2.3 用矩阵消元

1 第一步将方程 $A\boldsymbol{x}=\boldsymbol{b}$ 与矩阵 E_{21} 相乘以产生 $E_{21}A\boldsymbol{x}=E_{21}\boldsymbol{b}$。

2 由于 x_1 从方程 2 中被消去，因此矩阵 $E_{21}A$ 的行 2、列 1 处为 0。

3 E_{21} 为**单位矩阵**（全 1 对角阵）在行 2、列 1 处减去乘数 a_{21}/a_{11}。

4 矩阵—矩阵相乘就是 n 次矩阵—向量相乘：$\boldsymbol{EA}=[\boldsymbol{Ea_1}\ldots\boldsymbol{Ea_n}]$。

5 我们还必须把 $E\boldsymbol{b}$ 乘起来！这样一来 E 就乘以**增广矩阵** $[A\ \boldsymbol{b}]=[\boldsymbol{a_1}\ldots\boldsymbol{a_n}\boldsymbol{b}]$。

6 消元法将 $A\boldsymbol{x}=\boldsymbol{b}$ 乘上 $E_{21},E_{31},\ldots,E_{n1}$，再乘上 $E_{32},E_{42},\cdots,E_{n2}$，继续下去。

7 **换行矩阵**不是 E_{ij} 而是 P_{ij}。为求出 P_{ij}，那就交换 I 的 i 行与 j 列。

本节给出了**矩阵乘法**的第一个例子。自然地，我们从包含许多 0 的矩阵开始。我们的目标是理解矩阵的*所作所为*。E 作用于一个向量 $\boldsymbol{b}$ 或一个矩阵 A 来产生一个新向量 $E\boldsymbol{b}$ 或一个新矩阵 EA。

我们的第一个例子将是"**消元矩阵**"。它们执行消元步骤。第 j 个方程乘以 l_{ij} 然后从第 i 个方程中减去它。(这从方程 i 中消去 x_j。) 我们需要许多这样的简单矩阵 E_{ij}，它针对主对角线下每个要消去的非零元素。

幸运的是我们不会在后面的章节见到所有这些矩阵。它们是开始接触时的好例子，但它们太多了。它们可以组合成一个一次做所有步骤的总体矩阵 E。最简洁的方式是将它们的逆 $(E_{ij})^{-1}$ 组合成一个总体矩阵 $L=E^{-1}$。以下是下一页的打算。

1. 弄清每一个步骤怎么就是一次矩阵乘法的?

2. 将所有这些步骤 E_{ij} 整合成一个消元矩阵 E。

3. 弄清每个 E_{ij} 是如何由它的逆矩阵 E_{ij}^{-1} 逆转的?

4. 将所有这些逆 E_{ij}^{-1}（按正确顺序）整合成 L。

L 的特殊性质是所有的乘数 l_{ij} 都在其位。在 E 中这些数都混合在一起（从 A 到 U 的前向消元）。它们在 L（撤销消元，从 U 返回到 A）中是完美的。逆过程将步骤与它们的矩阵 E_{ij}^{-1} 置于相反的顺序并且它防止了混合。

本节求矩阵 E_{ij}。2.4 节介绍了四种矩阵乘法方式。2.5 节逆转每个步骤。(对于消元矩阵我们已经在这里见到了 E_{ij}^{-1}。) 然后这些逆成为 L。

矩阵乘以向量与 $A\boldsymbol{x}=\boldsymbol{b}$

上一节中的 3×3 例子具有简写形式 $A\boldsymbol{x}=\boldsymbol{b}$:

$$\begin{aligned}2x_1+4x_2-2x_3&=2\\4x_1+9x_2-3x_3&=8\\-2x_1-3x_2+7x_3&=10\end{aligned}\quad \text{和}\quad \begin{bmatrix}2&4&-2\\4&9&-3\\-2&-3&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2\\8\\10\end{bmatrix}\quad \text{一样。}\tag{1}$$

左边的 9 个数成为矩阵 A。这个矩阵不仅仅列于 $\boldsymbol{x}$ 旁。A *乘以* $\boldsymbol{x}$。"A 乘以 $\boldsymbol{x}$"的运算规则恰好可选来以产生该三个方程。

回顾 A 乘以 $\boldsymbol{x}$。一个矩阵乘以一个向量得出一个向量。当方程个数（3 个）与未知数个数（3 个）相一致时，其矩阵是方形的。我们的矩阵是 3×3 的。一般方阵是 3×3 的。那么向量 $\boldsymbol{x}$ 在 n 维空间中。

$$\textbf{未知数是}\quad \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\quad \textbf{且解为}\quad \boldsymbol{x}=\begin{bmatrix}-1\\2\\2\end{bmatrix}。$$

重点：$A\boldsymbol{x}=\boldsymbol{b}$ 表示方程的行形式也表示列形式。

$$\textbf{列形式}\qquad A\boldsymbol{x}=(-1)\begin{bmatrix}2\\4\\-2\end{bmatrix}+2\begin{bmatrix}4\\9\\-3\end{bmatrix}+2\begin{bmatrix}-2\\-3\\7\end{bmatrix}=\begin{bmatrix}2\\8\\10\end{bmatrix}=\boldsymbol{b}. \tag{2}$$

$A\boldsymbol{x}$ 是 A 的列组合。为了计算 $A\boldsymbol{x}$ 的每个分量，我们运用矩阵乘法的**行形式**。**$A\boldsymbol{x}$ 的分量都是与 $\boldsymbol{A}$ 的行的点积**。与 $\boldsymbol{x}$ 点积的简短公式应用"sigma 表示法"。

$$\text{上面 } A\boldsymbol{x} \text{ 的第一个分量为}\ (-1)(2)+(2)(4)+(2)(-2)。$$
$$A\boldsymbol{x} \text{ 的第 } i \text{ 个分量为}\ (\text{行}i)\cdot\boldsymbol{x}=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n。$$

有时用 sigma 符号写作 $\sum_{j=1}^{n}\boldsymbol{a_{ij}x_j}$。

Σ 是加法指令[1]。从 $j=1$ 开始到 $j=n$ 结束。求和从 $a_{i1}x_1$ 开始到 $a_{in}x_n$ 结束。它构成点积 $(\text{行}i)\cdot\boldsymbol{x}$。

再重复一遍关于矩阵的表示法：行 1、列 1 的元素（左上角）为 a_{11}。行 1、列 3 的元素为 a_{13}。行 3、列 1 的元素为 a_{31}。（行号位于列号之前。）矩阵的"元素"一词与向量的"分量"一词相对应。一般规则：**$\boldsymbol{a_{ij}=A(i,j)}$ 在行 i、列 j 上**。

例 1 这个矩阵具有 $a_{ij}=2i+j$。因此 $a_{11}=3$。并且 $a_{12}=4$ 及 $a_{21}=5$。以下是由号和字母表示的 $A\boldsymbol{x}$：

$$\begin{bmatrix}3&4\\5&6\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}3\cdot2+4\cdot1\\5\cdot2+6\cdot1\end{bmatrix}\qquad\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}a_{11}x_1+a_{12}x_2\\a_{21}x_1+a_{22}x_2\end{bmatrix}。$$

一行乘以一列得出一个点积。

[1] 爱因斯坦通过省略 Σ 来进一步缩短它。$a_{ij}x_j$ 中重复的 j 自动意味着加法。他还将求和写作 $a_i^jx_j$。我们不是爱因斯坦，我们得写上 Σ。

一次消元步骤的矩阵形式

$A\boldsymbol{x}=\boldsymbol{b}$ 是原始方程的简便形式。消元步骤是怎样的呢？在本例中，从第二个方程中减去 2 倍的第一个方程。在等号右边，再从第二个分量减去 2 倍 $\boldsymbol{b}$ 的第一个分量。

$$\textbf{第一步}\qquad \boldsymbol{b}=\begin{bmatrix}2\\\mathbf{8}\\10\end{bmatrix}\quad\text{变为}\quad \boldsymbol{b}_{\text{new}}=\begin{bmatrix}2\\\mathbf{4}\\10\end{bmatrix}。$$

我们想用矩阵来执行减法！当我们将一个"消元矩阵"E 乘以 $\boldsymbol{b}$ 时，实现同一结果 $\boldsymbol{b}_{\text{new}} = E\boldsymbol{b}$。它从 b_2 减去 $2b_1$：

$$\textbf{该消元矩阵是}\quad E = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}。$$

乘以 $\boldsymbol{E}$ 从行 2 减去 2 倍的行 1。行 1 和行 3 保持不变：

$$\begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & \mathbf{1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{2} \\ \mathbf{8} \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ \mathbf{4} \\ 10 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & \mathbf{1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ \boldsymbol{b_2 - 2b_1} \\ b_3 \end{bmatrix}$$

E 的第一和第三行来自于单位矩阵 I。它们不改变第一个与第三个数（2 与 10）。经过消元步骤后出现的新第二个分量是数 4。这是 $b_2 - 2b_1$。

表述像这个 E 这种的"初等矩阵"或"消元矩阵"是简单的。从单位矩阵 I 开始。*把它其中一个 0 改成乘数 $-l$*：

单位矩阵享有 1 在对角线上而且其余为 0。因此对于所有 $\boldsymbol{b}$ 有 $I\boldsymbol{b} = \boldsymbol{b}$。**初等矩阵或消元矩阵** E_{ij} 在 i, j 位置具有额外的非零元素 $-l$。于是 E_{ij} 为从行 i 减去 l 倍的行 j。

例 2 矩阵 E_{31} 在 3, 1 位置为 $-l$：

$$\textbf{单位矩阵}\quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \textbf{消元矩阵}\quad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \boldsymbol{-l} & 0 & 1 \end{bmatrix}。$$

当你将 I 乘以 $\boldsymbol{b}$ 时，你就会得到 $\boldsymbol{b}$。然而 E_{31} 是从第三个元素减去 l 倍的第一个元素。$l = 4$ 时本例得出 $9 - 4 = 5$：

$$I\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} \quad 与 \quad E\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{-4} & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ 3 \\ \mathbf{9} \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ \mathbf{5} \end{bmatrix}。$$

$A\boldsymbol{x} = \boldsymbol{b}$ 左边是怎么样的？两边都讲乘以该 E_{31}。**$\boldsymbol{E_{31}}$ 的目的是要在矩阵的 $(\mathbf{3}, \mathbf{1})$ 位置产生 $\mathbf{0}$**。

这个表示法符合此目的。从 A 开始。运用 E 在主元下面产生 0（第一个 E 是 E_{21}）。以三角矩阵 U 结束。我们现在来观察这些步骤的细节。

首先是个小要点。向量 $\boldsymbol{x}$ 保持不变。解 $\boldsymbol{x}$ 不会因为消元而改变。（这可能不仅仅是个小要点。）其系数矩阵是改变的。当我们从 $A\boldsymbol{x} = \boldsymbol{b}$ 开始然后乘以 E，其结果是 $EA\boldsymbol{x} = E\boldsymbol{b}$。这个新矩阵 EA 是 E *乘以* A 的结果。

声明 *消元矩阵* E_{ij} 是极好的例子，但是你将不会在后面见到它们。它们展示了一个矩阵是如何作用在行上的。通过做几次消元步骤，我们将领悟如何*乘以矩阵*（E 的顺序就变得很重要）。E 的**积与逆**都尤为清晰。本书将使用这两个思想。

有个大问题是：**我们如何将两个矩阵相乘？** 当第一个矩阵是 E 时，我们知道 EA 将会是什么。这个特定的 E 从行 2 减去 2 倍的行 1。其乘数是 $l=2$：

$$EA=\begin{bmatrix}1&0&0\\ \mathbf{-2}&\mathbf{1}&0\\ 0&0&1\end{bmatrix}\begin{bmatrix}\mathbf{2}&\mathbf{4}&\mathbf{-2}\\ \mathbf{4}&\mathbf{9}&\mathbf{-3}\\ -2&-3&7\end{bmatrix}=\begin{bmatrix}2&4&-2\\ \mathbf{0}&\mathbf{1}&\mathbf{1}\\ -2&-3&7\end{bmatrix}\quad \textit{(带有 0)} \tag{3}$$

该步骤不会改变 A 的行 1 和行 3。这些行未在 EA 中被改变——仅有行 2 不同了。**两倍的第一行被从第二行中减去**。矩阵乘法与消元一致——新的方程组就是 $EA\boldsymbol{x}=E\boldsymbol{b}$。

EA 很简单，但它包含一个微妙的思想。从 $A\boldsymbol{x}=\boldsymbol{b}$ 开始。等号两边乘上 E 得出 $E(A\boldsymbol{x})=E\boldsymbol{b}$。对于矩阵乘法，该式也是 $(EA)\boldsymbol{x}=E\boldsymbol{b}$。

第一个是 E 乘以 $A\boldsymbol{x}$，第二个是 EA 乘以 $\boldsymbol{x}$。它们是一样的。

括号不是必须的。我们只写 $EA\boldsymbol{x}$ 就好。

此规则拓展到具有多列向量的矩阵 C。当 ABC 相乘时，你可以先算 AC 或先算 EA。这是像 $3\times(4\times5)=(3\times4)\times5$ 的“结合律”要点。3 乘以 20 或 12 乘以 5。两个答案都是 60。这个法则好像如此明确，很难去设想它会是错的。

“交换律”$3\times4=4\times3$ 看上去甚至更显而易见。但是 EA 通常不同于 AE。当 E 乘在右边时，它作用于 A 的**列**——而不是行。AE 实际上是从列 1 减去两倍的列 2。因此 $EA\neq AE$。

结合律为真	$\boldsymbol{A(BC)=(AB)C}$
交换律为假	**通常** $\boldsymbol{AB\neq BA}$

矩阵乘法还有另一个要求。假设 B 只有一列（此列为 $\boldsymbol{b}$）。EB 矩阵—矩阵的乘法定律应与 $E\boldsymbol{b}$ 矩阵—向量乘法定律相同。更多地，我们应该能*一次乘以一列地将矩阵 EB 相乘*：

若 $\boldsymbol{B}$ 具有多个列 $\boldsymbol{b_1,b_2,b_3}$，则 $\boldsymbol{EB}$ 的列为 $\boldsymbol{Eb_1,Eb_2,Eb_3}$。

$$\textbf{矩阵乘法}\qquad AB=A[\boldsymbol{b_1}\ \boldsymbol{b_2}\ \boldsymbol{b_3}]=[A\boldsymbol{b_1}\ A\boldsymbol{b_2}\ A\boldsymbol{b_3}]. \tag{4}$$

这同样适用于（3）中的矩阵乘法。如果你将 E 乘以 A 的列 3，那么你就正确地得出 EA 的列 3：

$$\begin{bmatrix}1&0&0\\ -2&1&0\\ 0&0&1\end{bmatrix}\begin{bmatrix}-2\\ -3\\ 7\end{bmatrix}=\begin{bmatrix}-2\\ 1\\ 7\end{bmatrix}\qquad E(A\text{ 的列 }j)=EA\text{的列 }j\text{。}$$

这种方式要求处理列，然而消元法适用于行。**下一节描述每个乘积 AB 的各个元素**。矩阵乘法的美妙之处在于所有三种方法（*行，列，整个矩阵*）的结果都对。

行交换矩阵 P_{ij}

要从行 j 减去行 i 我们使用 E_{ij}。要交换或“置换”两列我们得用另一种矩阵 P_{ij}（置换矩阵）。当 0 在主元位时，行交换是必须的。在下方，主元列或许含有非零元素。通过交换两列，我们就有了主元并且消元得以继续。

什么矩阵 P_{23} 交换行 2 和行 3。我们可以通过交换单位矩阵 I 的行来求出它：

$$\textbf{置换矩阵}\qquad P_{23}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}。$$

这是一个**行交换矩阵**。乘上 P_{23} 可交换任一列向量的分量 2 与分量 3。因此它也能交换任一矩阵的行 2 与行 3：

$$\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1\\\mathbf{3}\\\mathbf{5}\end{bmatrix}=\begin{bmatrix}1\\\mathbf{5}\\\mathbf{3}\end{bmatrix}\quad 与 \quad \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}2&4&1\\\mathbf{0}&\mathbf{0}&\mathbf{3}\\0&6&5\end{bmatrix}=\begin{bmatrix}2&4&1\\0&6&5\\\mathbf{0}&\mathbf{0}&\mathbf{3}\end{bmatrix}。$$

在右边，P_{23} 在做其生而为之所为。由于 0 在第二个主元位且“6”在它下面，则交换行将 6 设为主元。

矩阵*采取行动*。它们不只是呆坐在那里。我们不久将接触其它排列矩阵，那些可改变多个列的顺序。行 $1,2,3$ 可移动到 $3,1,2$。我们的 P_{23} 是一个特定的排列矩阵——它交换行 2 与行 3。

行交换矩阵　　P_{ij} 是行 i 与行 j 颠倒的单位矩阵。当这个“**置换矩阵**”P_{ij} 乘以一个矩阵时，它交换行 i 与行 j。

$$\textbf{要交换方程 1 与方程 3 就乘上 } \boldsymbol{P_{13}}=\begin{bmatrix}\mathbf{0}&\mathbf{0}&\mathbf{1}\\\mathbf{0}&\mathbf{1}&\mathbf{0}\\\mathbf{1}&\mathbf{0}&\mathbf{0}\end{bmatrix}。$$

通常行交换不是必需的。消元法很大几率只用到 E_{ij}。但是就算需要将主元移动到对角线上，P_{ij} 也早就准备好了。

增广矩阵

本书最终远超消元范畴。乘上有各种特定应用的矩阵。我们的最佳出发点是一个方阵 E 乘以一个方阵 A，是因为我们在消元中遇见了它——并且我们知道 EA 期望的答案是什么。接下来一步是要接受一个*矩形矩阵*。它仍然来自于我们的原始方程，只是它包含了右边的 $\boldsymbol{b}$。

关键思想：消元法对 A 及 $\boldsymbol{b}$ 做相同的行运算。**我们可以囊括 $\boldsymbol{b}$ 使它成为额外的一列，然后经过消元来密切关注它**。矩阵 A 被扩大了或者说由额外的列 $\boldsymbol{b}$“增广了”：

$$\textbf{增广矩阵}\quad [A\ \boldsymbol{b}]=\begin{bmatrix}2&4&-2&\mathbf{2}\\4&9&-3&\mathbf{8}\\-2&-3&7&\mathbf{10}\end{bmatrix}。$$

消元法作用在这个矩阵的所有行上。为从方程 2 减去 2 倍的方程 1，等号左边和右边都乘上了 E。利用 $[A\ \boldsymbol{b}]$ 这些步骤一块起作用：

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 & \mathbf{2} \\ 4 & 9 & -3 & \mathbf{8} \\ -2 & -3 & 7 & \mathbf{10} \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & \mathbf{2} \\ 0 & 1 & 1 & \mathbf{4} \\ -2 & -3 & 7 & \mathbf{10} \end{bmatrix}。$$

新第二行包含 $0, 1, 1, 4$。新的第二个方程是 $x_2 + x_3 = 4$。矩阵乘法按行运算同时按列运算：

行　E 的各行作用于 $[A\ \boldsymbol{b}]$ 以得出 $[EA\ E\boldsymbol{b}]$ 的一行。

列　E 作用于 $[A\ \boldsymbol{b}]$ 的各列以得出 $[EA\ E\boldsymbol{b}]$ 的一列。

再次注意“行动”一词。这是必不可少的。矩阵做了一些事情！矩阵 A 作用于 $\boldsymbol{x}$ 以产生 $\boldsymbol{b}$。矩阵 E 对 A 运算得出 EA。消元的整个过程是一系列的行运算，也就是矩阵乘法。A 到 $E_{21}A$ 再到 $E_{32}E_{31}E_{21}A$。最后 $E_{32}E_{31}E_{21}A$ 是一个三角矩阵。

等号右边包含在增广矩阵中。最终结果是一个三角方程组。在写下对所有矩阵的乘法规则之前，我们停下来练习乘以 E。

■　复习关键点　■

1. $A\boldsymbol{x} = x_1$乘以列 $1 + \cdots + x_n$乘以列 n。且 $(A\boldsymbol{x})_i = \sum_{j=1}^{n} a_{ij}x_j$。
2. 单位矩阵 $= I$，利用 l_{ij} 的消元矩阵 $= E_{ij}$，置换矩阵 $= P_{ij}$。
3. 将 $A\boldsymbol{x} = \boldsymbol{b}$ 乘上 E_{21} 是从方程 2 减去 l_{21} 倍的方程 1。其中数 $-l_{21}$ 是消元矩阵 E_{21} 的 $(2, 1)$ 元素。
4. 对于增广矩阵 $[A\ \boldsymbol{b}]$，其消元一步得出 $[E_{21}A\ E_{21}\boldsymbol{b}]$。
5. 当 A 乘以任一矩阵 B 时，它分别乘以 B 的每一列。