

以线性方程组为中心诠释线性代数核心概念

刘海

上海对外经贸大学 统计与数据科学学院, 上海 201620

DOI:10.61369/ASDS.2026070006

摘 要：《线性代数》的理论体系起源于对线性方程组的求解，因此，以线性方程组为中心诠释《线性代数》中的核心概念是自然的、易懂的，这有助于学生理解抽象概念，帮助学生建构知识体系，提升教学效果。行列式是线性方程组唯一解公式的速记符号；矩阵是表示线性方程组的简便符号；初等行变换是线性方程组同解变形的操作；矩阵的秩是有效方程的个数；用矩阵运算表示线性方程是定义矩阵乘法规则的一个动机；向量是表示方程组解的简便符号。

关 键 词： 线性方程组；行列式；矩阵；向量

Illustrating the Core Concepts of Linear Algebra in Terms of Linear Equations

Liu Hai

School of Statistics and Data Science, Shanghai University of International Business and Economics, Shanghai 201620

Abstract： The theoretical system of "Linear Algebra" originates from solving linear equations. Therefore, it is natural and easy to interpret the core concepts of "Linear Algebra" in terms of linear equations, which helps students understand abstract concepts and construct knowledge system, and improves teaching effect. The determinant is a shorthand symbol for the unique solution formula of a linear system of equations. A matrix is a simple symbol for a group of linear equations. Elementary row transformation is the operation of deformation of linear equations with their solutions unchanged. The rank of the matrix is the number of valid equations. Representation of linear equations by matrix operations is one motivation for defining the rule of matrix multiplication. Vectors are simple notations for the solutions of a system of equations.

Keywords： linear equations; determinant; matrix; vector

引言

如何求解线性方程组，是《线性代数》课程要解决的一个核心问题。为了回答这个问题，《线性代数》课程讲述了行列式、矩阵和向量的相关知识，它们占据了该课程近三分之二的篇幅^[1]。运用这些知识，人们可以彻底回答线性方程组求解的理论问题。因此，以线性方程组为中心诠释《线性代数》中的核心概念是自然的、易懂的，这有助于学生理解抽象概念，帮助学生建构知识体系，提升教学效果。在此基础上，我们认为考研数学一范围内的《线性代数》课程的最佳教学章顺序应该是：行列式、矩阵、向量、矩阵的特征值和特征向量、二次型。前三章可以视为一个整体，其教学过程应围绕线性方程组的求解为中心，展开行列式、矩阵和向量的讲解，它们自身形成一个逻辑闭环。后两章可以看作是《线性代数》的另一主题，重在研究矩阵的性质及其在二次型中的应用。接下来，本文将基于线性方程组的视角对行列式、矩阵和向量等概念进行解释。

一、行列式是线性方程组唯一解公式的速记符号

1683年，日本数学家关孝和首次提出了行列式的概念。1693年，德国数学家莱布尼茨（Leibniz）在写给洛必达的信中也独立地使用了行列式，并给出了线性方程组系数行列式的定义。1750年，瑞士数学家克莱姆（Cramer）提出了克莱姆法则，这是行列式理论的一个重要应用，用于解线性方程组。18世纪末至19世纪

初，行列式理论逐步完善，数学家如拉普拉斯（Laplace）和高斯（Gauss）对其性质和计算方法进行了深入研究^[2]。

数学家是出于何种动机要提出行列式的概念？很多教科书都没有回答这个问题，而是直接给出行列式的定义，然后讲述行列式的性质和计算等知识，最后将行列式应用于克拉姆法则中，给出方程组的唯一解公式。

然而，在教学过程中，笔者认为，我们应该先澄清在解线性

基金项目：上海对外经贸大学统计与数据科学学院教学团队项目；2026年上海对外经贸大学校级重点课程建设项目。

作者简介：刘海，上海对外经贸大学统计与数据科学学院，副教授，研究方向：组合数学、线性代数；邮箱：liuhai@suibe.edu.cn。

方程组的过程中会出现什么问题?为了解决这个问题,人们有必要引进行列式,然后再去讲解行列式的定义、性质和计算等知识。在这样的逻辑框架下,学生们才能建立健全的行列式知识体系,知道行列式从何处来,到何处去,形成知识闭环。

实际上,行列式是为了快速记忆含 n 个方程和 n 个未知数(简记为 $n \times n$)的线性方程组唯一解公式而引进的符号。在这种符号的帮助下,人们可以方便快捷地给出并记住方程组的唯一解公式。方程组的解是数,因此行列式的本质也是数。如何由行列式的符号计算出它所表达的真实数字呢?这就是行列式理论需要研究的问题,于是同学们需要学习行列式的定义、性质和计算,这些知识的最终目标就是为了计算出行列式符号所表达的真实数字或代数式。于是,一方面,人们可以运用行列式的符号快速写出方程组唯一解公式;另一方面,人们又能运用行列式的理论计算出行列式所表达的真实数字;最终,人们就能获得方程组的唯一解。下面,我们来展示这个教学过程。

对于 2×2 的线性方程组: $\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$, 运用消元法, 人们容易得到: 当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时, 方程组有唯一解, 且

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}, \quad x_2 = \frac{b_2 a_{11} - a_{21} b_1}{a_{11} a_{22} - a_{12} a_{21}}. \quad (1)$$

公式(1)的规律性不明显,不方便记忆和运用,人们需要发明一种简便的符号来速记公式(1),在这种动机下,二阶行列式的符号就应运而生。

定义1: 符号 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 称为二阶行列式, 其值为 $a_{11}a_{22} - a_{12}a_{21}$, 即:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (2)$$

二阶行列式的符号是由4个数排成2行2列的阵列, 为了表示它是一个整体, 用一对竖线将其框起来。二阶行列式的本质是一个数, 是有值的, 其值可以通过“对角线法则”计算出来, 即主对角线元素的积减去副对角线元素的积。

利用二阶行列式, 公式(1)可以写成如下形式:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \quad (3)$$

记 $D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $D_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}$, $D_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$, 则当 $D \neq 0$ 时, 公式(3)可略写为:

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}. \quad (4)$$

由于 D 中的元素均由方程组的系数按其在方程组中的原始位置顺序构成, 所以称 D 为系数行列式。 D_1 可以看作将 D 的第1列元素换成方程组的常数项列、其它列保存不变而得到。 D_2 可以看作将 D 的第2列元素换成常数项列、其它列保存不变而得到。在这种规则默认下, 公式(4)就是公式(1)的速记公式。

类似地, 对于 3×3 的线性方程组: $\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$, 运用消元法, 可以得到, 当 $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \neq 0$ 时, 方程组有唯一解, 且

$$\begin{aligned} x_1 &= \frac{b_1 a_{22} a_{33} + a_{12} a_{23} b_3 + a_{13} b_2 a_{32} - b_1 a_{23} a_{32} - a_{12} b_2 a_{33} - a_{13} a_{22} b_3}{a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31}}, \\ x_2 &= \frac{a_{11} b_2 a_{33} + b_1 a_{23} a_{31} + a_{13} a_{21} b_3 - a_{11} a_{23} b_3 - b_1 a_{21} a_{33} - a_{13} b_2 a_{31}}{a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31}}, \\ x_3 &= \frac{a_{11} a_{22} b_3 + a_{12} b_2 a_{31} + b_1 a_{21} a_{32} - a_{11} b_2 a_{32} - a_{12} a_{21} b_3 - b_1 a_{22} a_{31}}{a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31}}. \end{aligned} \quad (5)$$

显然, 公式(5)的规律性很不明显, 非常不方便记忆和运用, 但运用三阶行列符号和上述规则对公式(5)改成, 将会得到类似公式(4)的公式。

定义2: 符号 $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ 称为三阶行列式, 其值为 $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$, 即:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \quad (6)$$

三阶行列式的符号是由9个数排成3行3列的阵列, 其本质也是一个数, 是有值的, 它的值也可以通过“对角线法则”计算出来, 只是规则较二阶行列式稍微复杂一些。对角线法则是计算二阶和三阶行列式的一种规则, 任何一本线性代数的教材都会讲授这个法则^[3]。在公式(2)或(6)中, 利用对角线法则, 人们可以容易地将左边行列式的符号展开为右边的值, 这个过程叫行列式的展开或计算。

需要说明的是, 在创造行列式的概念时, 人们是将(2)和(6)式的右边强制写成左边的简便符号, 但行列式的难点在于如何将左边展开成右边。在将二阶和三阶行列式推广到一般 n 阶行列式时, 人们容易写出左边行列式的符号, 却难以发现其右边的展开形式。实际上, 这是 n 阶行列式的定义要解决的问题。

$$\text{记 } D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix},$$

则当 $D \neq 0$ 时, 公式(5)可改写为:

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad x_3 = \frac{D_3}{D}. \quad (7)$$

公式(7)是公式(5)的速记公式, 其规则是: 每个未知数 $x_i (i=1, 2, 3)$ 的值都是一个分数形式, 分母是系数行列式 D , 分子 D_i 是将 D 的第 i 列元素换成方程组的常数项列、其它列保持不变而得到的行列式。

于是, 对于 3×3 的线性方程组, 其求解过程可分为三步: 第一、写成系数行列式 D , 并按公式(6)或对角线法则计算出其值; 第二, 当 $D \neq 0$ 时, 按公式(7)写出速记公式; 第三, 按上述规则分别写出行列式 D_1 、 D_2 和 D_3 , 根据公式(6)或对角线法则分别计算出它们的值, 再分别代入公式(7), 即可求出每个未知数的值。需要注意的是, 本文的主要目的是论述引入行列式的动机和作用, 即行列式可以速记方程组有唯一解时的公式, 因此我们不需要考虑系数行列式 $D=0$ 的情形。

现在的问题是，类似公式（4）和（7）的这种速记规则是否可以推广到一般的 $n \times n$ 的线性方程组？也就是说，对于方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \quad (8)$$

$$\text{记 } D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad D_i = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_i & a_{1,i+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{i,i-1} & b_i & a_{i,i+1} & \cdots & a_{in} \end{vmatrix} \quad (i=1,2,\dots,n),$$

当 $D \neq 0$ 时，是否有：

$$x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, \dots, x_n = \frac{D_n}{D}. \quad (9)$$

答案是肯定的。实际上，公式（9）就是著名的克拉姆法则，是当方程组（8）有唯一解时解的计算公式。由于行列式符号的使用，这个公式的规律非常明显，很容易被记住和运用。公式（9）中的每个行列式都是 n 阶行列式，它们都是由 n^2 个数字排成 n 行 n 列的阵列，它们的本质也是数、有值。如何计算它们的值，就是行列式理论需要回答的问题。

于是，教师们就可以开启行列式理论的教学，带领学生们一起去观察、归纳、发现一般 n 阶行列式的定义，探索行列式的性质，研究行列式计算的方法和技巧，并运用行列式的理论证明公式（9）。实际上，检验一套行列式理论是否正确的标准，就是看人们运用这套理论能否证明公式（9）。公式（9）是我们想要得到结论，一套正确的行列式理论应当证明它。

自此，行列式理论的教学就形成了一个闭环，学生们知道了行列式的概念形成于方程组，是方程组唯一解公式的速记符号，它的一个主要去向就是帮助人们快速解方程组。行列式理论的主要目标是为了计算行列式，获得行列式的值。

二、矩阵是表示线性方程组的简便符号

克拉姆法则的优势在于，当 $n \times n$ 的线性方程组有唯一解时，运用行列式给出了唯一解的公式。公式的好处在于便于被记住和运用；也便于算法化，让电脑也能运用公式解方程组。然而，克拉姆法则也存在局限性。首先，对于 $n \times n$ 的线性方程组，它要计算 $n+1$ 个 n 阶行列式，而每个高阶行列式的计算是费时的，所以它的计算复杂度较高，速度较慢。其次，它的适用范围较窄，只适用于 $n \times n$ 的线性方程组有唯一解时的情形，对一般的含 m 个方程和 n 个未知数（简记为 $m \times n$ ）的线性方程组，它是无效的。因此，仅有行列式的工具，人们无法完全解决线性方程组的求解，还需要其它的工具。

求解线性方程组的第二个工具是矩阵。矩阵的产生和发展经历了从古代萌芽到现代理论的漫长过程。它的正式提出和系统化主要归功于19世纪的数学家凯莱和西尔维斯特^[4]。随着数学和其他学科的发展，矩阵逐渐成为描述线性关系、解决实际问题的重要工具，并在现代科学和技术中发挥着不可替代的作用。矩阵的应用极其广泛，求解线性方程组只是其应用的一个方面。从线性

方程组的视角，解释矩阵及其相关概念和运算的产生，可以帮助学生理解抽象概念，提升教学效果。

第一，矩阵可以看作是为了简便表示线性方程组的符号而被提出的。矩阵的初等行变换是线性方程组同解变形的操作。行阶梯形矩阵对应着一种特殊形式的方程组，从它们身上可以直接看出方程组是否有解以及解的个数。行简化阶梯形矩阵对应着另一种特殊形式的方程组，从它们身上可以直接写出方程组的解。运用高斯消元法解线性方程组的过程，可看作是对矩阵进行初等行变换，消元化为行阶梯形矩阵，再回代化为行简化阶梯形矩阵的过程。

高斯消元法（Gaussian Elimination）是解线性方程组的一种经典方法，其命名来源于德国数学家高斯。虽然高斯并不是最早提出这种方法的人，但他在19世纪初对其进行了系统化和推广，使其成为线性代数中的核心工具之一。在图1中，我们以一个 4×3 的具体线性方程组为例，给出了高斯消元法解线性方程组的步骤。高斯消元法是一种可算法化的机械步骤，适用于所有线性方程组，其详细具体步骤描述可参见^[5, 第25页]。

在图1中，方程组①是待解的原始方程组；方程组②是行阶梯形方程组，从它可以看出方程组仅有1个解；方程组③是行简化阶梯形方程组，从它可以直接写出方程组的解为： $x_1=1, x_2=3, x_3=2$ 。从①到②的过程叫消元，从②到③的过程叫回代，中间涉及的操作包含三类：一是交换两个方程的位置；二是某个方程的两边乘以一个非零数；三是某个方程的倍数加到另一个方程上去。这三类操作叫方程组的初等变换，它们不会改变方程组的解。因此，图1中的9个方程组彼此等价，它们具有相同的解。高斯消元法的步骤就是对线性方程组进行初等变换，消元化为行阶梯形方程组，再回代化为行简化阶梯形方程组的过程。

$$\begin{aligned} & \begin{cases} 2x_1 + 2x_2 - x_3 = 6 \\ x_1 - 2x_2 + 4x_3 = 3 \\ 4x_1 - 2x_2 + 7x_3 = 12 \\ 5x_1 + 7x_2 + x_3 = 28 \end{cases} \text{①} \rightarrow \begin{cases} x_1 - 2x_2 + 4x_3 = 3 \\ 2x_1 + 2x_2 - x_3 = 6 \\ 4x_1 - 2x_2 + 7x_3 = 12 \\ 5x_1 + 7x_2 + x_3 = 28 \end{cases} \rightarrow \begin{cases} x_1 - 2x_2 + 4x_3 = 3 \\ 6x_2 - 9x_3 = 0 \\ 6x_2 - 9x_3 = 0 \\ 17x_2 - 19x_3 = 13 \end{cases} \\ & \rightarrow \begin{cases} x_1 - 2x_2 + 4x_3 = 3 \\ 6x_2 - 9x_3 = 0 \\ 0 = 0 \\ \frac{13}{2}x_3 = 13 \end{cases} \rightarrow \begin{cases} x_1 - 2x_2 + 4x_3 = 3 \\ 6x_2 - 9x_3 = 0 \\ \frac{13}{2}x_3 = 13 \\ 0 = 0 \end{cases} \rightarrow \begin{cases} x_1 - 2x_2 + 4x_3 = 3 \\ 2x_2 - 3x_3 = 0 \\ x_3 = 2 \\ 0 = 0 \end{cases} \text{②} \\ & \rightarrow \begin{cases} x_1 - 2x_2 = -5 \\ 2x_2 = 6 \\ x_3 = 2 \\ 0 = 0 \end{cases} \rightarrow \begin{cases} x_1 - 2x_2 = -5 \\ x_2 = 3 \\ x_3 = 2 \\ 0 = 0 \end{cases} \rightarrow \begin{cases} x_1 = 1 \\ x_2 = 3 \\ x_3 = 2 \\ 0 = 0 \end{cases} \text{③} \end{aligned}$$

图1：高斯消元法的例子

然而，图1解方程组的过程，在写法上有很多累赘，不简洁、不方便。大括号、等号和未知数字母都不是方程组的本质、是多余的，方程组的本质是未知数字母前面的系数和常数项。未知数字母可以默认用 x_1, x_2 等来表示，系数和常数项才是影响方程组解的决定因素。因此，现在的问题是：是否可以只选取系数和常数项来表示方程组？答案是肯定的。在回答这个问题的过程中，人们就会很自然地提出矩阵的概念。

$$\begin{aligned}
 &\begin{pmatrix} 2 & 2 & - & 6 \\ 1 & -2 & 4 & 3 \\ 4 & -2 & 7 & 12 \\ 5 & 7 & 1 & 28 \end{pmatrix} \textcircled{4} \rightarrow \begin{pmatrix} 1 & -2 & 4 & 3 \\ 2 & 2 & -1 & 6 \\ 4 & -2 & 7 & 12 \\ 5 & 7 & 1 & 28 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & 4 & 3 \\ 0 & 6 & -9 & 0 \\ 0 & 6 & -9 & 0 \\ 0 & 17 & -19 & 13 \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} 1 & -2 & 4 & 3 \\ 0 & 6 & -9 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{13}{2} & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & 4 & 3 \\ 0 & 6 & -9 & 0 \\ 0 & 0 & \frac{13}{2} & 13 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & 4 & 3 \\ 0 & 2 & -3 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \textcircled{5} \\
 &\rightarrow \begin{pmatrix} 1 & -2 & 0 & -5 \\ 0 & 2 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & 0 & -5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \textcircled{6}
 \end{aligned}$$

图2: 高斯消元法的矩阵表示

图2给出了图1中每个线性方程组的矩阵表示, 这些矩阵叫方程组的增广矩阵, 每个增广矩阵是由方程组的系数和常数项按其在方程组中的位置顺序排列而成的矩形阵列。仅由方程组的系数按其在方程组中的位置顺序排列而成的矩形阵列叫方程组的系数矩阵, 它比增广矩阵少了最后一列。系数矩阵或增广矩阵的行数等于方程组中方程的个数, 系数矩阵的列数等于方程组中未知数的个数。增广矩阵在如下规则下与方程组一一对应: 首先, 矩阵的第*i*行对应方程组中的第*i*个方程; 其次, 矩阵的第*j*列是方程组中第*j*个未知数 x_j 前面的系数, 矩阵的最后一列是方程组的常数项。

在图2中, 矩阵④是方程组①的增广矩阵, 它代表了原始待解的方程组。矩阵⑤是行阶梯形方程组②的增广矩阵, 于是类似形式⑤的矩阵叫行阶梯形矩阵。矩阵⑥是行简化阶梯形方程组③的增广矩阵, 类似形式③的矩阵叫行简化阶梯形矩阵。方程组的初等变换用在矩阵身上叫做矩阵的初等行变换。因此, 用矩阵的语言来说, 高斯消元法就是对原方程组的增广矩阵进行初等行变换, 消元化为行阶梯形矩阵, 再回代化为行简化阶梯形矩阵的过程。

用矩阵及其初等行变换表达高斯消元法解线性方程组的过程存在两大优势。首先, 它书写简洁、方便、省时, 省去了大括号、未知数和等号这些非本质的东西。其次, 它能让高斯消元法程序化, 让电脑也能解方程组。电脑能认识形式④的矩阵, 却不能认识形式①的方程组。如果要往数学软件中输入一个方程组, 我们通常输入的就是它的增广矩阵。电脑能对增广矩阵实施初等行变换的操作, 从而解出方程组。这两个优势体现了在解线性方程组中引入矩阵的好处和作用。

第二, 矩阵的秩是有效方程的个数。在《线性代数》中, 矩阵的秩是一个非常重要而抽象的概念, 它通常有四种解释。首先, 矩阵的秩是矩阵能生成的所有行列式中, 非零子式的最大阶数; 其次, 矩阵的秩等于与它等价行阶梯形矩阵中非零行的行数; 再次, 矩阵的秩等于其行(列)向量组的秩; 最后, 矩阵的秩等于其行(列)空间的维数^[6]。然而, 这四种解释都比较抽象、难懂, 也不能说明为何要提出秩的概念。实际上, 矩阵的秩还有另一种解释, 即它等于其对应线性方程组中有效方程的个数, 这

种解释简单直观, 能说明引入秩概念的动机, 还能直观展现秩的一个重要性质, 即初等行变换不改变矩阵的秩。

给定一个矩阵, 如矩阵④, 我们可以把它当作某个线性方程组的增广矩阵, 如方程组①。于是, 每个矩阵都等同于一个线性方程组。在一个线性方程组中, 每个方程都是对未知数的约束或限制, 能缩小未知数取值范围的方程叫有效方程, 不能缩小未知数取值范围的方程叫无效方程。不失一般性, 在方程组中, 人们可以按从上到下的顺序给方程组排序, 然后来指出哪个方程是有效的, 哪个是无效的。在方程组①中, 前两个方程是有效的; 但第三个方程是无效的, 它可看作由第一个方程加上第二个方程的2倍而得到, 因而不能缩小未知数的取值范围; 第四个方程是有效的。无效方程也可以理解为由它前面的某些方程通过初等变换而得到的方程。在高斯消元法的变换过程中, 无效方程在增广矩阵中最终会变为零行, 其对应的方程是: $0=0$, 不妨称之为零行方程, 它们恒成立, 对未知数没有实质性约束, 不能缩小其取值范围。由此可见, 有效方程才是方程组中最本质的方程, 有效方程的个数非常重要, 它会影响方程组解的个数。

$$\begin{aligned}
 &\begin{cases} x_1 + 5x_2 - x_3 - x_4 = -1 \\ x_1 - 2x_2 + x_3 + 3x_4 = 3 \\ 3x_1 + 8x_2 - x_3 + x_4 = 1 \\ x_1 - 9x_2 + 3x_3 + 7x_4 = 7 \end{cases} \textcircled{I}: \begin{pmatrix} 1 & 5 & -1 & -1 & -1 \\ 1 & -2 & 1 & 3 & 3 \\ 3 & 8 & -1 & 1 & 1 \\ 1 & -9 & 3 & 7 & 7 \end{pmatrix} \textcircled{A1} \rightarrow \begin{pmatrix} 1 & 5 & -1 & -1 & -1 \\ 0 & -7 & 2 & 4 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \textcircled{B1} \\
 &\text{方程组} \qquad \qquad \qquad \text{增广矩阵} \qquad \qquad \qquad \text{行阶梯形矩阵} \\
 &\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 = 1 \\ x_2 + x_3 - 4x_4 = 1 \\ x_1 + 2x_2 + 3x_3 - x_4 = 4 \\ 2x_1 - x_2 + x_3 + 18x_4 = -1 \end{cases} \textcircled{II}: \begin{pmatrix} 1 & 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -4 & 1 \\ 1 & 2 & 3 & -1 & 4 \\ 2 & -1 & 1 & 18 & -1 \end{pmatrix} \textcircled{A2} \rightarrow \begin{pmatrix} 1 & 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \textcircled{B2}
 \end{aligned}$$

图3: 两个线性方程组的例子

通常情况下, 譬如在线性方程组有解的前提下, 如果有效方程的个数等于未知数的个数, 那么该线性方程组有唯一解; 如果有效方程的个数少于未知数的个数, 那么它有无穷多解。在线性方程组中, 如果一个方程与它前面几个方程对未知数的约束产生矛盾或冲突, 那么可称这个方程为矛盾方程。只要存在1个矛盾方程, 这个方程组就是无解的。矛盾方程也是一种有效方程, 它能将方程组的解缩小为0个。反之, 如果一个线性方程组无解, 那么它就一定存在矛盾方程。在高斯消元法的变换过程中, 矛盾方程在增广矩阵中最终会变为“ $0, \dots, 0, b$ ”这种类型的行, 其中 $b \neq 0$, 即最后一个数是非零数, 其它数均为0的行, 其对应的方程是: $0=b$, 显然是一个矛盾式, 并且这样的行在行阶梯形矩阵中只会有一行。

在图3中, 方程组(I)有2个有效方程, 2个无效方程, 4个未知数, 没有矛盾方程; 有效方程的个数少于未知数的个数, 于是它有无穷多解。方程组(II)有3个有效方程, 1个无效方程, 但第三个方程是矛盾方程, 于是它无解。对于图1的方程组①, 它有3个有效方程, 1个无效方程, 3个未知数, 没有矛盾方程, 于是它有唯一解。需要注意的是, 当有效方程的个数多于未知数的个数时, 此时必然存在矛盾方程, 线性方程组总是无解的。因此, 在判断线性方程组解的个数时, 一个非常直观而显然的结论是: 方程组无解 $\Leftrightarrow$ 它存在矛盾方程; 方程组有解 $\Leftrightarrow$ 它没有矛盾方程; 方程组有唯一解 $\Leftrightarrow$ 它没有矛盾方程且有效方程个数等于未

知数个数；方程组有无穷多解 $\Leftrightarrow$ 它没有矛盾方程且有效方程个数少于未知数个数。

然而，给定一个线性方程组或它对应的增广矩阵，在不对它进行初等行变换之前，人们很难看出它有几个有效方程，以及是否存在矛盾方程。但通过初等行变换，将增广矩阵化为行阶梯型矩阵时，人们却可以容易看出行阶梯型矩阵对应方程组的有效方程的个数、以及是否存在矛盾方程：有效方程个数等于行阶梯型矩阵中非零行的行数；如果有“ $0, \dots, 0, b$ ”（ $b \neq 0$ ）类型的行，则存在矛盾方程；否则，不存在矛盾方程。由于初等行变换本质上是在对原方程组进行同解变形，因此它不会改变方程组有效方程的个数，也不改变是否存在矛盾方程，否则方程组的解也会发生改变。换言之，有效方程的个数和是否存在矛盾方程，是初等行变换下的不变量。原始的增广矩阵和它对应的行阶梯型矩阵，虽然它们对应的方程组存在不同形式，但它们的解相同，有效方程的个数相同。所以，给定一个矩阵，其对应线性方程组有效方程的个数就是这个矩阵的秩，秩是初等行变换下的不变量。

有了秩的概念，人们就可以直接针对原线性方程组或其增广矩阵来描述解的判定，而不需要先对它进行初等行变换化为行阶梯形，然后再描述，这也可以视为引入秩概念的一个动机。对于一个 $m \times n$ 的线性方程组，设其系数矩阵为 A ，增广矩阵为 $\bar{A}$ ，增广矩阵比系数矩阵多了最后 1 列，这一列由常数项构成。记 A 的秩为 $r(A)$ ， $\bar{A}$ 的秩为 $r(\bar{A})$ 。容易验证，如果在 $\bar{A}$ 对应的行阶梯形矩阵 B 中，存在“ $0, \dots, 0, b$ ”类型的行，则 $r(\bar{A}) = r(A) + 1$ ；如果在 B 中不存在“ $0, \dots, 0, b$ ”类型的行，则 $r(\bar{A}) = r(A)$ 。于是，线性方程组是否存在矛盾方程，可以用 $r(\bar{A})$ 和 $r(A)$ 的大小比较来描述。因此，我们可以得到《线性代数》的一个重要定理：线性方程组解的判定定理，即：方程组无解 $\Leftrightarrow r(\bar{A}) = r(A) + 1$ ；方程组有解 $\Leftrightarrow r(\bar{A}) = r(A)$ ；方程组有唯一解 $\Leftrightarrow r(\bar{A}) = r(A) = n$ ；方程组有无穷多解 $\Leftrightarrow r(\bar{A}) = r(A) < n$ 。

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \Leftrightarrow AX = B$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

图4：线性方程组的矩阵乘法表示

第三，用矩阵运算表示线性方程是定义矩阵乘法规则的一个动机。在初等代数中，一元一次方程是通过数的乘法来表示的，即 $ax=b$ ，系数 a 和未知数 x 之间是用数的乘法运算连接的。类似地，对于一个 $m \times n$ 的线性方程组，设其系数矩阵为 A ，未知数矩阵为 X ，常数项矩阵为 B ，如果我们也将它表示成 $AX=B$ 的形式，其中 A 与 X 之间的运算叫矩阵的乘法，那么这种乘法的规则是什么？

如图4，如果仔细观察原方程组与 $AX=B$ 的对应关系，人们将不难发现，矩阵乘法的规则应当是 A 的每一行元素“ $a_{11}, a_{12}, \dots, a_{1n}$ ”与 X 的第 1 列元素“ $x_1, x_2, \dots, x_n$ ”分别对应相

乘再相加得到： $a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n$ ；它等于 B 中元素 b_1 ，从而得到第 i 个方程： $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$ ，其中 $i=1, 2, \dots, m$ 。在这种思想的启发下，人们便能定义矩阵乘法的规则。设 A 是 $m \times s$ 的矩阵，即它有 m 行 s 列， B 是 $s \times n$ 的矩阵，则乘积 AB 将是 $m \times n$ 的矩阵，且乘积矩阵的每个元素 C_{ik} 等于 A 的第 i 行元素“ $a_{i1}, a_{i2}, \dots, a_{is}$ ”与 B 的第 k 列元素“ $b_{1k}, b_{2k}, \dots, b_{sk}$ ”对应相乘再相加而得到，即 $C_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{is}b_{sk}$ ，其中 $i=1, 2, \dots, m$ ， $k=1, 2, \dots, n$ 。

三、向量是表示线性方程组解的简便符号

运用行列式和矩阵的工具，人们不仅可以判断线性方程组解的个数，还可以求出它的所有解。但有一个问题是这两个工具无法解决的，即当线性方程组有无穷多解时，如何描述这些解之间的关系和结构？或者说，是否可以找出其中有限多个解，它们能表示或生成其它所有的解？这样人们就化无穷为有限，极大地减少要研究的解的个数，只需要研究这有限个解即可。为了回答这个问题，人们需要引入向量，用向量来表示线性方程组的解，用向量之间的关系来描述解的关系。所以，从线性方程组的视角看，向量是表示其解的简便符号。

$$\begin{cases} x_1 + x_2 - x_3 - x_4 = 0, \\ 2x_1 - 5x_2 + 3x_3 + 2x_4 = 0 \text{ ①}, \\ 7x_1 - 7x_2 + 3x_3 + x_4 = 0 \end{cases} \begin{cases} x_1 = \frac{2}{7} \\ x_2 = \frac{5}{7} \text{ ②}, \\ x_3 = 1 \\ x_4 = 0 \end{cases} \begin{cases} x_1 = \frac{3}{7} \\ x_2 = \frac{4}{7} \text{ ③}, \\ x_3 = 0 \\ x_4 = 1 \end{cases} \eta_1 = \begin{pmatrix} \frac{2}{7} \\ \frac{5}{7} \\ 1 \\ 0 \end{pmatrix}, \eta_2 = \begin{pmatrix} \frac{3}{7} \\ \frac{4}{7} \\ 0 \\ 1 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$\begin{cases} x_1 = \frac{2}{7}c_1 + \frac{3}{7}c_2 \\ x_2 = \frac{5}{7}c_1 + \frac{4}{7}c_2 \text{ (} c_1, c_2 \text{ 为任意数) ④} \\ x_3 = c_1 \\ x_4 = c_2 \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = c_1 \begin{pmatrix} \frac{2}{7} \\ \frac{5}{7} \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} \frac{3}{7} \\ \frac{4}{7} \\ 0 \\ 1 \end{pmatrix} \text{ ⑤} \Leftrightarrow x = c_1\eta_1 + c_2\eta_2 \text{ ⑥}$$

图5：线性方程组解的向量表示

起初，人们习惯用方程组的形式来表达方程组的解，但这种表示形式不利用描述解与解之间的关系。例如，在图5中，方程组②和③是方程组①的两个解，但②和③的关系似乎难以用语言来表达。方程组④表达了①的所有解，但它没有揭示这无穷多解的结构。如果把方程组②和③分别写成向量 η_1 和 η_2 的形式，把①的任意解写成向量 x 的形式，把④写成⑤或⑥的形式，那么借用向量理论的相关语言，人们就很容易描述解的关系和结构。为此，教师们此时就要先向同学们讲授线性组合、线性表示、线性相关、线性无关、极大线性无关组等向量知识，然后再用向量术语去描述解的关系和结构。

在图5中，运用向量理论的术语， η_1 和 η_2 叫方程组①的解向量，它们的关系是线性无关。⑤或⑥表达的含义是：①的任何解都可以写成 η_1 和 η_2 的线性组合，或者说， η_1 和 η_2 的所有线性组合生成了①的所有解。这就表明， η_1 和 η_2 这两个解是所有解中最本质的解，通过这两个解，人们就能掌握①的无穷多个解。

自此，我们清晰地看到，为了求解线性方程组，人们引入了行列式、矩阵和向量三个工具，运用这三个工具，人们彻底回答了线性方程组求解的理论问题。线性方程组、行列式、矩阵和向量，这四块内容占据了《线性代数》课程的主要篇幅，并形成了

完美的逻辑闭环：先提出如何解线性方程组的问题，然后引入工具解决问题。因此，以线性方程组为中心，引入并诠释行列式、矩阵和向量的相关知识，这种教学组织框架必定有助于学生理解抽象概念，帮助学生建构知识体系，提升教学效果。

四、结论

行列式和矩阵虽然起源于线性方程组，但它们还有其它的很多解释和作用。例如，在几何中，行列式可以表示有向面积或体积^[6]；矩阵可以表示线性变换、基变换、内积等。在数据科学中，矩阵也是一种数据结构。向量的概念不来源于线性方程组，其最

初的模型应是物理学中的力、速度、加速度等概念，以及几何中的有向线段和坐标。也就是说，向量也存在多种解释，如它可以表示力、速度、有向线段等^[7]，在数学科学中，向量也是一种数据结构。然而，从线性方程组的视角来解释行列式、矩阵和向量，或许是一种最简单、最容易被学生理解的解释方式，也能从某个侧面反映《线性代数》引入或研究行列式、矩阵和向量的动机。基于线性方程组的视角，行列式是线性方程组唯一解公式的速记符号；矩阵是表示线性方程组的简便符号；初等行变换是线性方程组同解变形的操作；矩阵的秩是有效方程的个数；用矩阵运算表示线性方程是定义矩阵乘法规则的一个动机；向量是表示方程组解的简便符号。

参考文献

[1] 吴赣昌. 线性代数 (理工类 第五版) [M]. 北京: 中国人民大学出版社, 2017.
[2] 杨浩菊. 行列式理论历史研究 [D]. 西北大学, 2004.
[3] 袁明生, 刘海, 唐国平. 线性代数 (第二版) [M]. 北京: 清华大学出版社, 2020.
[4] 李文林. 数学史概论 (第三版) [M]. 北京: 高等教育出版社, 2011.
[5] Leon S J. Linear Algebra with Applications (Ninth Edition) [M]. Edinburgh: Pearson Education, 2015.
[6] 吕林根, 许子道. 解析几何 (第五版) [M]. 北京: 高等教育出版社, 2019.
[7] [美] 莫里斯克莱因. 古今数学思想 (第二册) [M]. 张理京等译. 上海: 上海科学技术出版社, 2013.