就得到定理要求的形式 $\mathsf{S}_1 = \begin{pmatrix} \mathsf{D}_1 & 0 \\ 0 & 0 \end{pmatrix}$ 的形式, 其中 $\mathsf{D}_1 = \operatorname{diag}(d_2, \ldots, d_r)$, 且 $d_2 \mid \cdots \mid d_r$. 由于 d_2 是 B 中元素的线性组合, 所以 $d_1 \mid d_2$. 这样就经过有限次初等变换将 A 变为 S 的形式.

(ii) 再证唯一性. 由于初等变换不改变矩阵的秩, 故 $r = \operatorname{rank}(\mathsf{S}) = \operatorname{rank}(\mathsf{A})$. 由于初等变换不改变行列式因子, 故 $D_k(\mathsf{A}) = D_k(\mathsf{S})$, 但根据定义, $D_k(S) = d_1 \cdots d_k$, 故

$$d_k = d_k(\mathsf{A}) = D_k(\mathsf{A})/D_{k-1}(\mathsf{A}).$$ □

注 4. Smith 标准形的唯一性在一般线性代数教材是使用更复杂的 Binet-Cauchy 公式证明的 [1].

另外, 唯一性的证明还说明 $D_{k-1}(\mathsf{A}) \mid D_k(\mathsf{A})$, 这个事实也可以通过 Laplace 展开公式 [1] 得出, 只要说明每个 k 阶子式都是 $k-1$ 阶子式的线性组合即可.

定理 12. *考虑 R 上的方阵. 下列条件等价：*

(i) *方阵 A 可逆, 即存在 B 使得 $\mathsf{AB} = \mathsf{BA} = \mathsf{I}$.*

(ii) *A 的行列式 $\det(\mathsf{A}) \in R^\times$.*

(iii) *A 是有限个初等矩阵的乘积.*

证. (iii)⇒(i)⇒(ii) 是显然的. 下面证明 (ii)⇒(iii). 设 $\mathsf{A} = \mathsf{P}_1 \cdots \mathsf{P}_s \mathsf{S} \mathsf{Q}_1 \cdots \mathsf{Q}_t$, 其中 $\mathsf{P}_i, \mathsf{Q}_j$ 是初等阵而 S 是 A 的 Smith 标准形. 由 $\det(\mathsf{A}) \in R^\times$ 知 $\operatorname{rank}(\mathsf{S}) = n, S = \operatorname{diag}(d_1, \ldots, d_n)$ 且 $\det(\mathsf{S}) = d_1 \cdots d_n \in R^\times$. 由于 d_i 恒正, 这只能 $d_1 = \cdots = d_n = 1$, 即 $\mathsf{S} = \mathsf{I}_n$. □

2 结 论

在线性代数的教学中, 初等变换是一种非常重要的工具, 可以用来证明线性代数中的定理与结论. 通过使用初等变换, 可以将一般的矩阵转化为性质已知的、结构简单的矩阵, 从而证明各个主要结论和定理.

这种教学方式的优势在于, 它能够帮助学生克服对线性代数中抽象理论的恐惧. 相比于直接面对抽象的问题, 将问题转化为具体的行、列变换可以使学生更容易理解和解决问题. 这种直观的思路可以提高教师的教学效果, 同时也能够提高学生的学习效果.

此外, 通过使用初等变换进行证明, 学生还可以同步地提高他们的计算能力. 在进行初等变换的过程中, 学生需要进行矩阵的运算和计算, 这可以帮助他们巩固和提高他们的计算技巧. 因此, 这种教学方式不仅可以帮助学生理解线性代数的理论, 还可以培养他们的计算能力.

总之, 通过使用初等变换来证明线性代数中的定理与结论是一种有效的教学方法. 它可以帮助学生克服对抽象理论的恐惧, 提高教师的教学效果和学生的学习效果, 同时还可以同步地提高学生的计算能力. 因此, 在线性代数的教学中, 我们可以充分利用初等变换这一工具, 来帮助学生更好地理解和应用线性代数的知识.

致谢 作者非常感谢相关文献对本文的启发以及审稿专家提出的宝贵意见.

[参 考 文 献]

[1] 李尚志. 线性代数 [M]. 北京: 高等教育出版社, 2016:9–23,138,405–414.

[2] 李炯生, 查建国, 王新茂. 线性代数 [M]. 2 版. 合肥: 中国科学技术大学出版社, 2010:134–142.

[3] 唐烁, 朱士信. 线性代数 [M]. 北京: 高等教育出版社, 2018:95–111.

[4] 蓝以中. 线性代数引论 [M]. 2 版. 北京: 北京大学出版社, 1999:1–70.

[5] 王新茂, 线性代数讲义 [M]. 合肥：中国科学技术大学出版社, 2021:7–14.