由鸡兔同笼问题从方程组步入矩阵

作者： 姜鱼

日期：2025 年 3 月 1 日

考虑这么一个二元一次方程组：

$$\begin{cases}3x+7y=114\\5x-y=514\end{cases}$$

我们考虑的不是如何对它求解，而是我们如何写下的它。

1 譬如一个鸡兔同笼问题

现有10个头，28个脚，我们写下的方程应该是：

$$\begin{cases}x+y=10\\2x+4y=28\end{cases}$$

我们向来写方程，都是一行一行写。意思就是说，我们先设鸡有x个，兔有y个。第一行的思路是：我们总共有10个动物，就有10个头，其中鸡的个数对总头的贡献为1，兔头对总头贡献为1，于是写出了$1x+1y=10$。第二行的思路是：我们总共有28个脚，其中鸡对脚数的贡献是2，所以2作为系数乘在x边，而兔对脚数贡献为4，所以4作用于y，由此写出$2x+4y=28$。

于是分成两行，写了两个方程

$$\begin{cases}x+y=10 & \cdots\cdots\text{头方程}\\2x+4y=28 & \cdots\cdots\text{脚方程}\end{cases}$$

一行方程表示一个事件，可以看出，头和脚是两个完全不同维度的事物，我们以事件为线索，分别对两个维度的事件写等式，串联起两个独立维度的方程，便可以充分地求解出一个方程。但是一定要确保两个方程由两个完全独立的事件所给出，这样我们才能确保信息是有效的，如果我们写出这样的两个方程：$\begin{cases}x+y=10\\2x+2y=20\end{cases}$，虽然看似是联立两个二元一次方程，实际上两个方程根据一个事件所给出，是并不独立的两个维度，则不能对其有效求解。

为什么我们要刻意强调以事件为线索地列方程呢？因为事实上，真正的鸡兔同笼问题的求解思路，和我们这样列二元一次方程组的思路大相径庭。

我们这次区分的事件并非“头”和“脚”，而是真正从“鸡”和“兔”出发，字面意义地去思考“鸡兔”同笼的问题。假设笼中全是鸡，那么10头鸡总共有$10\times 2=20$只脚，然而总共有28个脚，多出来的8个脚由兔子提供，每个鸡换成兔子后，都会使脚多2只，因此多出的8只脚来自$8\div 2=4$只兔。这是传统的求解鸡兔同笼的思路，仔细审视其思路，其实依然并非完全不能

用方程组表示，只是要适当处理一下：

$$\begin{cases} x = 10 - y & \cdots\cdots \text{先带入}x = 10\text{，再去掉其中y的部分} \\ 2x = 2 \times (10 - y) - (2y - 8) & \cdots\cdots \text{由此导出}2y - 8 = 0\text{，算出兔数} \end{cases}$$

这次我们把鸡看做整体线索，同时考察“头系数=1”和“脚系数=2”对鸡数的作用效果，最终通过鸡兔自身性质（脚的数量差2）的比较，得出鸡和兔的数量。我们的计算离不开10和28两个结果导向的媒介，但是我们解析的方式截然不同。1头2脚整体算作属于“鸡”的性质，在这里共同作用于鸡数x的时候，再用方程组分成两行来表述，就违和了许多。因为我们的思路其实并没有如最初写方程组那般分成“两个维度”去思考和分析问题。

向量方程就是这样诞生的。

如果作用于x的整体系数作为鸡的性质，从而整体看待，用另一种方式记为$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \times x$，上面的数对应方程组中上面那行，下面的数对应方程组中下面那行。我们继续用这种思路把整个方程组写为：

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} (10 - y) - \begin{bmatrix} 0 \\ 2y - 8 \end{bmatrix}$$

进而回头再利用$(10 - y) = x$导出$\begin{bmatrix} 0 \\ 2y - 8 \end{bmatrix} = 0$ 或 $2y - 8 = 0$，最终解出y再进一步解出x来实现求解鸡兔同笼。

我们写出$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x$的结构并非难点，但是为了构造等式后面那一长串的内容可谓煞费苦心。我们虽然知晓二元一次方程组的解法，而仍愿意探究鸡兔同笼的原理，其在于原则上有趣的思路多多益善，而并不为其复杂的形式所遏制。但是一味地奇技淫巧也并非我们追逐的目标，适当的去芜存菁才是学习数学的本质。

我们的思路是以完全不同的结构$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x$代替原先两行方程组的结构，进一步说，是用这种别样的思路来表述变量之间的关系。其本质只是思路上从以**分解结果计算头和脚为核心线索**改为了**以变量元为核心线索**进行方程的书写，所以其实繁杂的形式并不是必要的，我们进行了一系列的构造，其最终的无非还得从10只头，28个脚入手。对于y，也无需将其进行复杂的构造，等同于x一般去书写对其作用的，代表其性质的系数组$\begin{bmatrix} 1 \\ 4 \end{bmatrix} y$。所以我们进一步借鉴二元一次方程组的简洁形式，将其以新的书写思路改造为：

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 4 \end{bmatrix} y = \begin{bmatrix} 10 \\ 28 \end{bmatrix}$$

这样的结构简直焕然一新！

2 从新的角度重新解构线性方程组

对于任意的方程，例如那个$\begin{cases}3x+7y=114\\5x-y=514\end{cases}$ 我们都可以将其改写为新的结构：$\begin{bmatrix}3\\5\end{bmatrix}x+\begin{bmatrix}7\\-1\end{bmatrix}y=\begin{bmatrix}114\\514\end{bmatrix}$。这样做并非我想自讨苦吃地专注于形式和结构的改造，而是以全新的思路去处理变量关系。提炼结构，是为了更好地挖掘其中的性质。最终即便我们任意写的方程中，每一行没有了一个实际场景的意义作为支撑，例如x,y未必再代表鸡和兔，数表$\begin{bmatrix}3\\5\end{bmatrix}$也不再代表鸡所具有的头和脚的数量性质。但是仍然有系数作用于变量的方程，我们都可以挖掘出其本质的共性，故鸡兔同笼的问题最终可以总结为：可以将方程组分离成变量与系数的两个部分，每一个变量元有一个对应的系数列表。

对于多元方程组$\begin{cases}a_1x_1+\cdots+d_1x_n & b_1\\ \vdots\ \ddots\ \vdots \quad = & \vdots\\ a_nx_1+\cdots+d_nx_n & b_n\end{cases}$，我们都可以将其转化为$\begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix}x_1+\cdots+\begin{bmatrix}d_1\\\vdots\\d_n\end{bmatrix}x_n=\begin{bmatrix}b_1\\\vdots\\b_n\end{bmatrix}$的形式，前者以每行为导向分析问题，解构结果条件。后者以变量元一列为导向按列书写方程组，侧重于表达与变量元有关的性质。

我们把横行的数量称为“维度”，体现的是事件，按维度书写方程，是以事件为线索集合不同变量的数据。我们把纵列的数量称为“自由度”，代表的是有多少未知的自由变量，按自由度书写方程，是以变量为线索集合不同维度事件中所有的性质的数据。

3 方程组与向量

进一步地分析结构，我们依旧可以再次给出方程组一个新的理解角度。中学阶段向量的形式写作$\vec{a}=(a_1,a_2,\cdots,a_n)$，本质是作为一群数据的集合体。同样是作为一群数据的集合体，我们可以把形式为$\vec{a}=(a_1,a_2,\cdots,a_n)$的向量改写为$\vec{a}=\begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix}$，也可以把形式为$\vec{a}=\begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix}$的向量改写为$\vec{a}=(a_1,a_2,\cdots,a_n)$。

于是之前我们竖向写的数表，本质和向量无异，则意味着可以当做向量处理（这也是为什么把横行的数量成为维度）。那么接下来有趣的就来了。首先，我们把

$$\begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix}x_1+\cdots+\begin{bmatrix}d_1\\\vdots\\d_n\end{bmatrix}x_n=\begin{bmatrix}b_1\\\vdots\\b_n\end{bmatrix}$$

的模样变为

$$x_1\begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix}+\cdots+x_n\begin{bmatrix}d_1\\\vdots\\d_n\end{bmatrix}=\begin{bmatrix}b_1\\\vdots\\b_n\end{bmatrix}$$

相当于之前把数表的一群东西当做了x_n的系数，如今位置互换，把x_n当成了一个个向量的系数。

总的头和脚的数量，分别由鸡头鸡脚的数量，和兔头兔脚的数量构成，而描述一只鸡的特征的向量$\boldsymbol{A}$，里面的两个维度分别包含了两个独立数据头和脚：$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$，那么鸡的数量则为其加权，去构成整体的头脚数$\begin{bmatrix} 1x \\ 2x \end{bmatrix}$。同样的，我们用兔数为兔的特征加权得到$\begin{bmatrix} 1x \\ 4x \end{bmatrix}$，并将鸡兔加权后的头脚数据相加，就可以构成整体的头脚数据$\begin{bmatrix} 10 \\ 28 \end{bmatrix}$。

广而推之，将每个变量x_n在不同维度的系数数据指标都整合为一个整体向量，令$\boldsymbol{A} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$，再试变量为赋向量的权。将权$x_1, x_2, ..., x_n$如法炮制地赋予向量$\boldsymbol{A}_1, \boldsymbol{A}_2, ..., \boldsymbol{A}_n$，我们的线性方程组则可以直接改写为

$$\boldsymbol{A}_1 x_1 + \cdots + \boldsymbol{A}_n x_n = \boldsymbol{b}$$

不断地精炼，浓缩结构，我们把一个线性方程组浓缩为了单行的一个方程。其中多行多维度的性质涵盖在了向量之内。接下来我们继续提炼结构，之前将向量作为x_n的系数，而今，我们也可以把将多个自由变量x_n视为系数向量的加权，认真思考如下这句话：解线性方程组，如同找到每个系数向量的权$x_1 \to x_n$，令n个基向量最终指向目标向量$\boldsymbol{b}$，并作定义：

$$\boldsymbol{A}_1 x_1 + \cdots + \boldsymbol{A}_n x_n \Longrightarrow \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix} = \boldsymbol{b}$$

其中$\begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n \end{bmatrix}$是代表不同权所具有的特征的向量，通常他们共同组成了n维空间的基底，而$\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix}$则为权向量，为不同的$A_1, ..., A_n$加权，生成n维空间中的$\boldsymbol{b}$向量。

同时由于$\boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{a}_{12} \\ \boldsymbol{a}_{22} \\ \vdots \\ \boldsymbol{a}_{n2} \end{bmatrix}$，所以将其每一位展开，$\begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix}$中的基向量部分，也可以表述为

$$\begin{bmatrix} \boldsymbol{A}_1 & \begin{matrix} \boldsymbol{a}_{12} \\ \boldsymbol{a}_{22} \\ \vdots \\ \boldsymbol{a}_{n2} \end{matrix} & \boldsymbol{A}_3 & \cdots & \boldsymbol{A}_n \end{bmatrix}$$

或者全部展开，形成更大的数表

$$\begin{bmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \cdots & \boldsymbol{a}_{1n} \\ \boldsymbol{a}_{21} & \ddots & & \boldsymbol{a}_{2n} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \cdots & \boldsymbol{a}_{nn} \end{bmatrix}$$

这样一个包含了一个方程组全部系数的数表，被命名叫系数矩阵。

矩阵表示的是一整个线性方程组的系数列表，矩阵运算被定义的目的是为了结构化运算。如同提炼结构是为了更好地挖掘其中的性质，我们做了许多复杂的操作，实际上只是顺着一些基本的逻辑，把一块方程组拆成了三个独立的部分：

(1)系数矩阵$A = \begin{bmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \cdots & \boldsymbol{a}_{1n} \\ \boldsymbol{a}_{21} & \ddots & & \boldsymbol{a}_{2n} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \cdots & \boldsymbol{a}_{nn} \end{bmatrix}$

(2)包含自由变量的权向量$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

(3)一个常数向量$\boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}$

于是我们可以将三个结构整合，把线性运算最终归纳为

$$A\boldsymbol{x} = \boldsymbol{b}$$

4 总结

我们捋了一遍从线性方程组到矩阵的过程：

$$\begin{aligned} \begin{cases} x + 2y + 7z = 11 \\ -2x + 5y + 4z = 45 \\ -5x + 6y - 3z = 14 \end{cases} &\implies \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} x + \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} y + \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} z = \begin{bmatrix} 11 \\ 45 \\ 14 \end{bmatrix} \\ &\implies \begin{bmatrix} \vec{A_1} & \vec{A_2} & \vec{A_3} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ 45 \\ 14 \end{bmatrix} \\ &\implies \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ 45 \\ 14 \end{bmatrix} \\ &\implies A\vec{x} = \vec{b} \end{aligned}$$

同时，也学到了三种看待方程组运算的角度:

一种是横着看：按照独立的事件集合不同元素的等式。

一种是竖着看：按照自由变量的性质集合不同维度的系数。

一种是从向量看，把线性方程组视为若干系数向量的线性组合。

这是一篇来自姜鱼的文章，于学习时偶然间顿有灵感，于是记录自己的学习心得，整理为文章。多有疏漏，欢迎各位与我交流，批评指正，谢谢大家阅读。