麻省理工大学公开课MIT OpenCourseWarehttp://ocw.mit.edu

18.06 **线性代数** Linear Algebra

第一单元 *A*x=b 和四个子空间

Unit 1 *A*x=b & the Four Subspaces

麻省理工大学公开课MIT OpenCourseWarehttp://ocw.mit.edu

18.06 **线性代数** Linear Algebra

红色字体是我扯淡的部分，谢谢观赏……

在漫长的求学过程中，总有那么几本教材最后是以“天书”的形态永久停留在我们的记忆当中，而数学课本成为“天书”的概率显然远高于其它科目。各大高校都流传着类似于“实变函数学十遍，随机过程随机过”的革命口号。

在众多凶名在外的同僚映衬下，“线性代数”显得那么普通，但我想说它应该是理工科高等教育基础课中最被低估的一门，恐怕没有之一。

我念书的时候，线性代数是从行列式讲起的，现在想来这是有点莫名其妙的事情。一上来就接触行列式，无疑会把它当成一种运算，等第二章立刻又接触到矩阵代数的时候，有些同学就懵 B 了。

后来大概是认识到了行列式已经变得没有从前那般重要了，大家已经很少用行列式来判定矩阵是否可逆，或者用代数余子式来求取逆矩阵 balabala，于是师弟跟我说线代教材已经改了，现在是从线性空间学起，但我拿过来一翻还是一堆定理开路……颇有一种似曾相识的坑爹感，这跟微积分从实数集什么上界上确界开始讲起，基本属于一个路数。所以说天朝在数学教育上一直秉持的理念就是希望你能夯(zhi)实（nan）基（er）础（tui）。

再看看大牛 Gilbert Strang 是怎么教的，他首先说了一个我上完线代好几年才明白的事情，就是线代这门课的基础作用是教你解线性方程。他从列图像引入对于矩阵和线性相关性的理解，而将矩阵运算的核心定位于对“行”或“列”进行独立操作。在第一单元，主要强调高斯消元法的意义以及矩阵的四个子空间的概念（如他所言当你对一个矩阵不知该如何是好的时候，那就消它吧）。第二单元主要介绍正交、行列式和特征值，当我们从矩阵开始进入课程，而将行列式的值当成是矩阵的一种性质时，思维会顺畅许多。第三单元介绍正定矩阵等概念以及线性代数在各种领域的应用。他所做的一切都是具象化每一种数学操作，在讲解中让你体会到每一部分内容相对于线代核心所处的位置。

如果读了 GS 老先生的书你就会发现它与微分方程、傅里叶变换、工程数值计算等等工科数学之间有着紧密的联系，特别是 Matlab 使得这些联系愈加紧密，并体现在“应用”的诸多方面。GS 关于线代有两本著作，其中“Introduction to Linear Algebra”就是 MIT 本科线代课的教材，还有一本是“Linear Algebra and Its Applications”也出到第四版了，相比而言前者更基础。GS 写作教材都用最简单直接的英语，甚至尽量避免使用复杂句（这点比科技文献强多了），在理解上比之其他外文教材容易很多。这两本书也都有中文版就是不好找，前者只有台湾天下文化在

2005 年出了繁体翻译版，译名为《線性代數的世界》，后者是 90 年由南开大学出了中文版《线性代数及其应用》，译者侯自新，我所见的电子版应该是从超星流出来的。这两年一直风闻要出 G 的中文书了，不知道会是哪一本。此外，超星还有一本 83 年的译本《有限元法分析》，应该也是当年从俄文版本弄过来的，一般人我不告诉他。

附 G 老先生的主页：http://math.mit.edu/~gs/

douTintin

第 01 讲 行图像和列图像

Row picture & Column picture

线性方程的几何图像 The geometry of linear equations

线性代数的基本问题就是解 n 元一次方程组。例如：二元一次方程组

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

写成矩阵形式就是

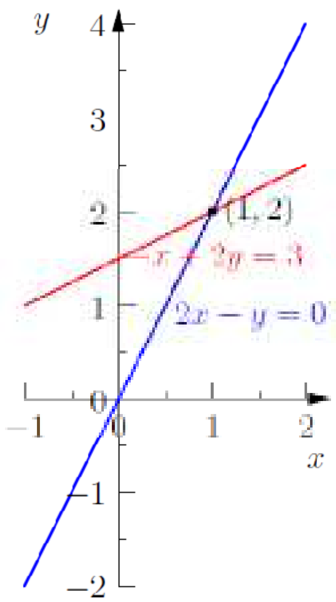
$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

其中 $\boldsymbol{A}=\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$被称为系数矩阵（coefficient matrix）。

未知数向量通常记为 $\mathbf{x}=\begin{bmatrix} x \\ y \end{bmatrix}$

而等号右侧的向量记为 **b**。线性方程组简记为 ***A*****x**=**b**。

行图像 Row Picture



行图像遵从解析几何的描述，每个方程在平面上的图像为一条直线。找到符合方程的两个数组，就可以确定出 x-y 平面上的两个点，连接两点可以画出该方程所代表的直线。两直线交点即为方程组的解 x=1,y=2。

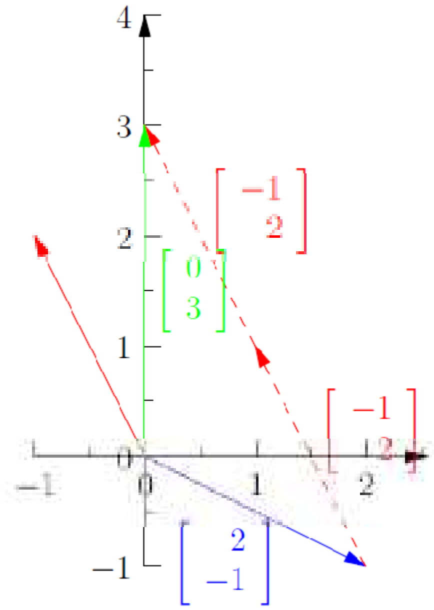
列图像 Column Picture

在列图像中，我们将系数矩阵写成列向量的形式，则求解原方程变为寻找列向量的线性组合（linear combination）来构成向量 **b**。

$$x\begin{bmatrix}2\\-1\end{bmatrix}+y\begin{bmatrix}-1\\2\end{bmatrix}=\begin{bmatrix}0\\3\end{bmatrix}$$

向量线性组合是贯穿本课程的重要概念。对于给定的向量 **c** 和 **d** 以及标量 x 和 y，我们将 x**c**+y**d** 称之为 **c** 和 **d** 的一个线性组合。

从几何上讲，我们是寻找满足如下要求的 x 和 y，使得两者分别数乘对应的列向量之后相加得到向量$\begin{bmatrix}0\\3\end{bmatrix}$。其几何图像如下图。

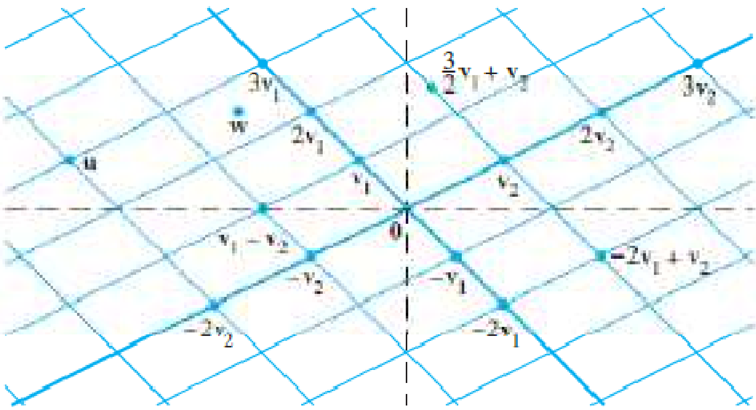


蓝色为向量$\begin{bmatrix}2\\-1\end{bmatrix}$；

红色为向量$\begin{bmatrix}-1\\2\end{bmatrix}$；

可以看到当蓝色的向量乘以 1 与红色的向量乘以 2（红色虚线）后做加法（首尾相接）就可以得到绿色的向量 **b**=$\begin{bmatrix}0\\3\end{bmatrix}$，由此可得到方程的解 x=1,y=2。

想象一下如果任意取 x,y，则得到的线性组合又是什么？其结果就是以上两个列向量的所有线性组合将会布满整个坐标平面。



D.C.Lay 的《线性代数及其应用》中，绘制向量$v_1=\begin{bmatrix}-1\\1\end{bmatrix}$和$v_2=\begin{bmatrix}1\\2\end{bmatrix}$的线性组合充满整个平面的图像，节点处为向量的整数倍线性组合。

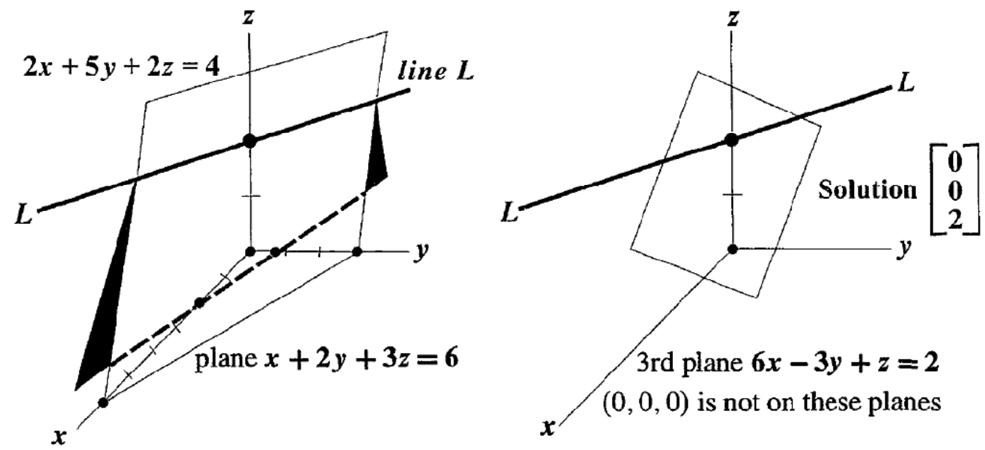
这本书也是难得的好书，作者喜欢利用几何图像来帮助读者理解线性代数中的

概念，英文版出到第 5 版了，华章出过中译本。（是不是觉得上面那个图片有点斜，是斜线造成的错觉呦！）

将以上讨论扩展到三元。图不好弄，所以用了 GS 书里的另一个方程，没有用视频中的那个！！！这样方程和配图是吻合的。

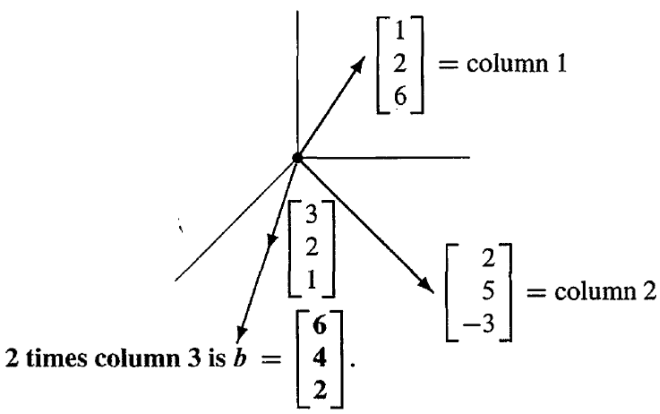
$$\begin{array}{rrrcl} x & +2y & +3z & = & 6 \\ 2x & +5y & +2z & = & 4 \\ 6x & -3y & +z & = & 2 \end{array} \qquad 矩阵形式\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 6 & -3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$$

方程的行图像比较复杂，每一个方程都是三维空间内的一个平面，方程组的解为三个平面的交点。



画图真不是 GS 的长项，在视频里画的就比较 shi，他自己也承认了。在课本里他用两个面相交于一条直线画了一个图，然后让这条直线和第三个平面相交画了第二个图。同样的事，D.C.Lay 一张图分分钟搞定。

方程组列图像为 $x\begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix} + z\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$



如果改变等号右侧的 **b** 的数值，那么对于行图像而言三个平面都改变了，而对于列图像而言，三个向量并没有发生变化，只是需要寻找一个新的组合。

那么问题来了，是否对于所有的 **b**，方程 ***A*x**=**b** 都有解？

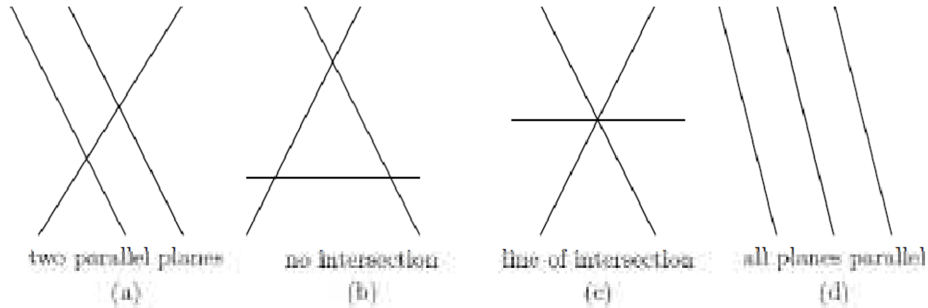
从列图像上看，问题转化为“列向量的线性组合是否覆盖整个三维空间？”

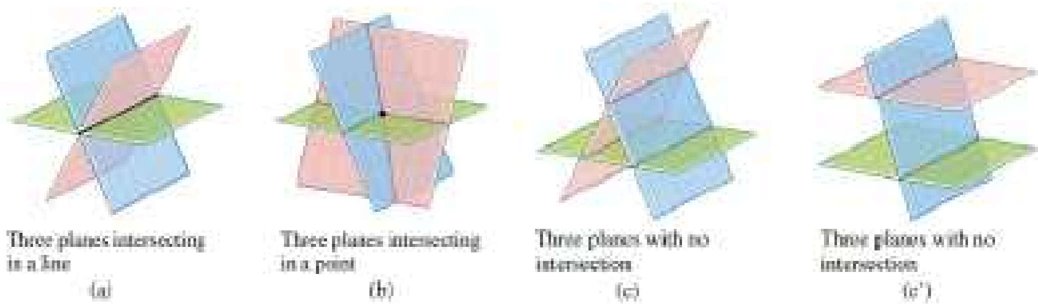
反例：若三个向量在同一平面内——比如“列 3”恰好等于“列 1”加“列 2”，而若 **b** 不在该平面内，则三个列向量无论怎么组合也得不到平面外的向量 **b**。此时矩阵 ***A*** 为奇异阵或称不可逆矩阵。在矩阵 ***A*** 不可逆条件下，不是所有的 **b** 都能令方程 ***A*x**=**b** 有解。

对 n 维情形则是，n 个列向量如果相互独立——“线性无关”，则方程组有解。否则这 n 个列向量起不到 n 个的作用，其线性组合无法充满 n 维空间，方程组未必有解。

从行图像的角度来看，三元方程组是否有解意味着什么？当方程所代表的三个平面相交于一点时方程有唯一解；三个平面中至少两个平行则方程无解；平面的两两交线互相平行方程也无解；三个平面交于一条直线则方程有无穷多解。

都是示意图，来看看 GS 和 Lay 的作图差异有多大吧……





矩阵与向量的乘法

列图像：***A*x** 是矩阵 ***A*** 列向量的线性组合 ：

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

也可以通过将矩阵 ***A*** 的行向量和 **x** 向量进行点积来计算：

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 2 \times 5 \\ 1 \times 1 + 2 \times 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

第 02 讲 矩阵消元

Elimination with matrices

消元法 Method of Elimination

消元法是计算机软件求解线形方程组所用的最常见的方法。任何情况下，只要是矩阵 $\boldsymbol{A}$ 可逆，均可以通过消元法求得 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的解。

此处给出的线性方程组为

$$\boldsymbol{A}=\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}\quad \mathbf{b}=\begin{bmatrix}2\\12\\2\end{bmatrix}$$

高斯消元法（Gauss elimination）就是通过对方程组中的某两个方程进行适当的数乘和加（jian）和（fa），以达到将某一未知数系数变为零，从而削减未知数个数的目的。

我们将矩阵左上角的 1 称之为“主元一”（the first pivot），第一步要通过消元将第一列中除了主元之外的数字均变化为 0。操作方法就是用之后的每一行减去第一行的适当倍数，此例中第二行应减去第一行的 3 倍。之后应对第三行做类似操作，本例中三行第一列数字已经为 0，故不用进行操作。

$$\boldsymbol{A}=\begin{bmatrix}\boxed{1}&2&1\\3&8&1\\0&4&1\end{bmatrix}\xrightarrow{(2,1)}\begin{bmatrix}\boxed{1}&2&1\\0&\boxed{2}&-2\\0&4&1\end{bmatrix}\xrightarrow{(3,2)}\mathrm{U}=\begin{bmatrix}\boxed{1}&2&1\\0&\boxed{2}&-2\\0&0&\boxed{5}\end{bmatrix}$$

处在第二行第二列的主元二为 2，因此用第三行减去第二行的两倍进行消元，得到第三个主元为 5。

矩阵 $\boldsymbol{A}$ 为可逆矩阵，消元结束后得到上三角阵 $\boldsymbol{U}$（Uppertriangular matrix），其左侧下半部分的元素均为 0，而主元 1,2,5 分列在 $\boldsymbol{U}$ 的对角线上。主元之积即行列式的值。

需要说明的是，主元不能为 0，如果恰好消元至某行，0 出现在了主元的位置上，应当通过与下方一行进行“行交换”使得非零数字出现在主元位置上。如果 0 出现在了主元位置上，并且下方没有对等位置为非 0 数字的行，则消元终止，并证明矩阵 A 为不可逆矩阵，且线性方程组没有唯一解。

$$\text{例如消成这样}\begin{bmatrix}*&*&*\\0&*&*\\0&0&0\end{bmatrix}$$

回代 Back-Substitution

做方程的高斯消元时，需要对等式右侧的 **b** 做同样的乘法和加减法。手工计算时比较有效率的方法是应用“增广矩阵”（augmented matrix），将 **b** 插入矩阵 ***A*** 之后形成最后一列，在消元过程中带着 **b** 一起操作。（Matlab 是算完系数矩阵再处理 **b** 的。）

$$\begin{bmatrix}1&2&1&2\\3&8&1&12\\0&4&1&2\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1&2\\0&2&-2&6\\0&4&1&2\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1&2\\0&2&-2&6\\0&0&5&-10\end{bmatrix}$$

此时我们将原方程 ***A*x**=**b** 转化为了新的方程 ***U*x**=**c**，其中 **c**=$\begin{bmatrix}2\\6\\-10\end{bmatrix}$。

从最后一行得到 z=-2，依次回代可以得到 y=1,和 x=2。

以上高斯消元法的内容基本是回忆求解线性方程的步骤，是我们很熟悉的东西。在线性代数中比较重要的就是将之前所说的“第二行减去第一行的 3 倍”这种操作条例变为矩阵化的数学语言。

消元矩阵 Elimination Matrices

矩阵运算的核心内容就是对“行”或者“列”进行独立操作。

如前一节课“列图像”部分所言，系数矩阵乘以未知数向量，相当于对系数矩阵的列向量进行线性组合。

例如
$$\begin{bmatrix}*&*&*\\*&*&*\\*&*&*\end{bmatrix}\begin{bmatrix}3\\4\\5\end{bmatrix}=3\underset{\text{列 1}}{\begin{bmatrix}*\\*\\*\end{bmatrix}}+4\underset{\text{列 2}}{\begin{bmatrix}*\\*\\*\end{bmatrix}}+5\underset{\text{列 3}}{\begin{bmatrix}*\\*\\*\end{bmatrix}}=\begin{bmatrix}\otimes\\\otimes\\\otimes\end{bmatrix}$$

与之相对称，矩阵左乘行向量则是对矩阵的行向量进行线性组合。

例如
$$\begin{bmatrix}1&2&7\end{bmatrix}\begin{bmatrix}*&*&*\\*&*&*\\*&*&*\end{bmatrix}=\begin{array}{c}1\underset{\text{行 1}}{\begin{bmatrix}*&*&*\end{bmatrix}}\\+\\2\underset{\text{行 2}}{\begin{bmatrix}*&*&*\end{bmatrix}}\\+\\7\underset{\text{行 3}}{\begin{bmatrix}*&*&*\end{bmatrix}}\end{array}=\begin{bmatrix}\oplus&\oplus&\oplus\end{bmatrix}$$

我学这部分的时候，列向量的线性组合非常容易就接受了，左乘行向量这个总觉得很别扭，偏偏行向量在下面介绍消元矩阵时比较重要。后来想想觉得“列”操作就像是把向量开进矩阵，而“行操作”这个就像把向量倒车进入矩阵（如图中箭头所示）。-_-! 虽然挺扯淡的，但是我反正突然就觉得习惯了。

矩阵消元的第一步是通过左乘矩阵 E_{21} 来实现原矩阵 ***A*** 的第二行减去第一行的

3 倍这一过程。E_{21}的第二行使矩阵 A 的行向量进行前述的线性组合，而其它两行为了保持与原矩阵相同，采用同阶单位阵 I 的行向量。左乘的这个矩阵为“初等矩阵”（Elementary Matrix），因此记做 E。我以为是消元矩阵，所以记做 E 呢。因为所乘行向量的倍数-3 出现在 E 矩阵的第二行第一列，因此将之标注为 21。完成操作后矩阵变为 $E_{21}A$。

$$\text{第一行：}\begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix}=\begin{bmatrix}1 & 2 & 1\end{bmatrix}$$

$$\text{第三行：}\begin{bmatrix}0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix}=\begin{bmatrix}0 & 4 & 1\end{bmatrix}$$

$$\text{关键第二行：}\begin{bmatrix}-3 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix}=\begin{matrix} & -3 & [1 & 2 & 1]\\ + & 1 & [3 & 8 & 1]\\ + & 0 & [0 & 4 & 1]\end{matrix}=\begin{bmatrix}0 & 2 & -2\end{bmatrix}$$

$$\underset{E_{21}}{\begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\0 & 0 & 1\end{bmatrix}}\underset{A}{\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix}}=\underset{E_{21}A}{\begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix}}$$

矩阵消元的第二步是完成矩阵 $E_{21}A$ 的第三行减去第二行的 2 倍，通过左乘矩阵 E_{32}来实现这一过程。

$$\underset{E_{32}}{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -2 & 1\end{bmatrix}}\underset{E_{21}A}{\begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix}}=\underset{E_{32}(E_{21}A)}{\begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix}}$$

3×3 矩阵的消元本来应该分三步完成，最终得到 $E_{32}(E_{31}(E_{21}A))$。本例中 $E_{31}=I$，所以结果变为 $E_{32}(E_{21}A)=U$，因为矩阵运算符合结合律，也可写作$(E_{32}E_{21})A=U$。可以记作 $EA=\mathbf{U}$。

方程 $A\mathbf{x}=\mathbf{b}$ 的解也满足方程 $U\mathbf{x}=EA\mathbf{x}=E\mathbf{b}=\mathbf{c}$，因此我们将问题转化为 $U\mathbf{x}=\mathbf{c}$。

置换矩阵 Permutation

左乘置换矩阵可以完成原矩阵的行变换，右乘置换矩阵则为列变换。例如

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}a & b\\c & d\end{bmatrix}=\begin{bmatrix}c & d\\a & b\end{bmatrix}$$

$$\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}=\begin{bmatrix}b & a\\d & c\end{bmatrix}$$

对于三阶矩阵，$\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}=\begin{bmatrix}g&h&i\\d&e&f\\a&b&c\end{bmatrix}$

构造 ***P*** 矩阵是通过对 ***I*** 矩阵进行“行交换”来实现的。

左右乘效果不同也展示了矩阵运算不符合交换律的性质。

逆矩阵 Inverse

这里主要讨论消元矩阵的逆矩阵。消元矩阵之逆矩阵的实施效果就是抵消原矩阵的消元操作。消元矩阵实现了对原矩阵 ***A*** 的操作，使第二行行向量[3,8,1]减掉了第一行[1,2,1]的 3 倍变为[0,2,-2]，则逆向操作就应该是把现在的第二行行向量[0,2,-2]加上第一行[1,2,1]的 3 倍，从而变回原来的第二行[3,8,1]。

所以对于 $\boldsymbol{E_{21}}=\begin{bmatrix}1&0&0\\-3&1&0\\0&0&1\end{bmatrix}$，有 $\boldsymbol{E_{21}}^{-1}=\begin{bmatrix}1&0&0\\3&1&0\\0&0&1\end{bmatrix}$。

满足 $\boldsymbol{E_{21}}^{-1}\,\boldsymbol{E_{21}}=\boldsymbol{I}$ $\begin{bmatrix}1&0&0\\3&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\-3&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$

第 03 讲 矩阵的乘法和逆矩阵

Multiplication & inverse matrices

矩阵乘法 Matrix multiplication

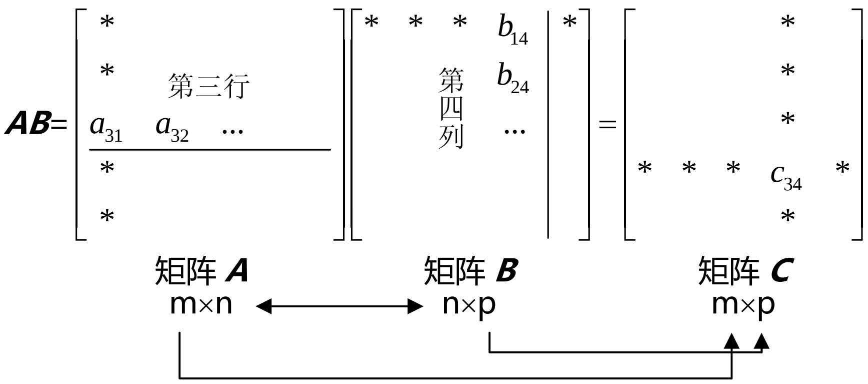
我们通过四种方法讨论如何使矩阵 ***A*** 与 ***B*** 相乘得到矩阵 ***C***。其中 ***A*** 为 m×n(m 行 n 列) 矩阵，而 ***B*** 为 n×p 矩阵，则 ***C*** 为 m×p 矩阵，记 c_{ij} 为矩阵 ***C*** 中第 i 行第 j 列的元素。

1)标准方法 (行乘以列)

矩阵乘法的标准计算方法是通过矩阵 ***A*** 第 i 行的行向量和矩阵 ***B*** 第 j 列的列向量点积得到 c_{ij}。

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots\ldots$$

举例



$$c_{34} = row3 \bullet col4 = \sum_{k=1}^{n} a_{3k} b_{k4} = a_{31}b_{14} + a_{32}b_{24} + a_{33}b_{34} + \ldots\ldots$$

所谓标准方法就是矩阵乘法的运算规则，下面扯几句矩阵代数运算的淡。念书的时候，学到矩阵代数，同桌问我：“矩阵运算怎么这么奇怪，只有同样型号的矩阵（指同为 m×n ）可以做加法，但是不同型号的矩阵却可以做乘法，话说这不同型号的矩阵为啥要做这种复杂的乘法？数乘一个矩阵为啥其中每一个元素都要乘以这个数，而行列式则不是？” 我后来找到了 A.D.亚历山大洛夫编著的《数学：它的内容，方法和意义》第三卷中找到了答案。

先补一下矩阵的加法和数乘规则：

Addition $A+B$

Multiplication $2A$

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 1 \\ 1 & 6 \end{bmatrix} \qquad 2\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & 0 \\ 0 & 8 \end{bmatrix}.$$

本质上矩阵的代数运算就是线性方程组的运算：

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \quad\ldots\ldots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{cases}$$

可以简写做矩阵形式 **y=*A*x**
A 为 m×n 矩阵

$$\begin{cases} z_1 = b_{11}x_1 + \ldots + b_{1n}x_n \\ \quad\ldots\ldots \\ z_m = b_{m1}x_1 + \ldots + b_{mn}x_n \end{cases}$$

z=*B*x

B 为 m×n 矩阵

$$\begin{cases} y_1 + z_1 = (a_{11} + b_{11})x_1 + \ldots + (a_{1n} + b_{1n})x_n \\ \quad\ldots\ldots \\ y_m + z_m = (a_{m1} + b_{m1})x_1 + \ldots + (a_{mn} + b_{mn})x_n \end{cases}$$

y+z=(*A*+*B*)x 所以矩阵加法运算规则就只能形状相同的矩阵进行加和

$$\begin{cases} \alpha y_1 = \alpha a_{11}x_1 + \ldots + \alpha a_{1n}x_n \\ \quad\ldots\ldots \\ \alpha y_m = \alpha a_{m1}x_1 + \ldots + \alpha a_{mn}x_n \end{cases}$$

α**y**=α***A*x**

数乘要在矩阵每个元素上乘以 α

而矩阵乘法可以视为给线性方程组做变量替换

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \quad\ldots\ldots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{cases}$$

y=*A*x　***A*** 为 m×n 矩阵

$$\begin{cases} w_1 = d_{11}y_1 + \ldots + d_{1m}y_m \\ \quad\ldots\ldots \\ w_k = d_{k1}y_1 + \ldots + d_{km}y_m \end{cases}$$

w=*D*y　***D*** 为 k×m 矩阵

将左侧 y_1 至 y_m 的表达式代入右侧的方程，则有

$$\begin{cases} w_1 = d_{11}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + d_{1m}(a_{m1}x_1 + \ldots + a_{mn}x_n) \\ \quad\ldots\ldots \\ w_k = d_{k1}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + d_{km}(a_{m1}x_1 + \ldots + a_{mn}x_n) \end{cases}$$

重新整理一下

$$\begin{cases} w_1 = (d_{11}a_{11} + \ldots + d_{1m}a_{m1})x_1 + \ldots + (d_{11}a_{1n} + \ldots + d_{1m}a_{mn})x_n = c_{11}x_1 + \ldots + c_{1n}x_n \\ \quad\ldots\ldots \\ w_k = (d_{k1}a_{11} + \ldots + d_{km}a_{m1})x_1 + \ldots + (d_{k1}a_{1n} + \ldots + d_{km}a_{mn})x_n = c_{k1}x_1 + \ldots + c_{kn}x_n \end{cases}$$

w=*D*y=*DA*x=*C*x

***C*=*DA* 为** k×n 矩阵，从上式可得到 c_{ij} 的公式，即矩阵的乘法运算规则，可以看到这是变量代换的结果。所以矩阵乘法中两矩阵形状可以不同，但要求左侧矩阵的列数要等于右侧矩阵的行数，即 **w=*D*y** 中 **y** 的分量个数要等于 **y=*A*x** 中 **y** 的分量数，不如此则线性方程的变量代换无法进行。

矩阵乘法符合结合律、分配率，但是不遵守交换律，这些通常通过定义加以证

明，即将等号两侧矩阵的元素乘开，比较对应元素，从而得出相等或者不等的结论。我们也可以从变量替换的角度思考一下矩阵乘法运算定律。比如结合律，若方程组有如下关系 **u**=***E*w**，**w**=***D*y**，**y**=***A*x**，则做变量替换可有 **u**=***E*w**=***ED*y**=***EDA*x**。结合律为(***ED***)***A***=***E***(***DA***)，它表达的意思就是替换环节无论是先把 **u** 表示成 ***ED*y**，再表示成 **x** 的表达式，又或者先把 **w** 表示成 ***DA*x**，再把公式带入 **u**=***E*w** 都是等效的。分配率也一样，先做加法再进行变量替换，或者先做变量替换再相加也是等效的。但是交换律就不行了，别说矩阵的尺寸可能导致交换顺序后不能进行乘法运算，即使能进行乘法，变量替换的关系也完全不对了。

2)列操作

列操作是指矩阵 ***C*** 的第 j 列是通过矩阵 ***A*** 乘以矩阵 ***B*** 第 j 列的列向量得到的。这表明矩阵 ***C*** 的列向量是矩阵 ***A*** 列向量的线性组合，组合的“权”就是矩阵 ***B*** 第 j 列的各个分量。

$$AB = A\begin{bmatrix} b_1 & b_2 & \dots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix}$$

上式中 $\mathbf{b}_i$，$\boldsymbol{A}\mathbf{b}_i$ 均为列向量，其中

$$Ab_j = \begin{bmatrix} | & | & & & | \\ a_1 & a_2 & \cdots & \cdots & a_n \\ | & | & & & | \end{bmatrix}\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ \vdots \\ b_{nj} \end{bmatrix} = b_{1j}\begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + b_{2j}\begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} + \cdots + b_{nj}\begin{bmatrix} | \\ a_n \\ | \end{bmatrix}$$

为矩阵 ***A*** 中列向量的线性组合

3)行操作

行操作是指矩阵 ***C*** 的第 i 行是通过矩阵 ***A*** 的第 i 行乘以矩阵 ***B*** 得到的。这表明矩阵 ***C*** 的行向量是矩阵 ***B*** 行向量的线性组合。

$$AB = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_m \end{bmatrix} B = \begin{bmatrix} a_1B \\ a_2B \\ \vdots \\ \vdots \\ a_mB \end{bmatrix}$$ 此处 $\mathbf{a}_i$ 和 $\mathbf{a}_i\boldsymbol{B}$ 为行向量

$$a_iB = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}\begin{bmatrix} \text{——} & b_1 & \text{——} \\ \text{——} & b_2 & \text{——} \\ & \vdots & \\ & \vdots & \\ \text{——} & b_n & \text{——} \end{bmatrix} = \begin{matrix} a_{i1}\begin{bmatrix} \text{——} & b_1 & \text{——} \end{bmatrix} \\ + \\ a_{i2}\begin{bmatrix} \text{——} & b_2 & \text{——} \end{bmatrix} \\ \vdots \\ + \\ a_{in}\begin{bmatrix} \text{——} & b_n & \text{——} \end{bmatrix} \end{matrix}$$

4)列乘以行

矩阵 $\boldsymbol{A}$ 的第 k 列是一个 m×1 的向量，而矩阵 $\boldsymbol{B}$ 第 k 行是一个 1×p 的向量，两向量相乘会得到一个矩阵 $\boldsymbol{C}_k$，将所有的 n 个矩阵相加记得到 $\boldsymbol{C}$。

$$AB=\sum_{k=1}^{n}\underbrace{\begin{bmatrix}a_{1k}\\ \vdots\\ a_{mk}\end{bmatrix}}_{m\times 1}\underbrace{\begin{bmatrix}b_{k1} & \cdots & b_{kp}\end{bmatrix}}_{1\times p}$$

可以从矩阵乘法对加法的分配率推导出来。

分块乘法 block multiplication

如果将矩阵 $\boldsymbol{A}$ 和矩阵 $\boldsymbol{B}$ 划分为严格匹配的区块，则矩阵乘法可以通过分块的乘法加以实现。

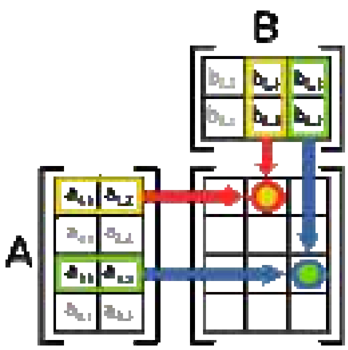
$$\begin{bmatrix}A_1 & A_2\\ A_3 & A_4\end{bmatrix}\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}=\begin{bmatrix}C_1 & C_2\\ C_3 & C_4\end{bmatrix}$$

其中 $\boldsymbol{C}_1=\boldsymbol{A}_1\boldsymbol{B}_1+\boldsymbol{A}_2\boldsymbol{B}_3$，计算方法与标准算法中矩阵里元素的操作方式相同。

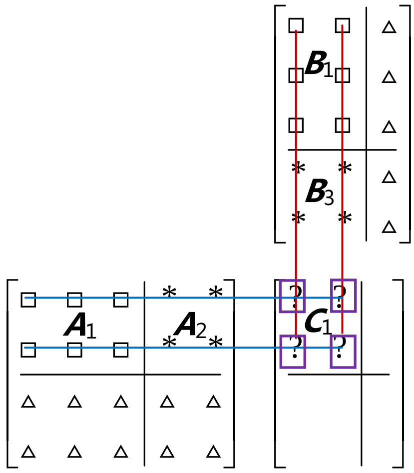
之前没有提到一点，就是因为矩阵乘法规则有点小复杂，在手算计算过程中有可能会串行，我们不可能每一次都如下图这样认真标注。

$$AB=\begin{bmatrix}a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\end{bmatrix}\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\\ b_{41} & b_{42}\end{bmatrix}=\begin{bmatrix}* & *\\ * & *\\ * & (AB)_{32}\end{bmatrix}$$

一个比较好的小技巧就是把 $\boldsymbol{B}$ 矩阵的位置改变一下，这样矩阵 $\boldsymbol{C}$ 中每一个元素所对应 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中的行与列就变得非常清楚了。我在 wiki 百科中看到了这张图，MIT 多元微积分课程上老师也提到了这种方法。



这里我们可以利用这种小技巧来帮助理解矩阵的分块乘法：



可以看到 C_1 就是经过 $C_1=A_1B_1+A_2B_3$ 的运算计算出来的，A_1 中的元素只和 B_1 的元素进行运算，A_2 只和 B_3 进行运算，C_1 为两者加和。

逆矩阵 Inverse

如果矩阵 A 是方阵，若存在逆矩阵 A^{-1}，使得 $A^{-1}A=I=A\,A^{-1}$（左逆矩阵等于右逆矩阵）。我们称矩阵 A 可逆（invertible）或者矩阵 A 非奇异（nonsingular）。反之，如果 A 为奇异（singular），则其没有逆矩阵。它的行列式为 **0**。另一个等价的说法是，A 为奇异阵，则方程 $A\mathbf{x}=\mathbf{0}$ 存在非零解 **x**。例如：

$$\begin{bmatrix}1 & 3\\ 2 & 6\end{bmatrix}\begin{bmatrix}-3\\ 1\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

在这个二阶矩阵的例子中，两个列向量排列在同一方向上。不可逆矩阵中总有列向量对生成线性组合没有贡献，等价的说法还有：不可逆矩阵的列向量可以通过线性组合得到 **0**。

换而言之，若矩阵 A 存在逆矩阵，则方程 $A\mathbf{x}=\mathbf{0}$ 只有零解。证明：反设其存在非零解 **x**，则有 $\mathbf{x}=(A^{-1}A)\mathbf{x}=A^{-1}(A\mathbf{x})=A^{-1}\mathbf{0}=\mathbf{0}$，矛盾。

对于可逆矩阵，求它的逆矩阵是一个重要的问题。

$$\underset{A}{\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}}\underset{A^{-1}}{\begin{bmatrix}a & c\\ b & d\end{bmatrix}}=\underset{I}{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}}$$

从“列操作”的角度来看，求逆矩阵过程其实和求 $A\mathbf{x}=\mathbf{b}$ 相同，只是这里 **x** 为矩阵 A^{-1} 的第 j 列，而 **b** 为单位阵 I 的第 j 列。

高斯-若尔当消元法 Gauss-Jordan Elimination

对于前面的二阶矩阵，求逆相当于两组方程：

$$\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}\text{和}\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}\begin{bmatrix}c\\ d\end{bmatrix}=\begin{bmatrix}0\\ 1\end{bmatrix}$$

Gauss-Jordan 消元法可以同时处理两个方程：

$$\underset{\boldsymbol{A}\qquad\quad\boldsymbol{I}}{\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right]} \to \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right] \to \underset{\boldsymbol{I}\qquad\quad\boldsymbol{A}^{-1}}{\left[\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right]}$$

在用高斯消元法的到上三角矩阵之后，按照若尔当的做法继续消元，用第一行减去第二行的若干倍，最后原矩阵变为单位阵，这时右侧的矩阵即为逆矩阵。

对 $\boldsymbol{A}$ 进行一系列消元操作，相当于左乘消元矩阵 $\boldsymbol{E}$，此时消元的结果为 $\boldsymbol{EA}=\boldsymbol{I}$，因为右侧矩阵也进行了同样消元操作，也等于左乘矩阵 $\boldsymbol{E}$，则右侧矩阵为 $\boldsymbol{EI}=\boldsymbol{E}$。由 $\boldsymbol{EA}=\boldsymbol{I}$ 可知这里的 $\boldsymbol{E}$ 就是 $\boldsymbol{A}$ 的逆矩阵，因此右侧矩阵给出的就是逆矩阵 $\boldsymbol{E}=\boldsymbol{A}^{-1}$。

$$E\left[A \mid I\right]=\left[EA \mid EI\right]=\left[I \mid A^{-1}\right]$$

第 04 讲 矩阵的 LU 分解

Factorization into $A=LU$

本节的主要目的是从矩阵的角度理解高斯消元法，最后找到所谓的 L 矩阵，使得矩阵 A 可以转变为上三角阵 U。即完成 LU 分解得到 $A=LU$。

首先继续了解一些矩阵乘法和逆矩阵的相关内容。

矩阵乘积的逆矩阵 Inverse of a product

（AB）$^{-1}=B^{-1}A^{-1}$

$ABB^{-1}A^{-1}=A(BB^{-1})A^{-1}=I$ 比较等式两端可得（AB）$^{-1}=B^{-1}A^{-1}$。

矩阵乘积的转置 Transpose of a product

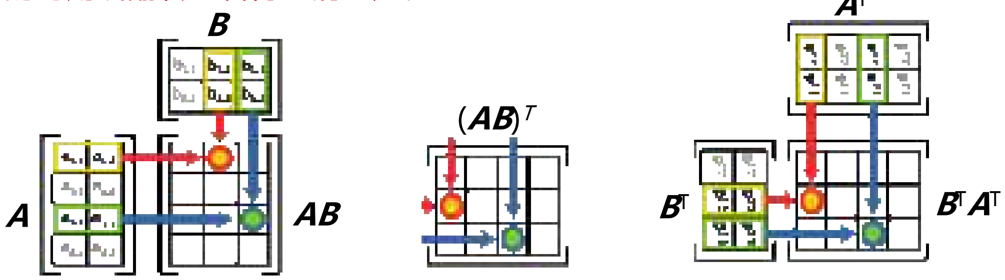
（本讲要用到的只是转置矩阵求逆的公式，把相关内容放在这里做个简介，方便大家理解。GS 在下一讲还会讲到转置。）

矩阵 A 的转置矩阵记为 A^T，对矩阵进行转置就是将 A 矩阵的行变为 A^T 的列，则完成后 A 的列也就成为了 A^T 的行，看起来矩阵如同沿着对角线进行了翻转。

$$A=\begin{bmatrix}2 & 1 & 4\\ 0 & 0 & 3\end{bmatrix} \quad 则\ A^T=\begin{bmatrix}2 & 0\\ 1 & 0\\ 4 & 3\end{bmatrix}$$

其数学表达式为 $(A^T)_{ij}=A_{ji}$，即 A^T 的第 i 行 j 列的元素为原矩阵 A 中第 j 行 i 列的元素。

$(AB)^T=B^TA^T$ 可以用定义证明。把上次课介绍的那个乘法小技巧的图片按照乘积矩阵的对角线翻转也看得一清二楚。



转置矩阵的逆矩阵 Inverse of a transpose

$(A^T)^{-1}=(A^{-1})^T$

$AA^{-1}=I$ 两边取转置，并应用积的转置公式，得到 $(A^{-1})^TA^T=I$。根据逆矩阵定义得到 $(A^T)^{-1}=(A^{-1})^T$

矩阵的 LU 分解

我们已经看到了如何应用消元法将矩阵 A 转变为上三角阵 U，这就引出了矩阵

的 ***LU*** 分解，它是理解矩阵 ***A*** 性质的重要方法。可以将矩阵的分解类比为多项式的因式分解，分解后的结果可以让我们更容易看清“解”的状态。

在没有行交换的情况下，矩阵 A 通过左乘一系列消元矩阵 E_{ij} 可以转化为 ***U***。在二阶矩阵中，进行一次消元操作即可达到这一效果。

$$\underset{E_{21}}{\begin{bmatrix}1 & 0\\ -4 & 1\end{bmatrix}}\underset{A}{\begin{bmatrix}2 & 1\\ 8 & 7\end{bmatrix}}=\underset{U}{\begin{bmatrix}2 & 1\\ 0 & 3\end{bmatrix}}$$

在等式两侧左乘，得到 $E_{21}^{-1}E_{21}A=E_{21}^{-1}U$ 即 $A=E_{21}^{-1}U$。就得到了矩阵 ***A*** 的 ***LU*** 分解结果。E_{21}^{-1} 是 E_{21} 的逆向操作，即左乘 E_{21} 中使得矩阵 ***A*** 第二行[8,7]减去第一行的 4 倍得到“新第二行”[0,3]，那么再左乘 E_{21}^{-1} 可以使得”新第二行”[0,3]加上“第一行”的 4 倍又变回原第二行的数值[8,7]。

$$\underset{A}{\begin{bmatrix}2 & 1\\ 8 & 7\end{bmatrix}}=\underset{L}{\begin{bmatrix}1 & 0\\ 4 & 1\end{bmatrix}}\underset{U}{\begin{bmatrix}2 & 1\\ 0 & 3\end{bmatrix}}$$

其中 ***U*** 为上三角阵（Upper triangular matrix），主元依次排列于它的对角线上，E_{21}^{-1} 即 ***L*** 为下三角阵（Lower triangular matrix）。有时我们也通过分解得到对角阵 ***D***（diagonal matrix）,例如

$$\underset{A}{\begin{bmatrix}2 & 1\\ 8 & 7\end{bmatrix}}=\underset{L}{\begin{bmatrix}1 & 0\\ 4 & 1\end{bmatrix}}\underset{D}{\begin{bmatrix}2 & 0\\ 0 & 3\end{bmatrix}}\underset{U'}{\begin{bmatrix}1 & 1/2\\ 0 & 1\end{bmatrix}}$$

对于三阶矩阵不需要换行进行消元的情况则有：

$E_{32}(E_{31}(E_{21}A))=U$，左乘逆矩阵可得 $A=E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}U=LU$

设定一组消元矩阵，其中 E_{31} 为单位阵 ***I***，其它两个消元矩阵如下：

$$\underset{E_{32}}{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -5 & 1\end{bmatrix}}\underset{E_{21}}{\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}=\underset{E}{\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 10 & -5 & 1\end{bmatrix}}$$

可以看到在消元矩阵 E 的左下角出现了数 10。它的出现是由于第一步操作 E_{21} 中“第二行”减去了 2 倍的“第一行”得到了“新第二行”。$row_2-2row_1=newrow_2$

而在第二步操作 E_{32} 中第三行减去了 5 倍的“新第二行”，这相当于减去 5 倍的“原第二行”并减去 5 倍的“(-2)倍原第一行”，**10** 就出现在这里。

$row_3-5newrow_2=row_3-5(row_2-2row_1)=row_3-5row_2+\mathbf{10}row_1$

在右侧操作则不会有这种情况发生，运算顺序会发生变化，$L=E^{-1}=E_{21}^{-1}E_{32}^{-1}$。

$$\underset{E_{21}^{-1}}{\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}\underset{E_{32}^{-1}}{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 5 & 1\end{bmatrix}}=\underset{L}{\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 0 & 5 & 1\end{bmatrix}}$$

如果没有行交换操作，则消元矩阵的因子可以直接写入矩阵 ***L***。没有多余的交叉项出现是 ***LU*** 分解要优于 ***EA***= ***U*** 这种形式的原因之一。

消元法所需运算量

在一些应用中我们需要处理超大型矩阵，即使用计算机来处理这一问题，也需要评估所需的计算量。

如果我们把“先乘后减”大致记为一次运算，那么对于一个 n×n 矩阵，对于一行进行消元要进行 n 次运算，由于有 n 行所以进行了 n^2 次运算，结果得到了第一列除第一主元外都消成 0 的矩阵。随后开始对除第一行第一列之外的剩余部分进行消元，这相当于一个 n-1×n-1 的矩阵，那么就要进行$(n-1)^2$ 次运算，以此类推。

最后需要的运算次数为 $1^2+2^2+......+n^2$，利用积分公式可以估算其数值。

$$1^2+2^2+...+n^2=\sum_{i=1}^{n} i^2 \approx \int_0^n x^2 dx = \frac{1}{3}n^2$$

算是个灵活使用数学工具的例子，我看到完全平方的求和，就一直在回忆公式，但其实作为估算采用积分公式就很方便了。有时候应用数学工具就像开车，保证你正常行驶主要靠驾驶技能，而不是你手边的一本《交规》，熟练驾驶但是不遵守交规很容易出错，交规倒背如流却没有驾驶技能则往往寸步难行。

再扯几句淡。在统计学中经常用到斯特林公式（Stirling's approximation），它用来评估和近似阶乘的大小 $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$，如果你去阅读 Stirling 公式的发展历程和证明过程，就会意识到级数、极限和微积分中间的紧密联系，然而在大多数同学那里这些概念都是割裂开的，我觉得这就是普通人和学霸有差距的地方，知识点互相联系不起来。Stirling 公式的证明过程也用到了积分，感兴趣可以看一看。

等号右侧向量 **b** 的行变换大致需要 n^2 次运算。

行转换 Row exchanges

如果主元的位置出现了 0，就需要进行“行交换”。我们可以通过左乘一个置换矩阵（Permutation Matrix）实现“行交换”的操作。例如

$$\boldsymbol{P}_{21}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$$

可以实现 3×3 矩阵的第一行与第二行的交换。所有的 3×3 的置换矩阵包括：

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$$

对于 n×n 矩阵存在着 n！个置换矩阵。置换矩阵每一行或者每一列只有一个元素是 1，其它都是 0，从第一行选一个位置设定为 1 有 n 个选择，第二行则只剩下 n-1 个选择，以此类推，最终有 n！种可能。

对于某阶的置换矩阵集合而言，置换矩阵的两两乘积仍在这个集合中，置换矩阵的逆矩阵也在此集合中。置换矩阵的逆矩阵即为它的转置 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{T}$。

可以用前面介绍过的乘法运算小技巧来理解这件事情

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \boldsymbol{P}^{-1}$$

$$\boldsymbol{P}=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$$

$\boldsymbol{P}$ 的第 i 行和 $\boldsymbol{P}^{-1}$ 的第 i 列相乘会得到 1，与其他列相乘都得到 0，所以 $\boldsymbol{P}^{-1}$ 第 i 列只能是 $\boldsymbol{P}$ 的第 i 行行向量的转置（让分量 1 出现在向量里的相同位置），其他各列以此类推，则 $\boldsymbol{P}$ 的每一个行向量转置就得到 $\boldsymbol{P}^{-1}$ 的列向量，这也就是矩阵转置 $\boldsymbol{P}^{T}$ 的定义。因此可得 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{T}$。

第 05 讲 转置、置换和空间

Transposes, permutations, spaces $\mathbf{R}^n$

本节的将引入向量空间（vector spaces）和子空间（subspaces）。

置换 Permutations

当应用消元法求解方程组的时候我们需要通过行交换将 0 从主元位置移走。左乘一个置换阵可以实现行交换的操作。（为了满足数值计算的要求，Matlab 甚至会对接近于 0 的非零主元做行交换）因此我们的 LU 分解由 $A=LU$ 变为 $PA=LU$。其中的 P 就是对 A 的行向量进行重新排序的置换矩阵。

我之前看到这里的时候有个疑惑，如果在消元的过程中进行“行交换”，那么不是意味着 P 矩阵存在于一堆 E 矩阵之间么 $EEEPEPEA=U$ 矩阵乘法不符合交换律，你怎么能把 P 提出来变成 $PA=LU$ 呢。同桌说可以假想在消元过程中你已经完全知道需要怎么进行“行交换”之后，我们重新开始做矩阵分解，这一次先对 A 进行“行交换”得到 A^*，这时候对 A^* 消元，就不用再进行“行交换”了，于是有 $PA=A^*=LU$。

置换矩阵 P 是通过对单位阵进行“行交换”得到的。对于 $n\times n$ 矩阵存在着 $n!$ 个置换矩阵。置换矩阵具有特殊性质 $P^{-1}=P^T$ 即 $P^TP=I$。

转置 Transposes

矩阵 A 的转置矩阵记为 A^T，对矩阵进行转置就是将 A 矩阵的行变为 A^T 的列，则完成后 A 的列也就成为了 A^T 的行，看起来矩阵如同沿着对角线进行了翻转。

其数学表达式为 $(A^T)_{ij}=A_{ji}$，即 A^T 的第 i 行 j 列的元素为原矩阵 A 中第 j 行 i 列的元素。

$$\begin{bmatrix}1 & 3\\ 2 & 3\\ 4 & 1\end{bmatrix}^T=\begin{bmatrix}1 & 2 & 4\\ 3 & 3 & 1\end{bmatrix}$$

若 A 是对称矩阵则有 $A^T=A$。

矩阵乘积的转置 $(AB)^T=B^TA^T$

给定一个矩阵 R，R 可以不是方阵，则乘积 R^TR 一定是对称阵。

$$(R^TR)^T=R^T(R^T)^T=R^TR$$

向量空间 Vector spaces

我们可以对向量进行所谓“线性运算”，即通过加和（$\mathbf{v}+\mathbf{w}$）与数乘运算（$3\mathbf{v}$）得到向量的线性组合。向量空间对线性运算封闭，即空间内向量进行线性运算得到的向量仍在空间之内。

$\mathbf{R}^2$ 即为向量空间，它是具有两个实数分量的所有向量（二维实向量）的集合。

例如$\begin{bmatrix}3\\2\end{bmatrix}$，$\begin{bmatrix}0\\0\end{bmatrix}$，$\begin{bmatrix}\pi\\e\end{bmatrix}$……

向量$\begin{bmatrix}a\\b\end{bmatrix}$的图像是从原点出发到点(a,b)的箭头，其中第一分量 a 为横轴坐标，第二分量 b 为纵轴坐标。空间 $\mathbf{R}^2$ 的图像为整个 x-y 平面。

所有向量空间必然包含零向量，因为任何向量数乘 0 或者加上反向量都会得到零向量，而因为向量空间对线性运算封闭，所以零向量必属于向量空间。

$\mathbf{R}^3$ 是向量空间，它是具有三个实数分量的所有向量的集合。

$\mathbf{R}^n$ 是向量空间，它是具有 n 个实数分量的所有向量的集合。

反例：$\mathbf{R}^2$ 中的第一象限则不是一个向量空间。

子空间 Subspaces

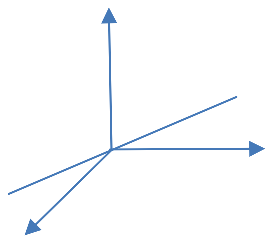
包含于向量空间之内的一个向量空间称为原向量空间的一个子空间。例如用实数 c 数乘 $\mathbf{R}^2$ 空间中向量 **v** 所得到的向量集合就是 $\mathbf{R}^2$ 空间的一个子空间，其图像为二维平面上穿过原点的一条直线，它对于线性运算封闭。

反例：$\mathbf{R}^2$ 中不穿过原点的直线就不是向量空间。子空间必须包含零向量，原因就是数乘 0 的到的零向量必须处于子空间中。

$\mathbf{R}^2$ 的子空间包括：

- $\mathbf{R}^2$ 空间本身
- 过原点的一条直线（这是 $\mathbf{R}^2$ 空间中的一条直线，与 $\mathbf{R}^1$ 空间有区别）
- 原点 仅包含 0 向量 Z

$\mathbf{R}^3$ 的子空间包括：



- $\mathbf{R}^3$ 空间本身 3 维
- 过原点的一个平面 2 维
- 过原点的一条直线 1 维
- 原点 仅包含 0 向量 Z

列空间 Column spaces

给定矩阵 ***A***，其列向量属于 $\mathbf{R}^3$ 空间，这些列向量和它们的线性组合张成了 $\mathbf{R}^3$ 空间中的一个子空间，即矩阵 ***A*** 的列空间 C(***A***)。

如果 $\boldsymbol{A}=\begin{bmatrix}1&3\\2&3\\4&1\end{bmatrix}$，则 ***A*** 的列空间是 $\mathbf{R}^3$ 空间中包含向量$\begin{bmatrix}1\\2\\4\end{bmatrix}$和$\begin{bmatrix}3\\3\\1\end{bmatrix}$并穿过原点的平面，空间内包含两向量的所有线性组合。

下面课程的任务就是在列空间和子空间的基础上理解 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$。

第 06 讲 列空间和零空间

Column space &Nullspace

本节继续研究子空间，特别是矩阵的列空间（column space）和零空间（nullspace）。

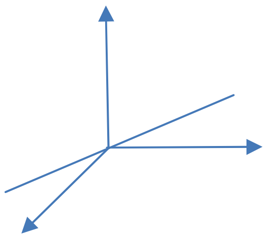
子空间综述

所谓的“向量空间”是对于线性运算封闭的向量集合。即对于空间中的任意向量 **v** 和 **w**，其和 **v**+**w** 和数乘 c**v** 必属于该空间；换而言之对于任何实数 c 和 d，线性组合 c**v**+d**w** 必属于该空间。

$\mathbf{R}^1$，$\mathbf{R}^2$，$\mathbf{R}^3$……都是重要的向量空间，$\mathbf{R}^n$ 代表的空间包含所有具有 n 个分量的向量。其中字母 R 表明分量均为实数（real）。

“子空间”为包含于向量空间内的一个向量空间。它是原向量空间的一个子集，而且本身也满足向量空间的要求。但是“子空间”和“子集”的概念有区别，所有元素都在原空间之内就可称之为子集，但是要满足对线性运算封闭的子集才能成为子空间。

例如



通过原点的平面 P 是 $\mathbf{R}^3$ 的一个子空间。

通过原点的直线 L 也是 $\mathbf{R}^3$ 的子空间。

P 并 L（union）通常并不是 $\mathbf{R}^3$ 的子空间。

P 交 L（intersection）是 $\mathbf{R}^3$ 子空间的特例——0 空间，只有零向量。

任意子空间 S 和 T 的交集都是子空间，可以通过 S 和 T 本身对线性组合封闭来证明。

列空间 Column space

矩阵 $\boldsymbol{A}$ 的列空间 C($\boldsymbol{A}$)是其列向量的所有线性组合所构成的空间。

求解 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的问题，对于给定的矩阵 $\boldsymbol{A}$，对于任意的 **b** 都能得到解么？

$$\text{如果 } \boldsymbol{A}=\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$$

显然并不是所有的 **b** 都能保证 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 有解，因为它有 4 个线性方程而只有 3 个未知数，矩阵 $\boldsymbol{A}$ 列向量的线性组合无法充满 $\mathbf{R}^4$，因此如果 **b** 不能被表示为 $\boldsymbol{A}$ 列向量的线性组合时，方程是无解的。只有当 **b** 在矩阵 $\boldsymbol{A}$ 列空间 C($\boldsymbol{A}$)里时，**x** 才有解。

对于我们所给定的矩阵 $\boldsymbol{A}$，由于列向量不是线性无关的，第三个列向量为前两个列向量之和，所以尽管有 3 个列向量，但是只有 2 个对张成向量空间有贡献。矩阵 $\boldsymbol{A}$ 的列空间为 $\mathbf{R}^4$ 内的一个二维子空间。

零空间（或化零空间）Nullspace

矩阵 $\boldsymbol{A}$ 的零空间 N($\boldsymbol{A}$)是指满足 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的所有解的集合。对于所给定这个矩阵 $\boldsymbol{A}$，其列向量含有 4 个分量，因此列空间是空间 $\mathbf{R}^4$ 的子空间，$\mathbf{x}$ 为含有 3 个分量的向量，故矩阵 $\boldsymbol{A}$ 的零空间是 $\mathbf{R}^3$ 的子空间。对于 $m\times n$ 矩阵，列空间为 $\mathbf{R}^m$ 的子空间，零空间为 $\mathbf{R}^n$ 空间的子空间。

本例中矩阵 $\boldsymbol{A}$ 的零空间 N($\boldsymbol{A}$)为包含$\begin{bmatrix}1\\1\\-1\end{bmatrix}$的任何倍数的集合，因为很容易看到第一列向量（1）和第二列向量（1）相加减去第三列向量（-1）为零。此零空间为 $\mathbf{R}^3$ 中的一条直线。

为了验证 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的解集是一个向量空间，我们可以检验它是否对线性运算封闭。若 $\mathbf{v}$ 和 $\mathbf{w}$ 为解集中的元素，则有：

$$\boldsymbol{A}(\mathbf{v}+\mathbf{w})=\boldsymbol{A}\mathbf{v}+\boldsymbol{A}\mathbf{w}=\mathbf{0}+\mathbf{0}=\mathbf{0},$$

$$\boldsymbol{A}(c\mathbf{v})=c\boldsymbol{A}\mathbf{v}=\mathbf{0},$$

因此得证 N($\boldsymbol{A}$)确实是 $\mathbf{R}^n$ 空间的一个子空间。

b 值的影响 Other values of b

若方程变为

$$\begin{bmatrix}1&1&2\\2&1&3\\3&1&4\\4&1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1\\2\\3\\4\end{bmatrix}$$

则其解集不能构成一个子空间。零向量并不在这个集合内。解集是空间 $\mathbf{R}^3$ 内过$\begin{bmatrix}1\\0\\0\end{bmatrix}$和$\begin{bmatrix}0\\-1\\1\end{bmatrix}$的一个平面，但是并不穿过原点$\begin{bmatrix}0\\0\\0\end{bmatrix}$。

本讲给出了关于矩阵的两种子空间，同时给出了两种构造子空间的方法。对于列空间，它是由列向量进行线性组合张成的空间；而零空间是从方程组出发，通过让 $\mathbf{x}$ 满足特定条件而得到的子空间。

第 07 讲 求解 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$：主变量，特解

Solving $\boldsymbol{A}\mathbf{x}=\mathbf{0}$: pivot variables, special solutions

我们定义了矩阵的列空间和零空间，那么如何求得这些子空间呢？本节课的内容即从定义转到算法。

计算零空间 Nullspace

矩阵 $\boldsymbol{A}$ 的零空间即满足 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的所有 $\mathbf{x}$ 构成的向量空间。

$$\text{取 } \boldsymbol{A}=\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}$$

（$\boldsymbol{A}$ 的列空间并不是线性无关的。）无论矩阵 A 是否可逆，我们都采用消元法作为计算零空间的算法。

对于矩阵 $\boldsymbol{A}$ 进行“行操作”并不会改变 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的解，因此也不会改变零空间。（但是会改变列空间。）此处不需要应用增广矩阵，因为等号右侧的向量 $\mathbf{b}=\mathbf{0}$。

第一步消元得到：

$$\boldsymbol{A}=\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}\rightarrow\begin{bmatrix}\boxed{1}&2&2&2\\0&0&2&4\\0&0&2&4\end{bmatrix}$$

第二列没有主元，因此主元二是第二行第三列的 2。

$$\begin{bmatrix}1&2&2&2\\0&0&2&4\\0&0&2&4\end{bmatrix}\rightarrow\begin{bmatrix}\boxed{1}&2&2&2\\0&0&\boxed{2}&4\\0&0&0&0\end{bmatrix}=\boldsymbol{U}$$

主元列 r=2

自由列 n-r=4-2

矩阵 $\boldsymbol{U}$ 为梯形矩阵。其第三行变为零，是因为第三行的行向量本身就是第一行和第二行行向量的线性组合。

矩阵的秩（rank）就是矩阵的主元的个数。本例中矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{U}$ 的秩均为 2。矩阵中包含主元的列为主元列（pivot column），不包含主元的列称为自由列（free column）。

特解 Special solutions

当我们将系数矩阵变换为上三角阵 $\boldsymbol{U}$ 时，就可以用回代求得方程 $\boldsymbol{U}\mathbf{x}=\mathbf{0}$ 的解。本例中，包含主元的矩阵第 1 列和第 3 列为主元列，而不包含主元的第 2 列和第 4 列为自由列。对自由变量（free variable）x_2 和 x_4 我们可以进行赋值。例如令 $x_2=1$

而 $x_4=0$。则有

$$2x_3+4x_4=0\Rightarrow x_3=0$$

$$x_1+2x_2+2x_3+2x_4=0\Rightarrow x_1=-1$$

因此可得一解 $\mathbf{x}=\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}$，其任意倍数均在矩阵的零空间之内。

取自由变量中 $x_2=0$ 而 $x_4=1$，则可得到另一解 $\mathbf{x}=\begin{bmatrix}2\\0\\-2\\1\end{bmatrix}$

矩阵 ***A*** 的零空间就是这些“特解”向量的线性组合所构成的向量空间。

矩阵的秩 r 等于其主元列的数目，因此自由列的数目就等于 n-r，即列的数目减去主元列的数目。这个数值等于特解的数目和零空间的维数。

主元列和自由列的一个重要区别就是，自由列可以表示为其左侧所有主元列的线性组合，而主元列则不可以。

$$\boldsymbol{U}=\begin{bmatrix}\boxed{\square} & * & * & * & * & * & * & *\\ 0 & 0 & \boxed{\square} & * & * & * & * & *\\ 0 & 0 & 0 & 0 & 0 & \boxed{\square} & * & *\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\square}\end{bmatrix}$$

例如，我们得到一个消元完成后的梯形矩阵 ***U***，其包含四个主元列。观察它的第五列，这是自由列，其左侧有两个主元列，这两个主元列显然线性无关，第五列也显然可以写成前两个主元列的线性组合。这里求的是 $\boldsymbol{U}\mathbf{x}=\mathbf{0}$ 的解，如果令对应第二列，第四列，第六列这三个自由列以及第五列右侧的两个主元列的 x 分量都为 0，而对应第五列的自由变量 $x_5=1$，则方程变为：

$$x_1\begin{bmatrix}\square\\0\\0\\0\end{bmatrix}+0+x_3\begin{bmatrix}*\\\square\\0\\0\end{bmatrix}+0+1\begin{bmatrix}*\\*\\0\\0\end{bmatrix}+0+0+0=0$$

相当于求第五列如何用前两个主元列进行线性组合，所得解 $\mathbf{x}^T=[x_1,0,x_3,0,1,0,0,0]$ 即为原方程的特解之一。对所有的自由列都进行此操作，就是上文求解过程中对自由变量的赋值过程。在本例中，四个自由变量分别取 1 会得到零空间的四个特解。如果把自由变量都赋值为 0 会怎么样？答案是求得的解为 **0** 向量。

行最简阶梯矩阵 Reduced row echelon form （rref）

通过继续消元我们可以将矩阵 ***U*** 转变为行最简阶梯矩阵形式 ***R***，其中主元为 1，

而主元列除主元外皆为 0。在 Matlab 中用命令 rref(***A***)实现这一过程。

$$\boldsymbol{U}=\begin{bmatrix}1&2&2&2\\0&0&2&4\\0&0&0&0\end{bmatrix}\rightarrow\begin{bmatrix}1&2&0&-2\\0&0&2&4\\0&0&0&0\end{bmatrix}\rightarrow\begin{bmatrix}\boxed{1}&2&0&-2\\0&0&\boxed{1}&2\\0&0&0&0\end{bmatrix}=\boldsymbol{R}$$

在矩阵中主元行和主元列的交汇处存在一个单位阵。通过“列交换”，可以将矩阵 ***R*** 中的主元列集中在左侧，从而在左上角形成这个单位阵，而将自由列集中在矩阵的右侧。如果矩阵 ***A*** 中的某些行是线性相关的，则在矩阵 ***R*** 的下半部分就会出现一些完全为 **0** 的行向量。

$$\boldsymbol{R}=\begin{array}{c}\begin{bmatrix}I&F\\0&0\end{bmatrix}\\ \text{r 主元列}\quad\text{n-r 自由列}\end{array}\begin{array}{l}\text{r 主元行}\\ \text{m-r }\mathbf{0}\text{ 向量行}\\ \\ \end{array}$$

这里的 ***I*** 是一个 r×r 的方阵。***F*** 即自由列消元后组成的部分。

原方程 ***A*x**=**0** 变为求解 ***R*** 的主元行乘以 **x**，$\begin{bmatrix}I & F\end{bmatrix}\begin{bmatrix}x_{pivot}\\x_{free}\end{bmatrix}=0$。我们将 ***A*x**=**0** 的特解作为列向量写成一个矩阵 ***N***，即零空间矩阵。则其形式为 ***N***=$\begin{bmatrix}\\I\end{bmatrix}$。这里的 ***I*** 为一个 n-r×n-r 的矩阵，就是对 n-r 个自由变量分别赋值为 1 所构造出来的，零空间矩阵满足 ***RN***=***0***，***0*** 矩阵是一个 m×(n-r)的矩阵。则从矩阵分块乘法运算可知零空间矩阵上半部分为-***F***，即 ***N*** 最终形式为

$$\boldsymbol{N}=\begin{bmatrix}-F\\I\end{bmatrix}\begin{array}{l}\text{r 主变量}\\ \text{n-r 自由变量}\end{array}$$

对于矩阵 ***R*** 而言，求零空间特解就变得非常简单，只需要将消元的到的 ***F*** 部分拼接上单位阵就可以得到所有的通解。注意如果在变换出 ***R*** 左上角的单位阵的过程中采用了列交换，则最后的解要完成逆变换。

第 08 讲 求解 ***A*x**=**b**：可解性与结构

Complete Solution of ***A*x**=**b**

可解的条件 Solvability conditions on b

我们仍取 $\boldsymbol{A}=\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}$为例，则方程为$\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$。

矩阵 ***A*** 的第三行为第一行和第二行的加和，因此 ***A*x**=**b** 中 **b** 的第 3 个分量也要等于其第 1 和第 2 个分量的和。若 **b** 不满足 $b_3=b_1+b_2$ 则方程组无解。

检验 ***A*x**=**b** 是否可解的方法是对增广矩阵进行行消元。如果矩阵 ***A*** 的行被完全消去的话，则对应的 **b** 的分量也要得 0。在本例中，矩阵 ***A*** 的第三行被消去：

$$\begin{bmatrix}1&2&2&2&b_1\\2&4&6&8&b_2\\3&6&8&10&b_3\end{bmatrix}\to\ldots\to\begin{bmatrix}1&2&2&2&b_1\\0&0&2&4&b_2-2b_1\\0&0&0&0&b_3-b_2-b_1\end{bmatrix}$$

如果 ***A*x**=**b** 有解，则 $b_3-b_1-b_2=0$。在本例中我们令 $\mathbf{b}=\begin{bmatrix}1\\5\\6\end{bmatrix}$

前几讲讨论过，只有当 **b** 处于矩阵的列空间 C(***A***)之中时，方程才有解。本讲推导出矩阵 ***A*** 的行向量若经过线性组合成为了零向量，则对应的 **b** 经同样的线性组合后也要等于 0。因此看起来我们有了两条关于 **b** 的限制条件，但实际上这两点是等价的。

通解 Complete solution

为求得 ***A*x**=**b** 的所有解，我们首先检验方程是否可解，然后找到一个特解。将特解和矩阵零空间的向量相加即为方程的通解。

特解 A particular solution

求 ***A*x**=**b** 特解的方法是将自由变量均赋值为 0，求解其主变量。

本例中，令 $x_2=x_4=0$ 得到方程组：

$$\begin{aligned}x_1+2x_3&=1\\2x_3&=3\end{aligned}$$

可解得 $x_3=3/2$，$x_1=-2$，

因此特解为 $\mathbf{x}_p=\begin{bmatrix}-2\\0\\3/2\\0\end{bmatrix}$。

上一讲说了主元列和自由列的一个重要区别就是，自由列可以表示为其左侧所有主元列的线性组合，而主元列则不可以，主元列是线性无关的。

$$\boldsymbol{U}=\begin{bmatrix}\boxed{\square} & * & * & * & * & * & * & * \\ 0 & 0 & \boxed{\square} & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \boxed{\square} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\square}\end{bmatrix}$$

我们仍以这个消元完成后的梯形矩阵 $\boldsymbol{U}$ 为例，其包含四个主元列。对于 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的求解转变为 $\boldsymbol{U}\mathbf{x}=\mathbf{c}$，其中 $\mathbf{c}$ 是向量 $\mathbf{b}$ 经过与左侧矩阵相同的行操作得到的向量。明显可以看到，此时四个主元列的线性组合可以组成任何 $\mathbf{R}^4$ 中的向量，我们将 $\mathbf{x}$ 中的自由变量赋值为 0 就可以去掉自由列列向量的干扰，求得方程的特解 $\mathbf{x}_p$。如果消元得到 $\boldsymbol{U}$ 最后 i 行为 0，就要求 $\mathbf{c}$ 的最后 i 个分量为 0，这时主元列才可以通过线性组合得到 $\mathbf{c}$，否则无解。

与零空间进行线性组合 Combined with nullspace

$\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的通解为 $\mathbf{x}_{complete}=\mathbf{x}_p+\mathbf{x}_n$，其中 $\mathbf{x}_n$ 为矩阵零空间中的一般向量。将 $\boldsymbol{A}\mathbf{x}_p=\mathbf{b}$ 和 $\boldsymbol{A}\mathbf{x}_n=\mathbf{0}$ 相加可得 $\boldsymbol{A}(\mathbf{x}_p+\mathbf{x}_n)=\mathbf{b}$。

上一讲我们得到了矩阵的零空间 N($\boldsymbol{A}$)就是其特解 $\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}$ 和 $\begin{bmatrix}2\\0\\-2\\1\end{bmatrix}$ 的线性组合的集合。因此方程 $\boldsymbol{A}\mathbf{x}=\begin{bmatrix}1\\5\\6\end{bmatrix}$ 的通解即为：

$$\mathbf{x}_{complete}=\begin{bmatrix}-2\\0\\3/2\\0\end{bmatrix}+c_1\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}2\\0\\-2\\1\end{bmatrix}$$

式中 c_1 和 c_2 为任意实数。

矩阵的零空间 N($\boldsymbol{A}$)是 $\mathbf{R}^4$ 空间中的二维子空间，方程的解 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 构成了穿过 $\mathbf{x}_p$ 点并和矩阵零空间平行的“平面“。但该”平面“并不是 $\mathbf{R}^4$ 空间的子空间。

前面求取 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的特解过程中，我们令所有自由变量赋值为 0。如果不赋值为 0，则等于带着自由列进行计算，但自由列其实也就是主元列的线性组合，这样求的

特解 $\mathbf{x}_p'$只不过是 $\mathbf{x}_p$ 与零空间特解的一个加和。如此例中若令 $x_2=1,x_4=0$，则求特解的方程变为：$\begin{matrix} x_1+2+2x_3=1 \\ 2x_3=3 \end{matrix}$，求得 $x_1=-4$，$x_3=3/2$。因此特解为 $\mathbf{x}_p'=\begin{bmatrix} -4 \\ 1 \\ 3/2 \\ 0 \end{bmatrix}$。而实际上 $\mathbf{x}_p'=x_p+x_{n1}=\begin{bmatrix} -2 \\ 0 \\ 3/2 \\ 0 \end{bmatrix}+\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$。

秩 Rank

矩阵的秩等于矩阵的主元数。如果 $m\times n$ 矩阵的秩为 r，则必有 $r\le m$ 且 $r\le n$。

讨论满秩（full rank）的情形：

- 列满秩：r=n。每列都有主元，**x** 的每一个分量都是主变量，没有自由变量。零空间 N(***A***)之内只有零向量。方程无解或者有唯一解 $\mathbf{x}_p$。

$$\text{例如 } \boldsymbol{A}=\begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 6 & 1 \\ 5 & 1 \end{bmatrix}\rightarrow\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}=\boldsymbol{R}$$

- 行满秩：r=m。每行都有主元，无论 **b** 取何值，方程 ***A*x**=**b** 都有解。主变量 r 个，自由变量 n-r 个。

$$\text{例如 } \boldsymbol{A}=\begin{bmatrix} 1 & 2 & 6 & 5 \\ 3 & 1 & 1 & 1 \end{bmatrix}\rightarrow\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{bmatrix}=\boldsymbol{R}$$

- 满秩 r=m=n，矩阵可逆。零空间只有零向量，无论 **b** 取何值，方程 ***A*x**=**b** 都有唯一解。

$$\text{例如 } \boldsymbol{A}=\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}\rightarrow\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=\boldsymbol{R}\text{ 为单位阵。}$$

总结：

r=m=n	r=n<m	r=m<n	r<n,r<m
R= I	***R***= $\begin{bmatrix} I \\ 0 \end{bmatrix}$	***R***= $\begin{bmatrix} I & F \end{bmatrix}$	***R***= $\begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$
唯一解	无解或唯一解	无穷多解	无解或无穷多解

秩决定了方程组解的数量。

$m\times n$ 给出了矩阵的尺寸，但是秩 r 给出的是矩阵的实际“大小”。关于这个随后会有讨论。

第 09 讲 线性无关，基和维数

Independence, basis, and dimension

向量的线性无关意味着什么？如何用线性无关的概念来帮助我们描述包括零空间在内的子空间。

线性无关 Independence

矩阵 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，其中 $m<n$（因此 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 中未知数个数多于方程数）。则 $\boldsymbol{A}$ 中具有至少一个自由变量，那么 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 一定具有非零解。$\boldsymbol{A}$ 的列向量可以线性组合得到零向量，所以 $\boldsymbol{A}$ 的列向量是线性相关的。

若 $c_1\mathbf{x}_1+c_2\mathbf{x}_2+\ldots\ldots+c_n\mathbf{x}_n=0$ 仅在 $c_1=c_2=\ldots\ldots=c_n=0$ 时才成立，则称向量 $\mathbf{x}_1$，$\mathbf{x}_2\ldots\ldots\mathbf{x}_n$ 是线性无关的。若这些向量作为列向量构成矩阵 $\boldsymbol{A}$，则方程 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 只有零解 $\mathbf{x}=\mathbf{0}$，或称矩阵 $\boldsymbol{A}$ 的零空间只有零向量。换而言之，若存在非零向量 $\mathbf{c}$，使得 $\boldsymbol{A}\mathbf{c}=\mathbf{0}$，则这个矩阵 $\boldsymbol{A}$ 的列向量线性相关。

在 $\mathbf{R}^2$ 中，两个向量只要不在一条直线上就是线性无关的。（在 $\mathbf{R}^3$ 中，三个向量线性无关的条件是它们不在一个平面上。）若选定空间 $\mathbf{R}^2$ 中的三个向量，则他们必然是线性相关的。例如，如下的三个向量 $\mathbf{v}_1$，$\mathbf{v}_2$ 和 $\mathbf{v}_3$ 是线性相关的。

$$\begin{array}{cc} & \begin{array}{ccc}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{array} \\ \boldsymbol{A}= & \begin{bmatrix} 2 & 1 & 2.5 \\ 1 & 2 & -1 \end{bmatrix}\end{array}$$

由我们研究方程组得到的结论，此矩阵构成的方程 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 必有非零解，即三个向量线性相关。

如果矩阵 $\boldsymbol{A}$ 的列向量为线性无关，则 $\boldsymbol{A}$ 所有的列均为主元列，没有自由列，矩阵的秩为 n。若 $\boldsymbol{A}$ 的列向量为线性相关，则矩阵的秩小于 n，并且存在自由列。

张成空间 Spanning a space

当一个空间是由向量 $\mathbf{v}_1$，$\mathbf{v}_2\ldots\ldots\mathbf{v}_k$ 的所有线性组合组成时，我们称这些向量张成了这个空间。例如矩阵的列向量张成了该矩阵的列空间。

如果向量 $\mathbf{v}_1$，$\mathbf{v}_2\ldots\ldots\mathbf{v}_k$ 张成空间 $\mathbf{S}$，则 $\mathbf{S}$ 是包含这些向量的最小空间。

基与维数 Basis &Dimension

向量空间的基是具有如下两个性质的一组向量 $\mathbf{v}_1$，$\mathbf{v}_2\ldots\ldots\mathbf{v}_d$：

- $\mathbf{v}_1$，$\mathbf{v}_2\ldots\ldots\mathbf{v}_d$ 线性无关
- $\mathbf{v}_1$，$\mathbf{v}_2\ldots\ldots\mathbf{v}_d$ 张成该向量空间

空间的基告诉我们了空间的一切信息。

例：$\mathbf{R}^3$ 空间

$\mathbf{R}^3$ 空间有一组基 $\begin{bmatrix}1\\0\\0\end{bmatrix}$，$\begin{bmatrix}0\\1\\0\end{bmatrix}$，$\begin{bmatrix}0\\0\\1\end{bmatrix}$，因为 $c_1\begin{bmatrix}1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}0\\1\\0\end{bmatrix}+c_3\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$ 只有零解。并且这三个向量可以张成 $\mathbf{R}^3$ 空间。

而 $\begin{bmatrix}1\\1\\2\end{bmatrix}$，$\begin{bmatrix}2\\2\\5\end{bmatrix}$，$\begin{bmatrix}3\\3\\8\end{bmatrix}$ 则不能构成一组基，因为以它们为列向量组成的矩阵，有两个相同的行，消元肯定有自由列存在，因此这三个向量并非线性无关。GS 在这里犯了个错误，他以为这三个是线性无关的，最后在第 10 讲中才进行了更正。当判定线性相关性时，可以随时在矩阵、空间和方程组的概念之间切换，哪个判据更容易判定就用哪个，这里显然矩阵不可逆更容易看出来，因为存在行向量重复的情况。从这里也可以看到行向量线性相关则列向量不可能线性无关，矩阵行空间的维数等于列空间的维数。

若以 $\mathbf{R}^n$ 空间中的 n 个向量为列向量构成的矩阵为可逆矩阵，则这些向量可以构成 $\mathbf{R}^n$ 空间中的一组基。

子空间的基 Basis for a subspace

向量 $\begin{bmatrix}1\\1\\2\end{bmatrix}$，$\begin{bmatrix}2\\2\\5\end{bmatrix}$ 可以张成 $\mathbf{R}^3$ 中的一个平面，但是它们无法成为 $\mathbf{R}^3$ 空间的一组基。

空间的每一组基都具有相同的向量数，这个数值就是空间的维数(dimension)。所以 $\mathbf{R}^n$ 空间的每组基都包含 n 个向量。

列空间和零空间的基 Basis of a column space and nullspace

$$取\ \boldsymbol{A}=\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}$$

讨论列空间： 矩阵 $\boldsymbol{A}$ 的四个列向量张成了矩阵 $\boldsymbol{A}$ 的列空间，其中第 3 列和第 4 列与前两列线性相关，而前两个列向量线性无关。因此前两列为主元列。他们组成了列空间 C($\boldsymbol{A}$)的一组基。矩阵的秩为 2。

实际上对于任何矩阵 $\boldsymbol{A}$ 均有：

矩阵的秩 r=矩阵主元列的数目=列空间的维数

注意：矩阵具有秩 rank 而不是维数 dimension，而空间有维数而不是秩。

当知道了列空间的维数，可以从矩阵列向量中随意选取足够数量的线性无关的向量，它们每一组都可以构成列空间的一组基。

讨论零空间： 本例中矩阵的列向量不是线性无关的，因此其零空间 N($\boldsymbol{A}$)不止包

含零向量。因为可以看出第 3 列是第 1 列和第 2 列的加和。所以向量$\begin{bmatrix}-1\\-1\\1\\0\end{bmatrix}$必然在零空间 N($\boldsymbol{A}$)之内。同样还可以对 x_4 赋值为 1，从而得到$\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$也在零空间之内。它们就是 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的两个特解。

零空间的维数=自由列的数目=n-r，因此本例中 N($\boldsymbol{A}$)的维数为 4-2=2。这两个特解就构成了零空间的一组基。

第 10 讲 四个基本子空间

Thefour fundamental subspaces

本讲讨论矩阵的四个基本子空间以及他们之间的关系。

四个子空间 Four subspaces

任意的 m×n 矩阵 $\boldsymbol{A}$ 都定义了四个子空间。

列空间 Column space C($\boldsymbol{A}$)

矩阵 $\boldsymbol{A}$ 的列空间是 $\boldsymbol{A}$ 的列向量的线性组合在 $\mathbf{R}^m$ 空间中构成的子空间。

零空间 Nullspace N($\boldsymbol{A}$)

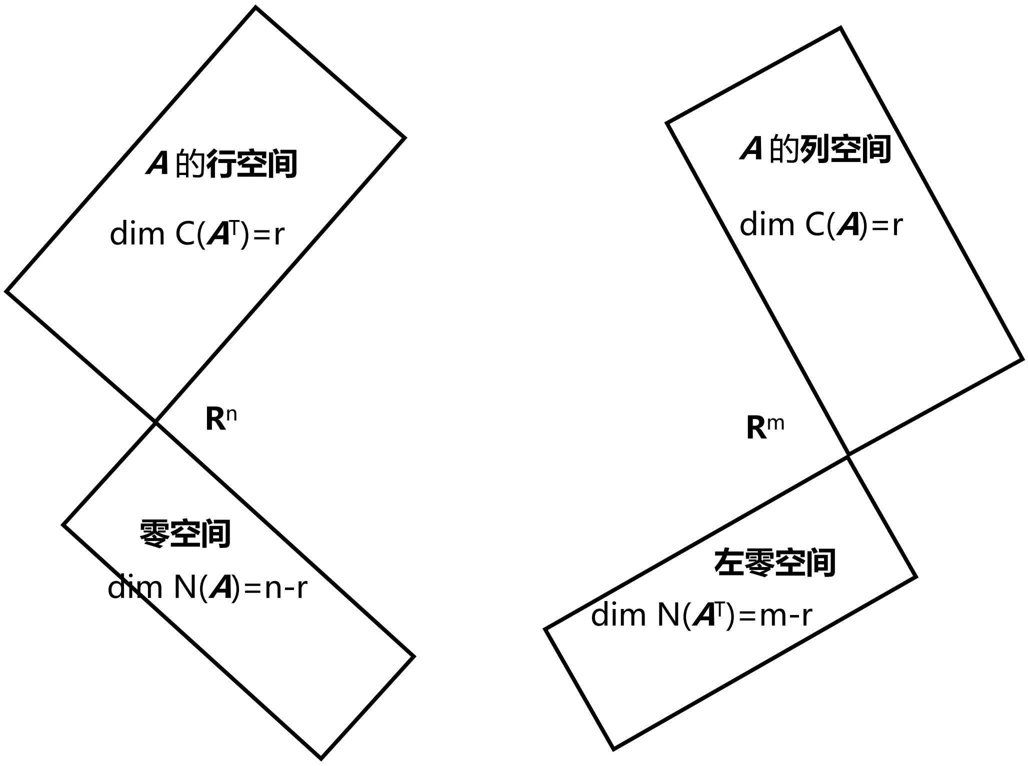
矩阵 $\boldsymbol{A}$ 的零空间是 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的所有解 $\mathbf{x}$ 在 $\mathbf{R}^n$ 空间中构成的子空间。

行空间 Row space C($\boldsymbol{A}^T$)

矩阵 $\boldsymbol{A}$ 的行空间是 $\boldsymbol{A}$ 的行向量的线性组合在 $\mathbf{R}^n$ 空间中构成的子空间，也就是矩阵 $\boldsymbol{A}^T$ 的列空间。

左零空间 Left nullspace N($\boldsymbol{A}^T$)

我们称矩阵 $\boldsymbol{A}^T$ 的零空间为矩阵 $\boldsymbol{A}$ 的左零空间，它是 $\mathbf{R}^m$ 空间中的子空间。



基和维数 Basis& Dimension

列空间

矩阵 ***A*** 的 r 个主元列构成了列空间 C(***A***)的一组基。dim C(***A***)=r

零空间

A*x=0** 的一组特解对应于矩阵 ***A 的 n-r 个自由列，并构成了零空间的一组基。dim N(***A***)=n-r

行空间

我们用矩阵 ***A*** 的化简的行阶梯矩阵 ***R***。

$$\boldsymbol{A}=\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}\to\ldots\to\begin{bmatrix}1&0&1&1\\0&1&1&0\\0&0&0&0\end{bmatrix}=\begin{bmatrix}I&F\\0&0\end{bmatrix}=\boldsymbol{R}$$

尽管矩阵 ***A*** 和矩阵 ***R*** 的列空间不同，但两者行空间相同。***R*** 的行向量来自于 ***A*** 的行向量的线性组合，因为消元操作是可逆的，所以 ***A*** 的向量也可以表示为 ***R*** 行向量的线性组合。

R 的前 r 行阶梯型“行向量”就是矩阵 ***A*** 行空间 C($\boldsymbol{A}^T$)的一组基。dim C($\boldsymbol{A}^T$)=r

左零空间

矩阵 $\boldsymbol{A}^T$ 有 m 列，而其秩为 r，因此其自由列数目为 m-r。所以 dim N($\boldsymbol{A}^T$)=m-r。

左零矩阵是满足 $\boldsymbol{A}^T\mathbf{y}=\mathbf{0}$ 的所有向量 **y** 的集合。称之为左零矩阵是因为该式可写作 $\mathbf{y}^T\boldsymbol{A}=\mathbf{0}$，而 **y** 出现在矩阵 ***A*** 左侧。

为找到左零空间的基，我们应用增广矩阵：

$$\begin{bmatrix}A_{m\times n} & I_{m\times n}\end{bmatrix}\to\begin{bmatrix}R_{m\times n} & E_{m\times n}\end{bmatrix}$$

我们将 ***A*** 通过消元得到矩阵 ***R***，其消元矩阵记为 ***E***，即 ***EA***=***R***。若 ***A*** 为方阵，且 ***R***=***I***，则有 ***E***=$\boldsymbol{A}^{-1}$。在本例中

$$\boldsymbol{EA}=\begin{bmatrix}-1&2&0\\1&-1&0\\-1&0&1\end{bmatrix}\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}=\begin{bmatrix}1&0&1&1\\0&1&1&0\\0&0&0&0\end{bmatrix}=\boldsymbol{R}$$

以“行操作”的观点来看矩阵 ***E*** 和 ***A*** 的乘法，则矩阵 ***E*** 最下面的 m-r 个行向量使得矩阵 ***A*** 的行向量线性组合成为 **0**，也就是矩阵 ***R*** 最下面的 m-r 个零向量。本例中，m-r=1。

矩阵 ***E*** 的这 m-r 个行向量满足 $\mathbf{y}^T\boldsymbol{A}=\mathbf{0}$，它组成了矩阵 A 左零空间的一组基。

新向量空间 New vector space

所有 3×3 矩阵构成的集合是一个向量空间，符合对于线性运算封闭，称之为 **M**。

M 的子空间包括：

- 所有的上三角阵

- 所有的对称阵
- 所有的对角阵

对角阵是前两个子空间的交集，其维数为 3，具有以下一组基：

$$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

当最初告诉我说，矩阵的列秩等于主元数，并且主元列构成了列空间的一组基时，其实我是拒绝的。主元这个东西不是行变换消元得来的么，消元过程列空间不是已经改变了么，为什么所得出 $\boldsymbol{U}$ 的主元数和主元列的位置还能够反映出矩阵 $\boldsymbol{A}$ 列空间的状态呢？

这里需要说明的是两点，其一是关于秩的定义有很多在数学上等价但是描述差异很大的说法，在这里我们把"秩"理解为行（列）向量中最大的线性无关向量组的向量数即可，在矩阵 $\boldsymbol{A}$ 行变换消元成梯形阵后，很容易看到行空间内极大无关组之一就是主元所在的那前 r 行，这 r 个行向量可以张成行空间，因此行空间的维数与主元数相等都是 r，并且前 r 行构成了行空间的一组基。

但是为什么列空间的维数也是 r，并且主元列可以构成列空间的一组基呢？这就是要说明的第二点，初等行变换不会改变列向量的线性相关性。为了叙述方便起见，我们假定矩阵 $\boldsymbol{A}$ 列向量的极大无关组就是 $\boldsymbol{A}$ 前 r' 列的向量（若否可以通过列交换而达成，列交换不会改变线性关系）。则矩阵 $\boldsymbol{A}$ 经过行消元得到梯形矩阵 $\boldsymbol{U}$ 的过程，可以用左乘可逆矩阵 $\boldsymbol{E}$ 来实现，即 $\boldsymbol{EA}=\boldsymbol{U}$。对等式左侧的矩阵乘法采用列操作的形式进行理解，$\boldsymbol{E}[\mathbf{a}_1,\mathbf{a}_2,\ldots\ldots\mathbf{a}_n]=\boldsymbol{U}$，其中 $\mathbf{a}_i$ 为矩阵 $\boldsymbol{A}$ 的第 i 列。则 $\boldsymbol{U}$ 的前 r' 列分别为 $\boldsymbol{E}\mathbf{a}_1\ldots\ldots\boldsymbol{E}\mathbf{a}_{r'}$。第一步，证明之前线性无关的列向量，行变换之后还无关：反设这些列向量线形相关，即存在非零 $c_1, c_2\ldots\ldots c_{r'}$ 使方程 $c_1\boldsymbol{E}\mathbf{a}_1+c_2\boldsymbol{E}\mathbf{a}_2\ldots\ldots+c_{r'}\boldsymbol{E}\mathbf{a}_{r'}=\mathbf{0}$，等式等价于 $\boldsymbol{E}(c_1\mathbf{a}_1+\ldots\ldots+c_{r'}\mathbf{a}_{r'})=\mathbf{0}$，等式两端同时乘以 $\boldsymbol{E}^{-1}$，则有 $c_1\mathbf{a}_1+c_2\mathbf{a}_2\ldots\ldots+c_{r'}\mathbf{a}_{r'}=\mathbf{0}$，但已知前 r' 列的向量是线性无关的，所以 $\mathbf{c}=0$，矛盾，因此经过左乘可逆矩阵 $\boldsymbol{E}$，即"行变换"，不会改变线性无关的状态。第二步证明类似，就是证明之前线性相关的还相关，没有生成新的线性无关项，此处略去。综上可知，经过行变换列向量极大无关组的向量数仍为 r'，即 $\boldsymbol{U}$ 的列空间维度为 r'。而从 $\boldsymbol{U}$ 的矩阵形态可知，它的 r 个主元列线性无关，其列空间的维度为 r，因此有 $r'=r$，并且主元列经过逆变换回去得到 $\boldsymbol{A}$ 的主元列也不会改变线性无关的关系，因此 $\boldsymbol{A}$ 的主元列就是 $\boldsymbol{A}$ 列空间的基。

我们大致说明了一件事情，就是矩阵的行秩等于其列秩。如果将矩阵 $\boldsymbol{A}$ 消元得到 rref，$\boldsymbol{PA}=\boldsymbol{R}$，其中 $\boldsymbol{R}=\begin{bmatrix}I&F\\0&0\end{bmatrix}$，对 $\boldsymbol{R}$ 再进行列变换，可得 $\boldsymbol{PAQ}=\boldsymbol{RQ}=\begin{bmatrix}I&0\\0&0\end{bmatrix}$，最后凹出的这个造型被称为矩阵 $\boldsymbol{A}$ 的相抵标准型，它和 $\boldsymbol{A}$ 具有相同的秩，具有同构

的行空间和列空间。你可以视这个标准型为“初代”，在对它进行了可逆的行变换或者列变换之后得到了一系列矩阵，其中就包括老 ***A***，这个初代的行列空间维数同为 r，所以它“繁衍”出来的一系列矩阵的行列空间维数就同为 r，所有的矩阵都有它的相抵标准型，所有的矩阵行秩都等于列秩。所以之前我们说过 m 和 n 给出了矩阵的尺寸，但是 r 给出了矩阵的“大小”，它告诉了你矩阵行向量和列向量中真正“管用的”有几个，它告诉了你方程组里管用的方程有几个，解集里面不能随便取的变量是几个，能随便赋值的又是哪几个。

再扯几句淡，如果去网上搜索“矩阵的行秩等于列秩究竟意味着什么”，那么你会发现很多看上去完完全全不同的说法，甚至完全不像是在讨论一个问题。之所以有这种感受，其中一个原因是对于秩的定义出发点不同，比如从行列式开始讲起的线代是用行列式来定义“秩”的；而另一个原因是，大家是在线代的不同层面在讨论这件事情。总体而言，GS 所讲的这套线代，是偏重于应用数学这个方向，希望通过深入理解矩阵运算，在工程应用的方面能走得更远，后面他引入马尔科夫链、傅里叶、微分方程的内容都是朝着这个方向努力，还有他开设的“计算科学与工程”的课程也是秉持这个理念；与之不同，还有一种讲解线性代数的思路是强调线性算子的核心作用，实际上它是在一个更加抽象的层面来进行讨论和推演，Sheldon Axler 的《Linear algebra done right》（中译本：线性代数应该这样学）就是以这个出发点来写作的书，引领我们走向线代更深入的部分，方便与其他数学分支建立联系。

其实数学的学习过程中，我们一直在做着“抽象”这件工作，我们从 10 个苹果、2 头牛这种实际事物中，抽象出数字的概念放下了实物的概念，随后我们抽象出未知数 x，y 而放下了具体数字的概念，再后来到函数连 x，y 的系数都已经被字母替代，到了微积分直接讨论函数之间的联系已经可以暂时放下函数的解析式，再到泛函……每前进一步我们都用一个更加抽象的概念替代一族具象的概念，使得一切操作变得更加简洁（但不一定就好懂），而结果更富普遍意义，但是对于理解能力就有更高的要求。

把话说回来，GS 讲的线性代数更 low 么（有人说 GS 的所讲的不是线性代数的核心而是 matlab 的核心-_-!）？显然不是的，他将线代的触角已经延伸出去很广的范围，在应用的领域做了非常细致有效的工作，开启了科学家和工程师对线代的正确理解方式。打个比方，就好像外门与内功，GS 的线代有如老顽童评价的“降龙十八掌”——已经是外门功夫的顶尖了，郭靖一招“亢龙有悔”基本上 S 级以下的出场人物分分钟打跑。反观张无忌身负九阳神通，被灭绝在昆仑山景区当众虐菜。都开着主角光环怎么差距就这么大呢！以我工科学渣的身份，估计内功真是了解一下就好了，把外功练到极致可能是更好的选择。

第 11 讲矩阵空间、秩 1 矩阵和小世界图

Matrix Spaces；Rank 1；Small World Graphs

扩展一下向量空间的含义。

新的向量空间 New vector spaces

3×3 矩阵空间 3 by 3 matrices

空间 **M** 是所有 3×3 矩阵所构成的空间，**M** 的部分子空间包括：

- 所有的上三角阵，记为 **U**。
- 所有的对称阵，记为 **S**。
- 所有的对角阵，记为 **D，**它是前两个子空间的交集。

空间 **M** 的维数为 9，与 $\mathbf{R}^9$ 空间很类似。我们可以选定它的一组基：

$$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix},\ldots\ldots\begin{bmatrix}0&0&0\\0&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

对角阵构成的子空间 **S** 维数为 6，它的一组基为：

$$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&1\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

上三角阵构成的子空间 **U** 维数也为 6，它的一组基为：

$$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&1\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

对角阵构成的子空间 **D** 维数为 3，可以选定 **S** 和 **U** 的基的交集为 **D** 的基。

S 和 **U** 的并集，即 3×3 矩阵中或为上三角阵或为对称阵的矩阵，构成 **M** 的子空间么？答案是否定的，这就如同在 $\mathbf{R}^2$ 空间中找出两条直线，询问它们的并集是否构成一个子空间。如果我们将 **S** 和 **U** 中所有元素可能构成的加和作为一个集合，可以称为和集 **S**+**U**，它是 **M** 的一个子空间。实际上 **S**+**U** 就是 **M** 本身，其维数为 9。

$$\dim \mathbf{S}+\dim \mathbf{U}=\dim \mathbf{S}\cap\mathbf{U}+\dim \mathbf{S}+\mathbf{U}$$

此处网上的笔记有误，等式右侧应是和集的维数与交集的维数相加。

微分方程 Differential equations

对于给定的微分方程 $\dfrac{d^2y}{dx^2}+y=0$，求解该方程可以视为求它的零空间。我们可以得到的解为：$y=\cos x$，$y=\sin x$，$y=e^{ix}$。

事实上，通解为 $y=c_1\cos x+c_2\sin x$，其中 c 可以取任意复数。也将解的线性组合构成的空间称为解空间，其维数为 2。cosx 和 sinx 可以成为解空间的一组基。这

些并不像是向量，它们是函数，但是可以对其进行线性运算，在线性代数的范畴内讨论之。

秩 1 矩阵 Rank one matrices

矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & 4 & 5\\ 2 & 8 & 10\end{bmatrix}=\begin{bmatrix}1\\ 2\end{bmatrix}\begin{bmatrix}1 & 4 & 5\end{bmatrix}$，矩阵的秩为 1。dim C($\boldsymbol{A}$)=r=dim C($\boldsymbol{A}^T$)

所有的秩 1 矩阵都可分解 $\boldsymbol{A}=\mathbf{U}\mathbf{V}^T$，其中 **U** 和 **V** 都是列向量。秩 1 矩阵的行列式和特征值都很简单，它可以被当作是构建其他矩阵的“积木”。其实我们在矩阵乘法的第四种形式里面见过它的作用。如果存在一个 5×17 的矩阵 $\boldsymbol{M}$，而其秩为 4，那么它可以由 4 个秩 1 矩阵组合而成。

若矩阵空间 **M** 为所有的 5×17 矩阵，那么 **M** 中所有的秩 4 矩阵所构成的集合是一个子空间么？答案是否定的，即使加入零矩阵也无法构成子空间，对于两个矩阵的加和，秩 4 矩阵集合并不封闭。

两个矩阵之和的秩小于等于两个矩阵的秩之和。

在 $\mathbf{R}^4$ 空间中，所有满足分量之和 $v_1+v_2+v_3+v_4=0$ 的向量 $\mathbf{v}=\begin{bmatrix}v_1\\ v_2\\ v_3\\ v_4\end{bmatrix}$ 构成了一个子空间 **S**，包含零向量并对线性运算封闭。它是矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}$ 的零空间，因为矩阵 $\boldsymbol{A}$ 的秩为 1，因此其零空间的秩为 n-r=3。零空间的基就是 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 特解：

$$\begin{bmatrix}-1\\ 1\\ 0\\ 0\end{bmatrix},\begin{bmatrix}-1\\ 0\\ 1\\ 0\end{bmatrix},\begin{bmatrix}-1\\ 0\\ 0\\ 1\end{bmatrix}。$$

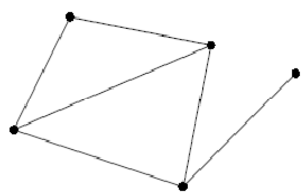
本例的一个意义就是寻找子空间维数的逆向思路，可以考虑它是不是某个方程的解空间，在这里它是矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}$ 的零空间，我们可以从矩阵推出这个空间的维数。

矩阵 $\boldsymbol{A}$ 的列空间为 $\mathbf{R}^1$。左零空间仅包含零向量，维数为 0。

小世界图 Small world graphs

介绍小世界图主要是引出图论和线性代数的联系。

在这里，“图” G 是结点和边的集合：G=｛结点（nodes），边（edges）｝



此图包含 5 个结点和 6 条边，我们可以利用一个 5×6 矩阵完全描述它。

我们可以用图来描述一个实际问题，如果每个人是一个结点，两个人互相认识为一个边，那么整个美国可以以此构成一张大图。我们可以通过这张图来确认两个人之间的最短距离是多少，即两个人需要通过最少几个朋友才能建立联系。G 本人和克林顿之间的距离为 2，他的一个朋友是参议员，他认识这个参议员朋友，那个人认识克林顿。班里的学生跟克林顿的距离因此不会大于 3。还可以继续算希拉里和莱温斯基……

所谓“六度分割理论”（six degrees of separation）猜想一个人和陌生人之间间隔的点不会超过六个。因此当陌生的两人聊起这种联系都会感叹：“世界真小啊！”这也是“小世界图”这个名字的由来。

尽管都在 $A\mathbf{x}=\mathbf{b}$ 这个单元之下，但是第 11 和第 12 两讲的内容与其他几节课程明显有所差异。之前的课程偏重于介绍线性代数的基础知识，主要是将矩阵运算和消元法解线性方程联系起来，告诉我们矩阵不是你常规理解的用于存储的数组，它可以是一种操作或者是空间的缩影。而这两节课程告诉我们线性方程组也不是你常规理解的，从一个非常明显的线性关系抽象出来的方程，它可以是从微分方程或者从图论的角度建立起来的关系式，而应用我们在线性代数里面学到的理论，可以处理这些问题，甚至得到一些更加精深的理解。

在应用数学工具来解决实际问题的过程中有两个基本步骤：建模和解方程。可以说前面的课程主要在解方程这边，而最后两节课稍稍向建模那个方面靠近了一点。

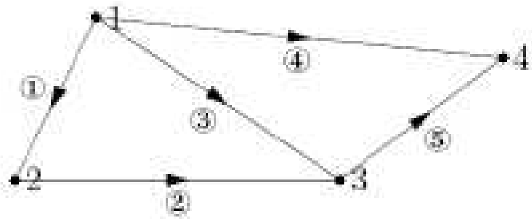
第 12 讲 图、网络、关联矩阵

Graphs，networks，incidence matrices

本讲讨论线性代数在物理系统中的应用。

图和网络 Graphs & Networks

“图”就是“结点”和“边”的一个集合。



边线上的箭头代表从结点流出的正方向。

关联矩阵（Incidence matrices）

构造一个矩阵来表示图的内在含义，此矩阵称为关联矩阵，图中每个结点代表一列，每边代表一行。则上图为 5×4 矩阵。反过来从这个矩阵出发我们也能画出图。

$$
\begin{array}{cc}
 & \begin{array}{cccc} \textbf{结点}\ 1 & 2 & 3 & 4 \end{array} \\
\boldsymbol{A}= & \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}
\end{array}
\begin{array}{c} \textbf{边}\ ① \\ ② \\ ③ \\ ④ \\ ⑤ \end{array}
$$

第一行代表边①，从结点 1 流出记为-1，从结点 2 流入记为 1。

边①,边②和边③构成了一个回路，称为环（loop）。反映在矩阵上是这三个行向量线性相关。

源于现实问题的关联矩阵，通常描述了问题的结构。如果我们研究一个很大的图，则会构建一个很大的矩阵，但这个矩阵会是稀疏矩阵。

考察矩阵的零空间，即求 $\boldsymbol{A}\mathbf{x}=\mathbf{0}$ 的解。零空间告诉我们列向量线性组合的状态。

$$\boldsymbol{A}\mathbf{x}=\begin{bmatrix} x_2-x_1 \\ x_3-x_2 \\ x_3-x_1 \\ x_4-x_1 \\ x_4-x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}。$$

如果 x 为结点上的电势，则 $\boldsymbol{A}\mathbf{x}$ 给出了每个边上的电势差。求解可以得到零空间为一维 dim N($\boldsymbol{A}$)=1，它的基就是 $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$，解集则是 $\mathbf{x}=c\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$，代表等电势，说明等电势条件下不会有电流产生。常数 c 的确定需要边界条件，比如我们将结点 4 接地，

则 $x_4=0$。

若求 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 的解，则相当于在给定了电压 **b** 的情况下，求各点的电势，但实际上我们得不到电势的准确值，因为零空间有常数解 **c**，各点得到的电势需要加上常数 c，这很类似于求积分要加上常函数，常数值需要边界条件来确定。

矩阵的列数为 4，而其零空间的维数为 1，则矩阵的秩为 3，矩阵第 1 列，第 2 列，第 4 列的列向量线性无关。

考察矩阵列空间，一个重要的问题就是对于什么样的 **b**，$\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 有解。边①,边②和边③构成了环，这三个行向量线性相关，同样的情况还有边④,边⑤和边③构成的环。于是 **b** 的分量需要满足 $b_1-b_2+b_3=0$ 以及 $b_3-b_4+b_5=0$。如果把边①,边②，边④,边⑤构成的大环也表示出来则还可以得到一个等式，但实际上这个等式就是之前这两个等式的组合。这两个等式就是基尔霍夫电压定律（Kirchhoff's Voltage law）,即环路电势差之和为零。

矩阵的左零空间是满足 $\boldsymbol{A}^{\mathrm{T}}\mathbf{y}=\mathbf{0}$ 的向量 **y** 的集合。因为矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 有 5 列，且矩阵的秩为 3，因此矩阵的左零空间维数为 2。这反应了行向量的线性关系，整个“图”中，环数为 2。

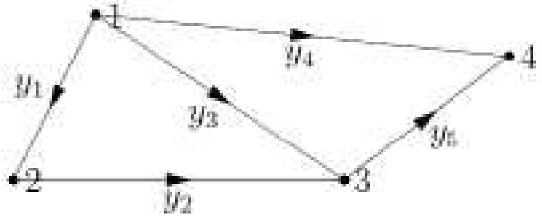
$$\boldsymbol{A}^{\mathrm{T}}\mathbf{y}=\begin{bmatrix}-1&0&-1&-1&0\\1&-1&0&0&0\\0&1&1&0&-1\\0&0&0&1&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\\y_4\\y_5\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

。**y** 的分量的值为“边”上的电流。

在电势差和电流之间建立联系就是欧姆定律（Ohm's Law）

$$\text{电势}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}\xrightarrow[\mathbf{e}=\boldsymbol{A}\mathbf{x}]{\text{矩阵 }\boldsymbol{A}}\text{电势差}\begin{bmatrix}x_2-x_1\\x_3-x_2\\x_3-x_1\\x_4-x_1\\x_4-x_3\end{bmatrix}\xrightarrow[\text{欧姆定律 }\mathbf{y}=\boldsymbol{C}\mathbf{e}]{\text{矩阵 }\boldsymbol{C}}\text{电流}\begin{bmatrix}y_1\\y_2\\y_3\\y_4\\y_5\end{bmatrix}\xrightarrow{\text{矩阵 }\boldsymbol{A}^{\mathrm{T}}}\underset{\text{基尔霍夫定律}}{\boldsymbol{A}^{\mathrm{T}}\mathbf{y}=\mathbf{0}}$$

我们求解 $\boldsymbol{A}^{\mathrm{T}}\mathbf{y}=\mathbf{0}$ 就是在求 5 个满足基尔霍夫电流定律（Kirchhoff's Law）的电流值。



$\boldsymbol{A}^T\mathbf{y}=\mathbf{0}$ 的方程形式 $\begin{aligned} -y_1-y_3-y_4 &= 0 \\ y_1-y_2 &= 0 \\ y_2+y_3-y_5 &= 0 \\ y_4+y_5 &= 0 \end{aligned}$，每一个方程关于一个结点，方程表示结点电流值为 0，即流入等于流出。

从图上解方程，而不是采用消元法解方程。如果我们设定 $y_1=1$，并且让 y_1，y_2 和 y_3 组成的回路的“环流“为 0，则有 $y_2=1$，$y_3=-1$。可解得 $\mathbf{y}=\begin{bmatrix}1\\1\\-1\\0\\0\end{bmatrix}$。取另一个回路的环流为 0，则有 $y_3=1$，$y_4=-1$，$y_5=1$。$\mathbf{y}=\begin{bmatrix}0\\0\\1\\-1\\1\end{bmatrix}$。

如果设定 y_1，y_2，y_4 和 y_5 组成的大回路环流为 0，则可以得到另一个向量 $\mathbf{y}$，而该向量在零空间内，是前两个向量的线性组合。

考察矩阵的行空间，因为矩阵 r=3，所以存在 3 个线性无关的向量。第 1 行，第 2 行和第 4 行为线性无关，在“图”中，边①,边②和边④构成了一张小图，这三个边没有形成回路。线性相关问题等价于形成回路。没有回路的小图包含 4 个结点和 3 条边，再添加一条边就会产生回路，在矩阵里表现为在第 1 行，第 2 行和第 4 行之上再添加一个行向量就会变为线性相关。没有回路的图称为“树”。

思考一下维数公式的在“图”中的意义：

左零空间维数 dim N($\boldsymbol{A}^T$)=m-r；

等价于“环”数量=“边”数量-（“结点”数量-1）；

即 Eular 公式：“结点”-“边”+“环”=1。对所有图都成立。

矩阵的秩 r=“结点”-1，因为 r 表示了线性无关的边的数目，也就是“树”中“边”的数目。

$$\text{电势}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}\xrightarrow[\mathbf{e}=\boldsymbol{A}\mathbf{x}]{\text{矩阵 }\boldsymbol{A}}\text{电势差}\begin{bmatrix}x_2-x_1\\x_3-x_2\\x_3-x_1\\x_4-x_1\\x_4-x_3\end{bmatrix}\xrightarrow[\text{欧姆定律 }\mathbf{y}=\boldsymbol{C}\mathbf{e}]{\text{矩阵 }\boldsymbol{C}}\text{电流}\begin{bmatrix}y_1\\y_2\\y_3\\y_4\\y_5\end{bmatrix}\xrightarrow[\text{基尔霍夫定律}]{\text{矩阵 }\boldsymbol{A}^T}\boldsymbol{A}^T\mathbf{y}=\mathbf{f}$$

之前的讨论都是针对于一个无源的电场，如果加入电源则情况又不同，例如加入电流源相当于将基尔霍夫定律的方程变为 $\boldsymbol{A}^T\mathbf{y}=\mathbf{f}$，$\mathbf{f}$ 就是外部流入的电流。

将 $\mathbf{e}=A\mathbf{x}$,$\mathbf{y}=C\mathbf{e}$,$A^{\mathrm{T}}\mathbf{y}=\mathbf{f}$,三个等式结合得到应用数学中的基本方程 $A^{\mathrm{T}}CA\mathbf{x}=\mathbf{f}$。

关于方程 $A^{\mathrm{T}}CA\mathbf{x}=\mathbf{f}$ 的更多内容可以阅读 GS 老先生 08 的书 “Computational science and engineering” 的第二章。

第 13 讲 复习一

Exam 1 review

本讲为考前复习课，考试范围就是 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 这个单元，重点是长方形矩阵，与此相关的概念包括零空间、左零空间、秩、向量空间、子空间，特别是四个基本子空间。当矩阵为可逆的方阵时，很多性质是一目了然的，但是对于长方形矩阵则需要多加注意。

问题一

向量 **u**，**v** 和 **w** 是 $\mathbf{R}^7$ 空间中的非零向量。它们张成了 $\mathbf{R}^7$ 空间中的一个子空间，那么这个子空间的维数可能是多少？

解答：1，2 或者 3。空间的基的个数不超过 3 个，所以其维数也不会超过 3。因为是非零向量，所以不能是 0。

问题二

给定矩阵 $\boldsymbol{U}$ 为 5×3 阶梯型矩阵，其秩 r=3。

GS 老师这里设定的矩阵 $\boldsymbol{U}$ 就是行最简阶梯型矩阵 $\boldsymbol{R}$，他在视频中说这是多年前卷子上的题目，那时候他都写 $\boldsymbol{U}$ 而不是 $\boldsymbol{R}$。而在 11 年的笔记中已经改为 $\boldsymbol{R}$ 了。此处我们与视频保持相同。

1.求 $\boldsymbol{U}$ 的零空间 N($\boldsymbol{U}$)

解答：因为列数为 3 且秩为 3，其列向量线性无关，则 $\boldsymbol{U}\mathbf{x}=\mathbf{0}$ 只有零解，所以其零空间 $N(\boldsymbol{U})=\{\mathbf{0}\}=\{\begin{bmatrix}0\\0\\0\end{bmatrix}\}$

2.令 $\boldsymbol{B}=\begin{bmatrix}U\\2U\end{bmatrix}$，求 $\boldsymbol{B}$ 的行最简阶梯矩阵？

解答：$\begin{bmatrix}U\\0\end{bmatrix}$

3.$\boldsymbol{B}$ 的秩？

解答：3。

4.令 $\boldsymbol{C}=\begin{bmatrix}U & U\\U & 0\end{bmatrix}$，求它的行最简阶梯矩阵？

解答：进行行消元，$\begin{bmatrix}U & U\\U & 0\end{bmatrix}\to\begin{bmatrix}U & U\\0 & -U\end{bmatrix}\to\begin{bmatrix}U & 0\\0 & -U\end{bmatrix}\to\begin{bmatrix}U & 0\\0 & U\end{bmatrix}$

最后还要做的就是将 $\boldsymbol{R}$ 中最后两行的零行向量下移到矩阵最下面。

5.$\boldsymbol{C}$ 的秩？

解答：6，***B***的秩为 3。

6.求 ***C*** 左零空间的维数 dim N($\boldsymbol{C}^T$)?

解答：m=10 而 r=6，所以 dim N($\boldsymbol{C}^T$)=4。

问题三

$\boldsymbol{A}\mathbf{x}=\begin{bmatrix}2\\4\\2\end{bmatrix}$，其中 $\mathbf{x}=\begin{bmatrix}2\\0\\0\end{bmatrix}+c\begin{bmatrix}1\\1\\0\end{bmatrix}+d\begin{bmatrix}0\\0\\1\end{bmatrix}$为通解。

1.矩阵 ***A*** 的形状？

解答：矩阵 ***A*** 为 3×3 矩阵，因为 **x** 的分量数即为 ***A*** 的列数，而 **b** 的分量数为 ***A*** 的行数。

2.矩阵 ***A*** 的行空间的维数？

解答：从通解的形式可以看出 ***A*** 零空间的维数为 2 则其行空间的维数为 3-2=1。

3.求矩阵 ***A***？

解答：代入特解 $\boldsymbol{A}\begin{bmatrix}2\\0\\0\end{bmatrix}$=2×（列 1）= $\begin{bmatrix}2\\4\\2\end{bmatrix}$可知，A 的第一列（列 1）= $\begin{bmatrix}1\\2\\1\end{bmatrix}$，代入零空间的特解 $\boldsymbol{A}\begin{bmatrix}1\\1\\0\end{bmatrix}$=（列 1）+（列 2）=0，可知（列 2）= $\begin{bmatrix}-1\\-2\\-1\end{bmatrix}$，代入另一特解 $\boldsymbol{A}\begin{bmatrix}0\\0\\1\end{bmatrix}$=（列 3）=0，可知 $\boldsymbol{A}=\begin{bmatrix}1&-1&0\\2&-2&0\\1&-1&0\end{bmatrix}$。

4.什么样的向量 **b** 使得方程 ***A*x**=**b** 有解 **x**？

解答：当 **b** 在矩阵 ***A*** 的列空间内的时候，方程有解。本题中 **b** 需为$\begin{bmatrix}1\\2\\1\end{bmatrix}$的倍数。

本题中矩阵的秩为 1，因此零空间很大。请注意系数矩阵满秩的情况，我们曾经花了不少时间进行讨论。

小问题

1.***A*** 是方阵，零空间只有零向量 **0**，$\boldsymbol{A}^T$ 的零空间？

解答：$\boldsymbol{A}^T$ 的零空间也只有零向量。

2.所有的 5×5 可逆矩阵，是否构成 5×5 矩阵空间的子空间？

解答：不行，不包含零矩阵 ***0***，不是子空间。顺便说一下奇异阵也不是一个子空间。

3.判断：如果 $\boldsymbol{B}^2=\boldsymbol{0}$，是否必有 $\boldsymbol{B}=\boldsymbol{0}$？

解答：否，反例 $\boldsymbol{B}=\begin{bmatrix}0&1\\0&0\end{bmatrix}$。

4.判断：方程组 $\boldsymbol{A}\mathbf{x}=\mathbf{b}$ 具有 n 个方程 n 个未知数，若矩阵 $\boldsymbol{A}$ 的列向量线性无关，则 $\mathbf{b}$ 取任意向量，方程均有解。

解答：是，$\boldsymbol{A}$ 为可逆矩阵，所以 $\mathbf{x}=\boldsymbol{A}^{-1}\mathbf{b}$ 是唯一解。

问题四

矩阵 $\boldsymbol{B}=\boldsymbol{C}\boldsymbol{D}=\begin{bmatrix}1&1&0\\0&1&0\\1&0&1\end{bmatrix}\begin{bmatrix}1&0&-1&2\\0&1&1&-1\\0&0&0&0\end{bmatrix}$

1.给出矩阵 $\boldsymbol{B}$ 零空间的一组基？

解答：因为矩阵 $\boldsymbol{C}$ 为可逆矩阵，因此矩阵 $\boldsymbol{B}$ 零空间与矩阵 $\boldsymbol{D}$ 的零空间相同。

矩阵 $\boldsymbol{D}$ 已经是行最简阶梯矩阵了，可解得其零空间的一组基为 $\begin{bmatrix}1\\-1\\1\\0\end{bmatrix}$ 和 $\begin{bmatrix}-2\\1\\0\\1\end{bmatrix}$。

左乘可逆矩阵，相当于对行向量进行可逆的线性组合。既不改变行空间，也不改变零空间。

若 $\mathbf{x}$ 满足 $\boldsymbol{B}\mathbf{x}=\mathbf{0}$，则有 $\boldsymbol{C}^{-1}\boldsymbol{B}\mathbf{x}=\mathbf{0}$，即 $\boldsymbol{D}\mathbf{x}=\mathbf{0}$，$\boldsymbol{B}$ 零空间中的向量均在 $\boldsymbol{D}$ 的零空间中；若 $\mathbf{y}$ 满足 $\boldsymbol{D}\mathbf{y}=\mathbf{0}$，则有 $\boldsymbol{C}\boldsymbol{D}\mathbf{y}=\mathbf{0}=\boldsymbol{B}\mathbf{y}$，则 $\boldsymbol{D}$ 零空间中的向量均在 $\boldsymbol{B}$ 的零空间中，因此 $\boldsymbol{B}$ 零空间与矩阵 $\boldsymbol{D}$ 的零空间相同。

2.求 $\boldsymbol{B}\mathbf{x}=\begin{bmatrix}1\\0\\1\end{bmatrix}$ 的通解？

解答：向量 $\mathbf{b}=\begin{bmatrix}1\\0\\1\end{bmatrix}$ 与矩阵 $\boldsymbol{C}$ 第一列列向量完全相同，则问题变为 $\boldsymbol{D}\mathbf{x}=\begin{bmatrix}1\\0\\0\end{bmatrix}$，求解方程得到一个特解为 $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$，则其通解就是 $\mathbf{x}=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}+c\begin{bmatrix}1\\-1\\1\\0\end{bmatrix}+d\begin{bmatrix}-2\\1\\0\\1\end{bmatrix}$

小问题

1.判断：$\boldsymbol{A}$ 为方阵，则它的行空间和列空间相同？

解答：否，只是空间的维数相等，但空间不一定相同。反例 $\boldsymbol{B}=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$。都是 $\mathbf{R}^n$ 空间中的 r 维子空间，但是彼此不一定相同。对称矩阵的行空间和列空间相同。不是对称矩阵也可以两者相同，比如可逆矩阵，再比如可逆的 $n-1\times n-1$ 矩阵加上半圈 0 变成的 $n\times n$ 矩阵，加上一圈 0……

2.判断：矩阵 $\boldsymbol{A}$ 和 $-\boldsymbol{A}$ 的四个子空间相同？

解答：是，相同。

3.判断：如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的四个子空间相同，则 $\boldsymbol{A}$ 一定是 $\boldsymbol{B}$ 的倍数？

解答：错误，$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的可以为同阶可逆方阵。

4.如果我们对矩阵 $\boldsymbol{A}$ 中的两行做行交换，四个子空间中不变的是哪些？

解答：是行空间和零空间。

5.为什么向量 $\mathbf{v}=\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ 不可能既在矩阵 $\boldsymbol{A}$ 的零空间中，又成为 $\boldsymbol{A}$ 的一个行向量？

解答：反设成立。$\mathbf{v}$ 在零空间中，则有 $\boldsymbol{A}\mathbf{v}=\mathbf{0}$，而它 $\mathbf{v}^T$ 也是 $\boldsymbol{A}$ 的行向量，则 $\mathbf{v}$ 和这个行向量 $\mathbf{v}^T$ 做乘法得到数值为 14 而不是 0，与 $\boldsymbol{A}\mathbf{v}=\mathbf{0}$ 矛盾。实际上 $\mathbf{v}$ 若在零空间中，则根本不会再零空间中，在下一章中我们将看到矩阵的零空间和其行空间是正交的。