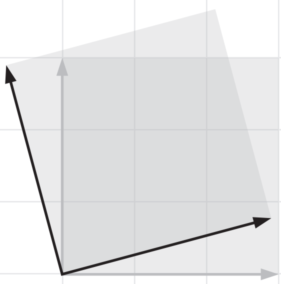


线性代数入门

梁 鑫　田 垠　杨一龙　编著

清華大學出版社
北 京

内 容 简 介

线性代数是当代大学生的必修科目，也是当前科学技术领域的数学基础和通用语言. 随着数据科学的发展，线性代数的地位越来越重要. 本书前六章是线性代数的基础，主要讨论 $\mathbb{R}^n$ 的结构，内容包括向量、矩阵、子空间、内积、行列式、特征值等；后两章是对基础内容的升华，主要讨论抽象的线性空间、线性映射、内积等内容.

本书可作为线性代数课程的教材或参考书.

图书在版编目(CIP)数据

线性代数入门/梁鑫，田垠，杨一龙编著. —北京：清华大学出版社，2022.6

ISBN 978-7-302-60971-1

Ⅰ. ①线…　Ⅱ. ①梁…　②田…　③杨…　Ⅲ. ①线性代数　Ⅳ. ①O151.2

中国版本图书馆 CIP 数据核字(2022)第 089506 号

责任编辑： 刘　颖
封面设计： 傅瑞学
责任校对： 王淑云
责任印制： 丛怀宇

出版发行： 清华大学出版社
网　　址： http://www.tup.com.cn, http://www.wqbook.com
地　　址： 北京清华大学学研大厦 A 座　　**邮　　编：** 100084
社 总 机： 010-83470000　　**邮　　购：** 010-62786544
投稿与读者服务： 010-62776969, c-service@tup.tsinghua.edu.cn
质 量 反 馈： 010-62772015, zhiliang@tup.tsinghua.edu.cn
印 装 者： 艺通印刷（天津）有限公司
经　　销： 全国新华书店
开　　本： 185mm × 260mm　　**印　　张：** 21　　**字　　数：** 468 千字
版　　次： 2022 年 7 月第 1 版　　**印　　次：** 2022 年 7 月第 1 次印刷
定　　价： 59.80 元

产品编号：097005-01

前　言

线性代数是当代高等院校学生的必修科目，也是当前科学技术领域的数学基础和通用语言．随着时代的发展和社会的进步，计算机技术和数据科学逐步渗透、应用甚至开始主导不少科学研究和工程技术领域．在新形势下，线性代数变得越来越重要．另一方面，为了更好地建设国家、服务社会，教师和学生都迫切地希望能更快更好地教会、学会基础课程．教好、学好线性代数，离不开体系完整、简明易懂的入门教材．本书正是为此而编写的．

线性代数研究线性空间和线性映射理论，其基础是向量和矩阵的性质．本书精心选择了教学内容，重点阐述最主要的、最常用的知识点，并辅以常见而易于理解的工程实例．

绪论简述了线性代数的整体框架，作者希望读者能够在学前和学后认真阅读，以便纲举目张，对线性代数有一个整体性认识．第 0 章主要用来应对中小学数学教育内容不断减少的问题，供读者了解数学的基本逻辑工具．第 1 章至第 4 章是最基本的主干内容，可以认为是中学数学的简单拓展，是各个专业的学生都务求精熟的部分；第 5 章至第 6 章是另一部分基本内容，是自然科学、工程技术、社会科学等专业的学生应当掌握的内容．这两部分是线性代数的初等部分，是抽象理论的具体实例．第 7 章至第 8 章是线性代数中抽象理论的介绍，是线性代数的真正入门．读者应注意随时对照初等部分，发现相同和不同，加深对抽象理论的理解．掌握这部分内容的读者应能在学习其他课程中不断发现和应用线性空间和线性映射理论．这样安排教学内容，既利于读者由浅入深，自具体而抽象地学习线性代数，又便于课时不足的读者只学习部分内容就能了解到线性代数的概貌．

“熟能生巧”“练习铸就大师”．学好线性代数，必须得有习题辅助．本书安排了大量不同类型、由易到难的习题，以供读者使用．只有当学生清楚地知道自己答题是否正确时，才能谈得上学会相应的知识．“学而不思则罔，思而不学则殆．”中学生常有些不好的学习习惯，如只看正文不做题，或不看正文只做题．作者希望读者能在练习中不断熟悉和复习相关知识，当习题做不出时反复阅读思考正文中的知识点，尝试使用相关知识点来解决问题．另外，读者应当认识到，有些较难的习题不是每个人都能做出，但只要尽力思考，即使没有找到解决问题的办法，也已经在思考过程中复习了相关知识，锻炼了分析能力，并非一无所获．

本书将定义、定理、命题、推论、例、注等统一编号为 $a.b.c$，表示第 a 章第 b 节第 c 项定义、定理或其他内容，其中加 ☕ 的章节属于选讲内容；将习题、阅读材料等集中放在每节之后，并统一编号为 $a.b.c$，表示第 a 章第 b 节第 c 项习题或阅读材料，其中

加 ☕ 或 ☕☕ 的习题是较难或更难的选做题. 另外, 为提示读者, 本书中的例以 ☺ 结束, 证以 □ 结束, 其他部分材料换用不同于正文的字体.

本书最后列出名词索引、人名表、符号表, 以供读者查阅. 正文中出现的人名除华人外一般以拉丁转写形式出现, 并在人名表中提供了原文和常见译法.

本书可作为高等院校线性代数课程的教材, 也可作为自学或复习的参考书. 本着体系的完整性和严谨性, 作者对所有命题都提供了证明, 有些命题还有多个证明, 从而使知识体系在逻辑上尽量扎实可靠. 教师可以根据教学要求和课时限制自行调整. 根据作者的授课经验, 完整讲述本书内容, 第 1 章至第 4 章用 35~40 学时, 第 5 章至第 6 章用 15~20 学时, 第 7 章至第 8 章用 16~20 学时, 供参考.

本书参考了不少国内外教材, 部分习题也来源于此. 现列出供读者参阅:

[1] G. Strang. 线性代数 [M]. 5 版. 北京: 清华大学出版社, 2019.
[2] А.И. Кострикин. 代数学引论（一）[M]. 2 版. 张英伯, 译. 北京: 高等教育出版社, 2006.
[3] А.И. Кострикин. 代数学引论（二）[M]. 3 版. 牛凤文, 译. 北京: 高等教育出版社, 2008.
[4] 丘维声. 简明线性代数 [M]. 北京: 北京大学出版社, 2002.
[5] 蓝以中. 高等代数简明教程（上、下）[M]. 北京: 北京大学出版社, 2002.
[6] 史明仁. 线性代数六百证明题详解 [M]. 北京: 北京科学技术出版社, 1985.

本书是在清华大学数学科学系和丘成桐数学科学中心的领导与线性代数教学团队的大力支持下完成的, 并在 2020 年初稿完成后在清华大学大范围试用了两年, 获得了大量颇有价值的意见. 作者感谢张友金、朱敏娴、刘思齐、王浩然、曹晋、蔚辉、孙晟昊、袁瑶、刘余及其他同事对写书的倡议和支持、对初稿的指正与对体系的建议, 并感谢探微书院化 12 班刘桓瑀同学对书稿的建议. 作者感谢本书的责任编辑刘颖为本书付出的辛勤劳动.

本书得到了清华大学教学改革项目的资助.

作者始终努力使本书准确可靠, 但精力和水平所限, 书中的错误在所难免. 欢迎广大读者对本书提出宝贵意见, 指出本书的逻辑上的、文字上的、排版上的任何错误. 作者向所有关心本书的人致以诚挚的感谢!

梁 鑫　田 垠　杨一龙

2022 年 1 月于北京

目　　录

绪　论

线性代数这门学科，以研究**线性系统**为目的，是一个从理论到应用都相当完善的数学分支. 线性系统是什么呢?

我们先从系统说起. 系统是指由若干个相互作用和相互依赖的事物组合而成的具有特定功能的一个整体. 例如一个由若干部件组成的机器，可以对“输入”进行运算或处理，得到“输出”:

$$\text{输入} \to \text{系统/机器} \to \text{输出}.$$

而线性系统是指，这个系统的输入与输出符合叠加原理，即如下两条性质：输入放大或缩小某一倍数，则产生的输出也放大或缩小相同倍数，这称为**齐次性**；两组输入所产生的总输出是二者分别产生的独立输出之和，这称为**可加性**.

下面来看几个具体的例子.

- 一个大型代工厂是一个系统：以工人的工时作为输入，以不同产品的数量作为输出. 二者的关系满足叠加原理.
- 由刚性杆组成的复杂支架是一个系统：在支架上端放置若干重物，将重物重量作为输入;支架下端以若干地秤支撑,将地秤测得的重量作为输出.根据 Newton 定律，二者的关系满足叠加原理.
- 由电阻组成的电路是一个系统：在若干对节点间接入电压表，将电压值作为输入；在若干节点处接入电流表，将电流值作为输出. 根据 Kirchhoff 定律，二者的关系满足叠加原理.

这些例子是大量实际问题中的几个代表. 今后读者会发现，日常生产生活中总会遇到类似的问题. 这些问题有时规模很大、未知数很多，但都可以利用线性代数知识使用计算机来解决. 我们将看到，线性系统虽然应用十分丰富，但它的数学本质并不复杂.

给定一个线性系统，它内部的运作方式我们还不清楚. 能够清楚知道的是，它需要 n 个输入 $x_1, x_2, \cdots, x_n$，并生成 m 个输出 $y_1, y_2, \cdots, y_m$. 换句话说，目前它还是一个黑盒子，给出任意输入，我们不知道它会输出什么.

想要搞清楚系统如何运作，我们可以对它进行测试，即给定一组输入，测量其输出. 我们先考虑最简单的输入，即只对某一个输入项输入:

- 先指定 $x_1 = 1$，其他输入项是 0，这时我们测出输出项依次为 $y_1 = a_{11}, y_2 = a_{21}, \cdots, y_m = a_{m1}$.
- 再指定 $x_2 = 1$,其他输入项是 0,这时输出项依次为 $y_1 = a_{12}, y_2 = a_{22}, \cdots, y_m =$

a_{m2}.

- 依次类推,直到指定 $x_n = 1$,其他输入项是 0,这时输出项依次为 $y_1 = a_{1n}, y_2 = a_{2n}, \cdots, y_m = a_{mn}$.

根据叠加原理中的齐次性，我们可以得到以下结论:

- 当 x_1 任意,其他输入项是 0 时,输出项必然为 $y_1 = a_{11}x_1, y_2 = a_{21}x_1, \cdots, y_m = a_{m1}x_1$.
- 当 x_2 任意,其他输入项是 0 时,输出项必然为 $y_1 = a_{12}x_2, y_2 = a_{22}x_2, \cdots, y_m = a_{m2}x_2$.
- 依次类推，当 x_n 任意，其他输入项是 0 时，输出项必然为 $y_1 = a_{1n}x_n, y_2 = a_{2n}x_n, \cdots, y_m = a_{mn}x_n$.

再根据叠加原理中的可加性，我们就知道，任意输入项 $x_1, x_2, \cdots, x_n$ 与其输出项 $y_1, y_2, \cdots, y_m$ 满足:

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n = y_1, \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n = y_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m. \end{cases} \tag{0.0.0}$$

可以看到，任意输入项和输出项之间的关系构成了一个多元一次方程组，又称为线性方程组. 这就是线性系统得名的原因.

容易看到，整个线性系统由 a_{ij} 这 mn 个数唯一确定. 换句话说，只要进行了前述的 n 组测试，得到了这些 a_{ij}，这个线性系统对我们来说就变得透明，不再是一个黑盒子了.

四个基本问题 现在假设某个线性系统已知，或者说任意输入项与输出项之间的关系已知. 如果系统有一个未知的输入，能否通过观察其输出来反推这个输入呢? 对于不同的系统，可能会出现以下三种情形:

1. 存在唯一一组输入得到给定的输出. 例如，由电阻组成的电路这一系统，对输出的电流，只可能有唯一的输入电压.
2. 存在不只一组输入得到给定的输出. 那么根据叠加原理，一定有无穷多组输入都能得到给定的输出，这是因为如果两组输入 $u_1, u_2, \cdots, u_n$ 和 $v_1, v_2, \cdots, v_n$ 得到同一输出，那么 $tu_1 + (1-t)v_1, tu_2 + (1-t)v_2, \cdots, tu_n + (1-t)v_n$ 也将会得到相同的输出. 例如，服装厂这一系统，对输出指定数量的衣服，既可以是这个工人花费工时做的，也可以是那个工人花费工时做的，也可能是这些工人按照任意比例分配工时完成的. 这里的输入有无数种可能.
3. 不存在输入得到给定的输出. 事实上这意味着，无论输入什么都不可能得到这种输出! 尽管听起来略显怪异，但实践中仍有可能发生. 例如，CT 扫描机这一

系统，输入人，输出 X 光片. 受灰尘、机械错位、计算机误差等因素影响，可能得到的 X 光片和真实情况有微小偏差. 这个微小偏差很可能使得这个 X 光片不能百分之百地符合任何人的真实情况. 另一方面，实际上这个误差并不影响医生诊断，X 光片不需要完全真实，只要它和真实情况足够接近就足够了. 换句话说，测量总有误差，因此测量结果并不一定是系统的真实输出. 但理论上讲，测量值和真实的输出应该非常接近. 尽管测量结果不可能从输入得到，但是我们可以在所有可能的输出结果中找到最接近测量结果的那个输出. 它应该是最有可能的真实的系统输出.

在数学上，我们是在研究线性方程组 (0.0.0)，其中输出 $y_1, y_2, \cdots, y_m$ 已知，需要求解的是未知的输入 $x_1, x_2, \cdots, x_n$.

在上面提到的三种情形中，前两种情形分别意味着线性方程组 (0.0.0) 有唯一解或有无穷多解，此时我们希望计算出解. 对这两种情形的考察构成了线性代数的一个核心问题——**线性方程组求解问题**.

而后一种情形意味着线性方程组 (0.0.0) 无解，这时我们希望计算出某组 $x_1, x_2, \cdots, x_n$，使得方程组左端的值和右端给定的值最接近. 对这种情形的考察构成了线性代数的一个重要问题——**最小二乘问题**.

在很多实际问题中，我们常常需要考虑自反馈系统，即将给定输入通过系统演化产生的输出作为新的输入继续演化，并不断重复这一过程直到达到我们需要的演化次数为止，或者直到系统稳定下来，即输出和输入相同. 例如，对某个范围的人群，将每一年的人口数量作为输入，都可以得到作为输出的第二年的人口数量，而这个输出可以再次作为输入，得到第三年的人口数量，以此类推. 我们希望得到第三十年的人口数量，或者当人口数量不再变化时的人口数量. 我们可以提出问题:

1. 如何计算出达到演化次数时的输出? 当然我们可以逐次计算，但是否有快速计算的方法?
2. 如何计算系统稳定时的输入或输出?

对这两个问题的考察构成了线性代数的另一个核心问题——**特征值问题**.

有时我们希望用简单的系统来模拟一个复杂的系统，即用远少于 mn 个数来描述由前述 mn 个数所代表的系统. 例如，一张数码照片，由于计算机存储不足或网络带宽限制，我们希望对它进行压缩. 压缩的过程中可能会丢失一些信息，但我们希望，在保证压缩程度的同时丢失的信息尽量少. 这需要考虑这样的问题:

1. 如何计算出这样的简单系统?
2. 如何衡量模拟的好坏?

对这两个问题的考察构成了线性代数的另一个重要问题——**奇异值问题**.

线性方程组求解问题、最小二乘问题、特征值问题和奇异值问题，是线性代数理论与应用中的**基本问题**，矩阵计算就是围绕着这四个问题展开的。

两个核心概念 数学上，为了便于开展对前述的系统、输入、输出等理念及上面提出的基本问题进行研究，我们需要将其数学化.

在考虑线性系统时，我们常常把一组输入或一组输出称为一个**向量**；所有可以叠加的向量构成的集合称为一个**线性空间**. 由于每个输入项和输出项都是一个数，当它们都只有有限多个时，向量，作为一组输入或者一组输出，常常可以写成一个数组.

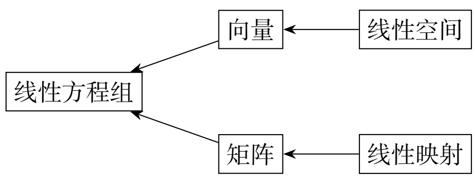
向量，也就是线性空间中的元素，最重要的特点是可以进行线性运算. 这是叠加原理赋予向量的运算.

而线性系统的处理过程，或者说从输入到输出的对应关系，称为一个**线性映射**. 特别地，如果输入项和输出项都只有有限多个，对应的线性映射常常可以写成一个数表，称为**矩阵**.

以线性方程组 (0.0.0) 为例，输入 x_j 组成的数组就是向量 $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$，输出 y_i 组成的数组也是向量 $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$. 而描述系统只需要知道 a_{ij} 的值，它们组成的数表就是矩阵 $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$. 我们还可以把线性方程组 (0.0.0) 写成：

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

可以看到，线性方程组可以把向量和矩阵有机地联系起来.



上面介绍了许多概念. 这些概念总结如下:

- 线性空间和线性映射是对所研究的问题提出的高层次概念, 是我们推导的基石和思考问题的凭借;
- 向量和矩阵是所研究的问题自然导出的低层次概念, 是我们计算的对象和获得结果的媒介;
- 线性方程组是向量和矩阵之间的联系, 很多问题的解决方案常常归结于对一个线性方程组求解.

线性空间和线性映射是线性代数研究中的**核心概念**, 可以说线性代数就是关于线性空间和线性映射的代数学.

研究思想和常用技巧 在学习、掌握到精通线性代数的过程中, 在利用线性代数尝试解决实际问题时, 我们常常使用许多数学思想来攻克问题, 获得解决方案.

(1) 简化问题, 化全局为局部 例如, 前面提到线性方程组的解如果不止一组, 就必然有无穷多组, 但是这无穷多组解是由两组解获得的. 反过来问, 如果我们知道有无穷多组解, 那么这些解能否由有限组解获得? 又如前所述, 我们获得一个有 n 个输入项的线性系统的全部信息, 并不需要知道所有可能的输入对应的输出, 只需要知道 n 组输入对应的输出就足够了. 这二者启发了线性空间的基和子空间分解等概念, 落实在矩阵上就是矩阵分块的技巧.

(2) 分解问题, 化复杂为简单 如前述, 一个有 n 个输入项、m 个输出项的线性系统可以用 mn 个数来表示, 这可能是很复杂的; 但有些线性系统的演化可能产生简单的效果. 如果一个线性系统可以分解成一系列简单的系统的组合, 那么线性系统的特性将容易把握得多. 这启发了映射标准形的概念, 落实在矩阵上就是矩阵分解的技巧.

(3) 归约问题, 化未知为已知 面对一个解决方案尚不明确的问题, 最直接的指导方针并不是随便尝试以期起效, 而是将问题划归成一个解决方案已经存在的问题. 有些读者熟知的数学归纳法就是一例. 如果我们能把某个问题都化归成已获解决的若干类简单问题之一, 例如已提及的标准形, 那么问题将容易获得解决, 这启发了等价类的概念, 落实在矩阵上就是矩阵变换的技巧.

矩阵分块、矩阵分解、矩阵变换是线性代数中最常用的推理和计算技巧. 可以看到, 矩阵是利用线性代数解决问题的**计算重点**.

总的来说, 我们这门课程将围绕着上述四个基本问题、两个核心概念展开, 利用上述研究思想和推理技巧, 来阐述完善的数学理论, 推导有效的计算方法, 并为后续的理论科学和应用技术课程提供必要的基础.

下面我们再举一些小例子, 读者可以先自行思考一下问题的解决之道.

- 解方程组: 一百根参数已知的相同弹簧串联在一起, 两端挂在给定间距的天花板和地板上, 如何计算每根弹簧拉伸的长度.
- 最小二乘: 有人说严刑峻法能威慑潜在犯罪分子从而遏制犯罪; 有人说严刑峻法变相鼓励犯罪分子犯重罪; 现有世界上 58 个国家死刑执行率与重罪率的统

计数据，如何用数据来判断不同说法的正确性.

- 特征值：甲有 100 元，乙和丙没有钱；甲将自己的钱的一半给乙，乙再将自己的钱的三分之一给丙，丙再将自己的钱的四分之一给甲，然后甲再将自己的钱的一半给乙，依次循环；如何计算最终达到稳定状态时三人的钱数.
- 奇异值：给定十个村庄的位置，修一条笔直的高速公路，再修一些小路将村庄和高速公路连通，如何修路能够使各小路长度的平方和最小.

在学习过线性代数后，读者将能够用一张纸徒手计算，得到上述问题的答案.

还有一些读者在中学时可能思考过但不曾得到结论的数学问题，在学习过线性代数后，将完全能够口述其中的原理.

- 多元一次方程组的所有解如何只用有限多个解表示；
- 如何获得一个一元高次方程的所有解的近似值；
- 平面上的二次曲线只有三种非退化的情形；
- 平面上保持距离的变换只有平移、旋转、镜射和滑移镜射；
- 平面上不改变共线关系的变换一定是保持距离的变换和正压缩的复合.

现在就让我们开始学习精巧而美妙的线性代数吧！

第 0 章 预备知识

0.1 逻辑与集合

首先回顾数学中常用的逻辑术语.

数学中, 可以判断真假的陈述句称为**命题**. 其中判断为真的命题称为**真命题**, 判断为假的命题称为**假命题**. 一个命题为真, 常称为命题成立.

在进行逻辑推理时, 最常使用的语句是

如果命题 A 成立, 那么命题 B 成立　或简写成　若 A, 则 B.

其中命题 A 常称为**条件**, 命题 B 常称为**结论**. 如果由 A 通过推理可以得出 B, 则上述语句是真命题. 此时称 A **蕴涵** B, 又称 A 是 B 的**充分条件**, 也称 B 是 A 的**必要条件**, 记作 $A \Rightarrow B$, 或 $B \Leftarrow A$. 如果 A 既是 B 的充分条件又是 B 的必要条件, 则称 A 是 B 的**充分必要条件**, 也称 A 和 B **等价**, 或 A **当且仅当** B, 记作 $A \Leftrightarrow B$.

例如, 平面上两个三角形, 如果全等, 则一定相似. 因此全等是相似的充分条件, 相似是全等的必要条件, 即 "全等 $\Rightarrow$ 相似". 反过来, 如果两个三角形相似, 并不一定全等. 因此相似不是全等的充分条件, 全等也不是相似的必要条件. 进而可知, 相似和全等并不是充分必要条件.

命题添加逻辑联结词可以得到新的命题. 用 "且" 联结命题 A, B, 得到新命题 "A 且 B", 记作 $A \wedge B$. A 且 B 成立当且仅当 A 和 B 同时成立. 用 "或" 联结命题 A, B, 得到新命题 "A 或 B", 记作 $A \vee B$. A 或 B 成立当且仅当 A 和 B 至少有一个成立. 用 "非" 联结命题 A, 得到新命题 "非 A", 记作 $\neg A$. 非 A 成立当且仅当 A 不成立.

含有变量的陈述句不一定能判断真假, 也就不一定是命题, 例如, "矩形是正方形". 如果在这种句子中添加一些对变量的限定词 (称为**量词**), 就能使其成为命题. 例如, "存在一个矩形是正方形" 或 "至少有一个矩形是正方形" 是真命题. "存在" 和 "至少有一个" 这种限定词称为**存在量词**, 用 "$\exists$" 表示. 又如, "所有矩形都是正方形" 或 "任意矩形都是正方形" 或 "每一个矩形都是正方形" 是假命题. "所有" 和 "任意" 这种限定词称为**全称量词**, 用 "$\forall$" 表示.

其次回顾关于集合的基本知识.

在数学研究中, 研究对象常称为**元素**, 而所研究对象的全体称为一个**集合**. 集合最重要的基本性质有两条:

1. 给定集合的元素是确定的，即给定一个集合和一个元素，必须能够对元素是否在集合中这一问题做出明确的判断；
2. 给定集合中的元素互不相同，即集合中的元素不重复出现.

不含任何元素的集合称为**空集合**，记作 $\varnothing$. 最常用的集合是由数组成的集合. 自然数的全体构成的集合称为自然数集，记作 $\mathbb{N}$. 类似地，有整数集 $\mathbb{Z}$、有理数集 $\mathbb{Q}$、实数集 $\mathbb{R}$、复数集 $\mathbb{C}$.

集合可以用列举法或描述法表示. 例如，在平面上取定平面直角坐标系 Oxy 之后，以 O 点为圆心的单位圆上横纵坐标都是整数的点所组成的集合 A 可以用如下记号表示：

$$A=\{(1,0),(0,1),(-1,0),(0,-1)\}=\big\{(x,y)\,\big|\,x,y\in\mathbb{Z},x^2+y^2=1\big\}.$$

如果 x 是集合 X 的元素，就称 x **属于**集合 X，记作 $x\in X$；如果 x 不是集合 X 的元素，称 x **不属于**集合 X，记作 $x\notin X$.

给定集合 X 和 Y，如果 X 中的任意元素都在 Y 中，则称 X 是 Y 的**子集**，或 Y **包含** X，记作 $X\subseteq Y$ 或 $Y\supseteq X$. 如果 X 和 Y 互为子集，则两个集合的元素完全相同，则称 X 与 Y **相等**，记作 $X=Y$. 显然，它与 $X\subseteq Y$ 且 $Y\subseteq X$ 等价. 证明 $X\subseteq Y$ 和 $Y\subseteq X$ 都成立是今后证明 X,Y 两个集合相等的一个基本方法. 如果 $X\subseteq Y$，但 $X\neq Y$，则称 X 是 Y 的**真子集**，记作 $X\subsetneq Y$. 空集认为是任何集合的子集.

集合 X 与 Y 的公共元素组成的集合称为 X 与 Y 的**交（集）**，记作 $X\cap Y$. 显然，$X\cap Y\subseteq X, X\cap Y\subseteq Y$. 如果 X 与 Y 没有公共元素，则 $X\cap Y=\varnothing$. 容易看出，$X\cap Y=X$ 与 $X\subseteq Y$ 等价. 把 X 与 Y 中的所有元素合并在一起组成的集合称为 X 与 Y 的**并（集）**，记作 $X\cup Y$. 显然，$X\subseteq X\cup Y, Y\subseteq X\cup Y$. 容易看出，$X\cup Y=Y$ 与 $X\subseteq Y$ 等价. 从集合 X 中去掉属于 Y 的元素之后剩下的元素组成的集合，称为 X 对 Y 的**差（集）**，记作 $X\backslash Y$. 特别地，如果 $Y\subseteq X$，则称 $X\backslash Y$ 是 X 中 Y 的**补（集）**.

容易发现，逻辑运算的且、或、非，直接对应于集合运算中的交、并、补. 而命题的蕴涵关系对应于集合的包含关系. 我们有以下类比.

例 0.1.1 给定命题 A,B 和集合 S，令

$$X=\{s\in S\,|\,s \text{ 满足条件 } A\},\quad Y=\{s\in S\,|\,s \text{ 满足条件 } B\}.$$

1. $X\cap Y=\{s\in S\,|\,s \text{ 满足条件 } A\wedge B\}$.
2. $X\cup Y=\{s\in S\,|\,s \text{ 满足条件 } A\vee B\}$.
3. $S\backslash X=\{s\in S\,|\,s \text{ 满足条件 } \neg A\}$.
4. $X\subseteq Y$ 和 $A\Rightarrow B$ 等价.

设 X,Y,Z 是集合 S 的子集，容易验证以下运算律：

1. 交满足交换律和结合律，即 $X\cap Y=Y\cap X,(X\cap Y)\cap Z=X\cap(Y\cap Z)$；
2. 并满足交换律和结合律，即 $X\cup Y=Y\cup X,(X\cup Y)\cup Z=X\cup(Y\cup Z)$；

3. 交对并满足分配律，即 $(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z)$；

4. 并对交满足分配律，即 $(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z)$；

5. De Morgan 定律, 即 $S \backslash (X \cap Y) = (S \backslash X) \cup (S \backslash Y), S \backslash (X \cup Y) = (S \backslash X) \cap (S \backslash Y)$；

6. 包含关系具有传递性，即如果 $X \subseteq Y, Y \subseteq Z$，那么 $X \subseteq Z$.

类似地，对命题 A, B, C，有以下运算律：

1. 且满足交换律和结合律，即 $A \wedge B \Leftrightarrow B \wedge A, (A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$；

2. 或满足交换律和结合律，即 $A \vee B \Leftrightarrow B \vee A, (A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$；

3. 且对或满足分配律，即 $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$；

4. 或对且满足分配律，即 $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$；

5. De Morgan 定律，即 $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B), \neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$.

6. 蕴涵关系具有传递性，即如果 $A \Rightarrow B, B \Rightarrow C$，那么 $A \Rightarrow C$.

以上运算规律的证明均留给读者. ☺

习题

练习 0.1.1 画出例 0.1.1 中的所有集合运算律的 Venn 图.

练习 0.1.2 对任意两个命题 A, B，$A \Rightarrow B$ 可以定义为 $B \vee (\neg A)$. 直观上可以如此理解：$B \vee (\neg A)$ 成立等价于 A 不成立或 B 成立，于是当 A 成立时必有 B 成立. 运用例 0.1.1 中的逻辑运算律，证明 $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$.

练习 0.1.3 对任意三个命题 A, B, C，运用例 0.1.1 中的逻辑运算律，证明逻辑运算符合模律，即当 A 是 C 的必要条件时，$(A \wedge B) \vee C \Leftrightarrow A \wedge (B \vee C)$.

练习 0.1.4 对任意两个命题 A, B，我们定义运算异或 $A \oplus B$ 为 $(A \wedge (\neg B)) \vee (B \wedge (\neg A))$.

1. 按例 0.1.1 的条件画出 $A \oplus B$ 对应的 Venn 图.
2. 证明异或满足交换律和结合律.

0.2 间接证明法

数学结论的正确性必须通过逻辑推理加以证明. 除直接从条件出发推出结论外，也可以利用逻辑进行间接证明.

反证法 容易看出，“如果命题 A 成立，则命题 B 成立”，和“如果命题 B 不成立，则命题 A 不成立”这两个命题等价. 这就是**反证法**的理论基础. 反证法是间接证明的方法之一，即当需要证明 $A \Rightarrow B$ 时，先假设命题 B 不成立，再想办法证明命题 A 也不成立. 特别地，当没有命题 A，只有命题 B 时，只需假设命题 B 不成立，然后推出与已知条件或事实相抵触的条件（称为***矛盾***）即可.

例 0.2.1 证明：$\sqrt{2}$ 是无理数. ☺

证 假设命题不成立，即 $\sqrt{2}$ 是有理数，则可令 $\sqrt{2}=\dfrac{n}{m}$，其中 n,m 是互素的整数. 易得 $2m^2=n^2$，于是 n 是偶数. 设 $n=2s$，则 $2m^2=4s^2, m^2=2s^2$，因此 m 也是偶数. 这与 n,m 互素的假设矛盾. □

例 0.2.2 证明：素数有无穷多个. ☺

证 假设命题不成立，即只有有限个素数，设为 $p_1,p_2,\cdots,p_n$. 令 $N=p_1p_2\cdots p_n+1>1$. 显然任何 p_i 都不能整除 N. 这与任何大于 1 的正整数都能写成素数乘积这一结论矛盾. 根据反证法推出，一定存在无穷多个素数. □

由上面两个例子可以看出，反证法往往适用于“命题 B 不成立”比较好处理的情形.

数学归纳法 **数学归纳法**是一种数学证明方法，通常被用于证明某个给定命题在整个自然数范围内成立，包括第一数学归纳法和第二数学归纳法.

（1）第一数学归纳法：给定一个与自然数 n 有关的命题 $P(n)$，如果

1. 可以证明当 n 取初始值 n_0 时命题 $P(n_0)$ 成立；
2. 假设当 $n=k\,(k\geqslant n_0)$ 时命题 $P(k)$ 成立，可以证明当 $n=k+1$ 时命题 $P(k+1)$ 也成立；

则对一切自然数 $n\geqslant n_0$，命题 $P(n)$ 都成立.

例 0.2.3 证明平方求和公式：$1^2+2^2+\cdots+n^2=\dfrac{n(n+1)(2n+1)}{6}$. ☺

证 当 $n=1$ 时，命题显然成立.

假设当 $n=k$ 时命题成立，即 $1^2+2^2+\cdots+k^2=\dfrac{k(k+1)(2k+1)}{6}$. 当 $n=k+1$ 时，可得 $1^2+2^2+\cdots+k^2+(k+1)^2=\dfrac{k(k+1)(2k+1)}{6}+(k+1)^2=\dfrac{(k+1)(k+2)(2k+3)}{6}$，此时命题也成立.

根据第一数学归纳法可知，命题对所有正整数 n 都成立. □

（2）第二数学归纳法：给定一个与自然数 n 有关的命题 $P(n)$，如果

1. 可以证明当 n 取初始值 n_0 时命题 $P(n_0)$ 成立；
2. 假设当 $n\leqslant k\,(k\geqslant n_0)$ 时命题 $P(n)$ 成立，可以证明当 $n=k+1$ 时命题 $P(k+1)$ 也成立；

则对一切自然数 $n\geqslant n_0$，命题 $P(n)$ 都成立.

例 0.2.4 考虑 Fibonacci 数列 $F_0=0,F_1=1,F_n=F_{n-1}+F_{n-2}$，$n=2,3,\cdots$，证明其通项公式 $F_n=\dfrac{x_1^n-x_2^n}{x_1-x_2}$，其中 $x_1=\dfrac{1+\sqrt{5}}{2},x_2=\dfrac{1-\sqrt{5}}{2}$ 是方程 $x^2-x-1=0$ 的两个根. ☺

证　当 $n=0$ 和 $n=1$ 时，命题显然成立.

假设当 $n \leqslant k\,(k \geqslant 2)$ 时命题成立，特别地，当 $n=k-1,k$ 时命题成立. 当 $n=k+1$ 时，有

$$\begin{aligned} F_{k+1} &= F_k + F_{k-1} \\ &= \frac{x_1^k - x_2^k + x_1^{k-1} - x_2^{k-1}}{x_1 - x_2} \\ &= \frac{x_1^{k-1}(x_1+1) - x_2^{k-1}(x_2+1)}{x_1 - x_2} \\ &= \frac{x_1^{k-1}x_1^2 - x_2^{k-1}x_2^2}{x_1 - x_2} = \frac{x_1^{k+1} - x_2^{k+1}}{x_1 - x_2}, \end{aligned}$$

即命题也成立.

根据第二数学归纳法可知，命题对所有自然数 n 都成立. □

由这个例子可以看出，对比第一数学归纳法，第二数学归纳法第二条中的假设条件更强，因此适用范围更广.

读者需要注意，由于在例 0.2.4 的递推步骤中，$n=k+1$ 被化简成了 $n=k-1,k$ 的情形，这就要求 $k \geqslant 2$. 也就是说，在递推步骤中，$k=0,1$ 都是不适用的. 因此，在验证初始值时，既需要验证 $k=0$，也需要验证 $k=1$. 总之，凡是递推步骤不适用的 n 的取值，全都需要单独进行验证.

下面举一个错误使用数学归纳法的例子，其错误的核心就是忘记检验全部初始条件. 读者务必引以为戒！

例 0.2.5　我们将（错误地）证明，世界上所有的马都是同一种颜色的.

当 $n=1$ 时，只有一匹马，命题显然成立.

假设当 $n \leqslant k$ 时命题成立. 当 $n=k+1$ 时，在 $k+1$ 匹马中，除第一匹马外，剩下的 k 匹马是同一种颜色的，另一方面，除第二匹马外，剩下的 k 匹马也是同一种颜色的. 这样一来，全部 $k+1$ 匹马都只能是同一种颜色的.

于是命题对所有正整数 n 都成立.

这个证明的错误之处就在于，递推步骤中实际上隐含了假设 $n \geqslant 3$（读者可以思考一下为什么）. 因此这个归纳过程既需要验证 $n=1$ 的情形，也需要验证 $n=2$ 的情形. 而验证一下很容易发现，$n=2$ 时命题为假. 因此后面的递推步骤毫无意义. ☺

我们这里再给两个例子. 这两个例子仅需验证一种初始条件，但是递推步骤时，每个 $n=k$ 都需要用到各个 k 之前所有情形.

例 0.2.6　假设有一个 $a \times b$ 的巧克力排块. 那么需要掰几次才能把巧克力排块掰成 1×1 的小块呢？下面证明，无论怎样掰，都需要 $ab-1$ 次才能做到（参见图 0.2.1）.

当 $a=b=1$ 时，命题显然成立.

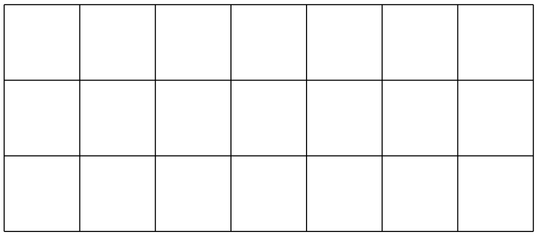


图 0.2.1 掰巧克力

假设 a, b 中至少有一个大于 1，不妨设 $a > 1$. 如果将 a 行排块掰成 k 行和 $a-k$ 行，那么还剩 $k \times b$ 的排块和 $(a-k) \times b$ 的排块需要掰. 根据归纳假设，还需要 $(kb-1)+((a-k)b-1)$ 次才能掰成 1×1 的小块. 于是共需要掰 $1+(kb-1)+((a-k)b-1) = ab-1$ 次. ☺

例 0.2.7 令 p_n 为自然数中的第 n 个素数. 我们来证明，第 n 个素数 p_n 一定小于等于 $2^{2^{n-1}}$. ☺

证 当 $n=1$ 时，命题显然成立.

假设当 $n \leqslant k$ 时命题成立，下面考虑 p_{k+1}. 由例 0.2.2 可得，$p_{k+1} \leqslant p_1 p_2 \cdots p_k + 1$. 根据归纳假设，我们有

$$p_{k+1} \leqslant p_1 p_2 \cdots p_k + 1 \leqslant 2^{2^0} 2^{2^1} \cdots 2^{2^{k-1}} + 1 \leqslant 2^{2^0+2^1+\cdots+2^{k-1}} + 1 \leqslant 2^{2^k-1} + 1 \leqslant 2^{2^k}.$$

因此命题成立. □

习题

练习 0.2.1 请用反证法证明以下命题.

1. 若 $p > 2$ 是素数，则 p 是奇数.
2. 对正整数 n，若 n^2 是奇数，则 n 是奇数.
3. 不存在最小的正有理数.

练习 0.2.2 经典逻辑的基本公理之一是排中律，即对任意命题 A，A 不能既不真又不假. 反证法可以看作排中律的推论，即如果我们发现 A 不能是错的，那么 A 就只能是对的. 因此，想要证明命题 A 真，不妨挑一个命题 B，先证明 $B \Rightarrow A$，再证明 $(\neg B) \Rightarrow A$. 那么 A 就必须是真的了. 给定命题：存在两个无理数 a, b，满足 a^b 是一个有理数. 请先假设 $\sqrt{2}^{\sqrt{2}}$ 是有理数，再假设 $\sqrt{2}^{\sqrt{2}}$ 是无理数，针对两种情形讨论来证明这个命题.

练习 0.2.3 证明 $2^{2^{n-1}}-1$ 必然至少有互异的 $n-1$ 个奇素因数.（因此第 n 个素数 p_n 一定小于等于 $2^{2^{n-1}}$.）

提示：运用因式分解 $x^2-1=(x+1)(x-1)$；并注意对任意 $m \neq n, 2^{2^m}+1$ 与 $2^{2^n}+1$ 互素.

练习 0.2.4 假设有一个 $a \times b$ 的巧克力排块，我们需要将它掰成 1×1 的小块，在掰的同时会被打分. 假设掰一次将 $x+y$ 块掰成了 x 块和 y 块，则得分 xy. 证明：当巧克力被完全掰成 1×1 的小块时，总得分永远为 $\dfrac{1}{2}ab(ab-1)$.

0.3 映　　射

最后回顾关于映射的基本概念.

设有集合 X 和 Y, 如果 X 中的任意元素 x, 都以某种法则 f 对应于 Y 中唯一一个元素, 则称这个对应的法则 f 是集合 X 到集合 Y 的一个**映射**, X 中的元素 x 所对应的 Y 中的元素常记作 $f(x)$.

映射用如下记号表示:

$$\begin{aligned} f\colon\ X &\to Y, \\ x &\mapsto f(x). \end{aligned}$$

这里, 集合 X 称为映射 f 的**定义域**, 集合 Y 称为映射 f 的**陪域**. 设 $x \in X, y = f(x) \in Y$, 我们称 y 为 x 在映射 f 下的**像**, x 为 y 在 f 下的**原像**. 注意, Y 中的元素不一定都有原像; 有原像时, 原像也可能不唯一. 定义域 X 中所有元素在 f 下的像构成 Y 的一个子集 $\{f(x) \mid x \in X\}$, 称为映射 f 的**像集**或**值域**, 记作 $f(X)$.

两个映射 $f, g\colon X \to Y$ 相等, 当且仅当对任意 $x \in X$, $f(x) = g(x)$.

根据映射的像集和原像的性质, 我们考虑如下三种特殊情形:

1. 如果对 X 中的任意两个不同元素 x_1, x_2, 都有 $f(x_1) \neq f(x_2)$, 则称 f 是一个**单射**; 即, 任何元素的原像至多只有一个.
2. 如果对 Y 中的任意元素 y, 都存在 X 中的某个 x, 使得 $y = f(x)$, 则称 f 是一个**满射**; 即, f 的像集 $f(X) = Y$.
3. 映射 f 既是单射又是满射, 则称 f 是一个**双射**. 双射 f 给出了集合 X 和 Y 之间的一个一一对应.

映射之间最重要的运算是两个映射的复合. 设 X, Y, Z 是三个集合, $f\colon X \to Y$, $g\colon Y \to Z$ 是两个映射, 则可以定义一个映射

$$\begin{aligned} h\colon\ X &\to Z, \\ x &\mapsto h(x) = g(f(x)), \end{aligned}$$

称其为 g 与 f 的**复合**, 记作 $h = g \circ f$. 可以如下表示:

$$\begin{array}{ccccc} X & \xrightarrow{f} & Y & \xrightarrow{g} & Z, \\ x & \mapsto & f(x) & \mapsto & g(f(x)). \end{array}$$

注意, 只有当 f 的陪域与 g 的定义域相同时, 才能定义二者的复合. 设 X, Y, Z, W 是四个集合, $f\colon X \to Y, g\colon Y \to Z, h\colon Z \to W$ 是三个映射, 则两种不同顺序的复合是同一个 X 到 W 的映射: $h \circ (g \circ f) = (h \circ g) \circ f$, 即映射的复合运算满足结合律. 因此, 这个映射可以写成 $h \circ g \circ f$ 而不会引起混淆.

当映射 f 的定义域和陪域都是 X 时, 称 f 是集合 X 上的一个**变换**. 容易看到, 一个变换永远可以跟自身复合, 记 $f^2 = f \circ f, f^n = f^{n-1} \circ f, n \geqslant 1$. 任意集合 X 上都有

一个特殊的变换，它把 X 中的任意元素 x 映到 x 本身，称为 X 上的**恒同变换**，记作 id_X，即 $\mathrm{id}_X(x)=x$，对任意 $x\in X$ 都成立. 特别地，恒同变换与映射的复合不改变映射：对任意映射 $f\colon X\to Y$，都有 $f\circ\mathrm{id}_X=f=\mathrm{id}_Y\circ f$.

定义 0.3.1 设 $f\colon X\to Y$ 是一个映射，若存在另一个映射 $g\colon Y\to X$，满足 $f\circ g=\mathrm{id}_Y, g\circ f=\mathrm{id}_X$，则称映射 f **可逆**，称 g 是 f 的一个**逆（映射）**.

容易看出，当 f 可逆时，它的逆映射是唯一的，称为 f 的逆，记作 f^{-1}.

利用映射复合和恒同变换，可以给出单射、满射和双射的等价描述.

命题 0.3.2 对映射 $f\colon X\to Y$，若 $X\neq\varnothing$，则有

1. f 是单射当且仅当存在另一个映射 $g\colon Y\to X$，使得 $g\circ f=\mathrm{id}_X$；
2. f 是满射当且仅当存在另一个映射 $g\colon Y\to X$，使得 $f\circ g=\mathrm{id}_Y$；
3. f 是双射当且仅当 f 可逆，即存在另一个映射 $g\colon Y\to X$，使得 $f\circ g=\mathrm{id}_Y$，$g\circ f=\mathrm{id}_X$.

命题的证明留作习题.

习题

练习 0.3.1 判断下列映射是否为单射、满射、双射，并写出双射的逆映射：

1. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto x+1$.
2. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 2x$.
3. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 3$.
4. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto x^2$.
5. $f\colon (-\infty,0]\to\mathbb{R}$, $x\mapsto x^2$.
6. $f\colon \mathbb{R}\to(0,\infty)$, $x\mapsto \mathrm{e}^x$.
7. $f\colon \left[-\dfrac{3\pi}{2},-\dfrac{\pi}{2}\right]\to[-1,1]$, $x\mapsto\sin x$.

练习 0.3.2 对复合映射 $f=g\circ h$，证明或举出反例.

1. g,h 都是满射，则 f 是满射.
2. g,h 都是单射，则 f 是单射.
3. h 不是满射，则 f 不是满射.
4. g 不是满射，则 f 不是满射.
5. g 不是单射，则 f 不是单射.
6. h 不是单射，则 f 不是单射.
7. g,h 都不是双射，则 f 不是双射.

练习 0.3.3

1. 在不改变对应法则和定义域的前提下，$\mathbb{R}$ 上的变换 $f(x)=x^2+2x+3$，把陪域换成哪个集合，得到的映射是满射?
2. 证明映射 $f\colon[0,1]\to\mathbb{R}, f(x)=x$ 是单射.

练习 0.3.4 下列 $\mathbb{R}$ 上的变换，哪些满足交换律 $f\circ g=g\circ f$?

1. $f(x)=x+1, g(x)=2x$.
2. $f(x)=x^2, g(x)=x^3$.
3. $f(x)=2^x, g(x)=3^x$.
4. $f(x)=2x+1, g(x)=3x+2$.
5. $f(x)=2x+1, g(x)=3x+1$.
6. $f(x)=\sin x, g(x)=\cos x$.

练习 0.3.5 对 $\mathbb{R}$ 上的变换 $f(x)=2x+1$ 和 $g(x)=ax+b$，求实数 a,b，使得 $f\circ g=g\circ f$.

练习 0.3.6 在化简函数 $\arccos\circ\sin$：$\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\to[0,\pi]$ 的下列步骤中，分别利用了映射的哪些性质？

1. 如果 $y=(\arccos\circ\sin)(x)$，则 $\cos y=(\cos\circ(\arccos\circ\sin))(x)$.
2. $\cos\circ(\arccos\circ\sin)=(\cos\circ\arccos)\circ\sin$.
3. $(\cos\circ\arccos)\circ\sin=\mathrm{id}_{[0,1]}\circ\sin$.
4. $\mathrm{id}_{[0,1]}\circ\sin=\sin$.

综上得 $\sin x=\cos y$，由此推断化简结果.

练习 0.3.7 用数学归纳法证明，任取有限个映射 $f_1,f_2,\cdots,f_n$，如果对 $1\leqslant i\leqslant n-1$，$f_{i+1}$ 的定义域都等于 f_i 的陪域，则复合映射 $f_n\circ\cdots\circ f_2\circ f_1$ 不因映射复合的计算次序而改变.

练习 0.3.8 证明，任意映射 f 都存在分解 $f=g\circ h$，其中 h 是满射，g 是单射.

练习 0.3.9 给定映射 h,g 和 $f=g\circ h$，证明，若 f 是双射，则 h 是单射，g 是满射.

练习 0.3.10 证明命题 0.3.2.

第 1 章 线性映射和矩阵

线性代数这门学科，以研究**线性系统**为目的，是一个从理论到应用都相当完善的数学分支．我们先从系统说起．系统是指由若干个互相作用和互相依赖的事物组合而成的具有特定功能的一个整体．比较简单的一种是集合之间的映射：

$$\text{输入} \xrightarrow{\text{系统}} \text{输出}.$$

读者在中学学习的函数是一种特殊的映射，即从实数集到实数集的映射．在本书中，我们用 $\mathbb{R}$ 来表示实数集，而 $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{C}$ 分别表示自然数集、整数集、有理数集和复集．

1.1 基本概念

我们从一个简单的系统开始．

例 1.1.1 (桥墩载荷) 如图 1.1.1 所示，有一座桥，桥长为 d，假设桥两端各有一桥墩，不考虑桥面和桥墩的重力、厚度．在距左侧 S_1 桥墩 l_1 处放置一重物，重力为 F_1．问两个桥墩各承担了多少载荷?

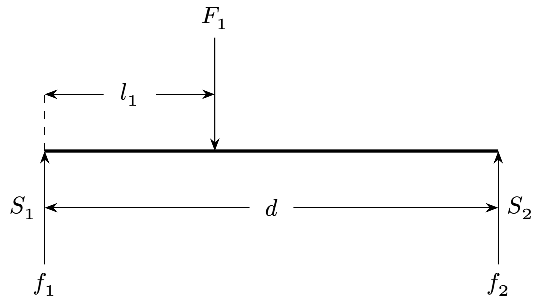


图 1.1.1 桥墩载荷：单重物

只需杠杆原理我们就可以求出答案．考虑以 S_1 为支点、桥面为杆臂的杠杆，显然 $F_1 l_1 = f_2 d$，因此 $f_2 = \dfrac{l_1}{d} F_1$．再考虑以 S_2 为支点的杠杆，有 $F_1(d - l_1) = f_1 d$，因此 $f_1 = \dfrac{d - l_1}{d} F_1$．也可以利用力平衡，$f_1 = F_1 - f_2$ 求出 f_1．

下面考虑两个重物的情形．如图 1.1.2 所示，在距左侧 S_1 桥墩 l_1 处放置一重物，重力为 F_1；在距左侧 S_1 桥墩 l_2 处放置一重物，重力为 F_2．问两个桥墩各承担了多少载荷?

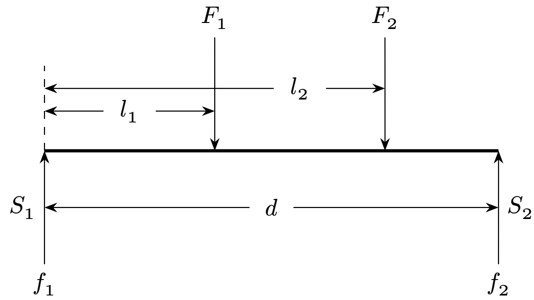


图 1.1.2　桥墩载荷：双重物

利用杠杆原理，考虑以 S_1 为支点、桥面为杆臂的杠杆，显然 $F_1l_1+F_2l_2=f_2d$，因此 $f_2=\frac{l_1}{d}F_1+\frac{l_2}{d}F_2$. 考虑以 S_2 为支点的杠杆，有 $F_1(d-l_1)+F_2(d-l_2)=f_1d$，因此 $f_1=\frac{d-l_1}{d}F_1+\frac{d-l_2}{d}F_2$. 也可以利用力平衡，$f_1=F_1+F_2-f_2$ 求出 f_1.

固定重物的位置，把重力 F_1,F_2 视为系统的输入，两个桥墩的载荷 f_1,f_2 视为系统的输出. 系统中的输入和输出满足**叠加原理**，即，

1. 输入放大或缩小某一倍数时，输出也放大或缩小同一倍数；
2. 两组输入所产生的输出是二者分别产生的独立输出之和.

满足叠加原理的系统称为**线性系统**.

更一般地，在任意多位置上放置任意重物，我们都可以直接写出两个桥墩的载荷的表达式. ☺

上面描述的线性系统里，输入和输出只有有限多个，二者都可以表示成有序数组，我们把有序数组称为向量.

定义 1.1.2 (向量)　一个 m 元有序数组 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}$ 称为一个 m 维**向量**，实数 $a_1,a_2,\cdots,a_m$ 称为向量 $\boldsymbol{a}$ 的**分量**或**坐标**.

分量都是实数的 m 维向量的全体构成的集合记为 $\mathbb{R}^m$.

注 1.1.3　分量都是复数的 m 维向量的全体构成的集合记为 $\mathbb{C}^m$. 分量也可以取其他范围内的数，本书主要讨论分量为实数和复数的两种情形.

两个向量**相等**是指二者的每个分量都相等. 向量常用黑体小写字母表示[①]，如 $\boldsymbol{a}$. 由于分量纵向排列，向量又称为列向量. 根据需要，向量也可以把分量横向排列，称为行向量. 为

① 手写时可以直接写字母，也可以在字母上加箭头，如 $\vec{a}$.

了与列向量区分,用符号 $\boldsymbol{a}^{\mathrm{T}}$ 表示行向量,即如果 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}$,则 $\boldsymbol{a}^{\mathrm{T}}=\begin{bmatrix}a_1 & a_2 & \cdots & a_m\end{bmatrix}$.

现在,例 1.1.1 中两个重物的线性系统可以表示成如下映射:

$$\begin{aligned}\mathbb{R}^2 &\to \mathbb{R}^2,\\ \begin{bmatrix}F_1\\F_2\end{bmatrix} &\mapsto \begin{bmatrix}f_1\\f_2\end{bmatrix}=\begin{bmatrix}\dfrac{d-l_1}{d}F_1+\dfrac{d-l_2}{d}F_2\\ \dfrac{l_1}{d}F_1+\dfrac{l_2}{d}F_2\end{bmatrix}.\end{aligned}$$

一般地,n 个输入 m 个输出的线性系统可以表示成如下映射:

$$\begin{aligned}f:\quad \mathbb{R}^n &\to \mathbb{R}^m,\\ \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} &\mapsto \boldsymbol{y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_m\end{bmatrix}=\begin{bmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n\end{bmatrix}.\end{aligned}\tag{1.1.1}$$

我们将在下节说明该表示的正确性.

线性系统中的输入和输出,也就是向量,需要满足叠加原理:可加性要求任意两个向量之间可以相加,而齐次性要求任意一个向量可以放缩某一倍数.

定义 1.1.4 (线性运算) 为 $\mathbb{R}^m$ 中的向量定义两种运算[①]

1. 两个 m 维向量的**向量加法**: $\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}+\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}:=\begin{bmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_m+b_m\end{bmatrix}$;

2. 一个 m 维向量与一个实数的**数乘**: $k\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}:=\begin{bmatrix}ka_1\\ka_2\\\vdots\\ka_m\end{bmatrix}$;

向量的加法和数乘统称向量的**线性运算**.

带有线性运算的集合 $\mathbb{R}^m$,称为**向量空间** $\mathbb{R}^m$ 或**线性空间** $\mathbb{R}^m$.

注意,“线性空间 $\mathbb{R}^m$”中的向量能做加法和数乘,而“集合 $\mathbb{R}^m$”则不能,二者并不是同一个数学对象. 另外,线性空间 $\mathbb{R}^n$ 和线性空间 $\mathbb{R}^m$ 中的向量在 $m\neq n$ 时无法做加法;线性空间 $\mathbb{R}^m$ 上的数乘运算中的数,需要是实数.

① 表达式中,$A:=B$ 表示令 $A=B$ 或者把 B 记作 A;类似地,$A=:B$ 表示 $B:=A$.

线性运算本身需要满足一定的运算法则.

命题 1.1.5 线性空间 $\mathbb{R}^m$ 中的向量加法和数乘满足如下八条运算法则:

1. 加法结合律: $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;
2. 加法交换律: $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;
3. 零向量: 存在 m 维**零向量** $\mathbf{0}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$, 满足 $\boldsymbol{a}+\mathbf{0}=\boldsymbol{a}$;
4. 负向量: 对任意 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_m\end{bmatrix}$, 记 $-\boldsymbol{a}=\begin{bmatrix}-a_1\\-a_2\\\vdots\\-a_m\end{bmatrix}$, 则 $\boldsymbol{a}+(-\boldsymbol{a})=\mathbf{0}$, 称 $-\boldsymbol{a}$ 为 $\boldsymbol{a}$ 的**负向量**;
5. 单位数: $1\boldsymbol{a}=\boldsymbol{a}$;
6. 数乘结合律: $(kl)\boldsymbol{a}=k(l\boldsymbol{a})$;
7. 数乘关于数的分配律: $(k+l)\boldsymbol{a}=k\boldsymbol{a}+l\boldsymbol{a}$;
8. 数乘关于向量的分配律: $k(\boldsymbol{a}+\boldsymbol{b})=k\boldsymbol{a}+k\boldsymbol{b}$.

其中, $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 是 $\mathbb{R}^m$ 中的向量, k,l 是实数.

命题的验证是直接的, 我们留给读者. 下面讨论运算法则的一些结果和注意事项.

1. 根据负向量性质, 我们可以定义向量的**减法**: $\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b})$.
2. 根据加法结合律, 我们不需要区分加法运算的先后次序, 因此可以写 $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$. 类似地, $kl\boldsymbol{a}=(kl)\boldsymbol{a}=k(l\boldsymbol{a})$.
3. 上面的运算法则使得我们能够"合并同类项". 例如
$$2\boldsymbol{a}+4(2\boldsymbol{a}+(3\boldsymbol{b}+\mathbf{0}))+4((\boldsymbol{a}+(2+3)\boldsymbol{c})-\boldsymbol{a})=10\boldsymbol{a}+12\boldsymbol{b}+20\boldsymbol{c}.$$
4. 有关数乘的其他结果: $0\boldsymbol{a}=\mathbf{0}$, $(-1)\boldsymbol{a}=-\boldsymbol{a}$.
5. 特别注意, 向量之间**没有**直接的乘除法. 有时我们会写 $\dfrac{\boldsymbol{a}}{k}$, 这时 k 必须是实数而非向量. 向量出现在分母上, 意味着已经犯了十分严重的概念错误.

例 1.1.6 读者在中学学习过平面向量. 把平面向量的起点固定在原点, 它的坐标表示就是二维数组; 而平面向量线性运算的坐标表示就是上面定义的两种线性运算. 因此, 配备上线性运算, 所有以原点为起点的平面向量全体构成的集合就是线性空间 $\mathbb{R}^2$. ☺

有了线性空间中的线性运算，线性系统 (1.1.1) 的叠加原理可以具体地描述如下：

$$f(\boldsymbol{x}+\boldsymbol{x}')=\begin{bmatrix}a_{11}(x_1+x_1')+\cdots+a_{1n}(x_n+x_n')\\a_{21}(x_1+x_1')+\cdots+a_{2n}(x_n+x_n')\\\vdots\\a_{m1}(x_1+x_1')+\cdots+a_{mn}(x_n+x_n')\end{bmatrix}$$

$$=\begin{bmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\a_{21}x_1+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+\cdots+a_{mn}x_n\end{bmatrix}+\begin{bmatrix}a_{11}x_1'+\cdots+a_{1n}x_n'\\a_{21}x_1'+\cdots+a_{2n}x_n'\\\vdots\\a_{m1}x_1'+\cdots+a_{mn}x_n'\end{bmatrix}=f(\boldsymbol{x})+f(\boldsymbol{x}'),$$

$$f(k\boldsymbol{x})=\begin{bmatrix}a_{11}kx_1+\cdots+a_{1n}kx_n\\a_{21}kx_1+\cdots+a_{2n}kx_n\\\vdots\\a_{m1}kx_1+\cdots+a_{mn}kx_n\end{bmatrix}=k\begin{bmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\a_{21}x_1+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+\cdots+a_{mn}x_n\end{bmatrix}=kf(\boldsymbol{x}),$$

其中[①]，$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\boldsymbol{x}'=\begin{bmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{bmatrix}\in\mathbb{R}^n,k\in\mathbb{R}$. 由观察可知，等式左边的 $\boldsymbol{x}+\boldsymbol{x}',k\boldsymbol{x}$ 是 $\mathbb{R}^n$ 中的线性运算，而等式右边的 $f(\boldsymbol{x})+f(\boldsymbol{x}'),kf(\boldsymbol{x})$ 是 $\mathbb{R}^m$ 中的线性运算.

定义 1.1.7 (线性映射)　映射 $f\colon\mathbb{R}^n\to\mathbb{R}^m$，如果满足

1. 任取 $\boldsymbol{x},\boldsymbol{x}'\in\mathbb{R}^n$，都有 $f(\boldsymbol{x}+\boldsymbol{x}')=f(\boldsymbol{x})+f(\boldsymbol{x}')$；
2. 任取 $\boldsymbol{x}\in\mathbb{R}^n,k\in\mathbb{R}$，都有 $f(k\boldsymbol{x})=kf(\boldsymbol{x})$；

则称 f 为从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的**线性映射**.

因此，(1.1.1) 式中的映射是线性映射.

线性映射把定义域上的加法映射成陪域上的加法，把定义域上的数乘映射成陪域上同一个数的数乘.简单来说,线性映射保持线性运算.例如,对任意线性映射 $f\colon\mathbb{R}^n\to\mathbb{R}^m$，都有 $f(k\boldsymbol{x}+k'\boldsymbol{x}'+k''\boldsymbol{x}'')=kf(\boldsymbol{x})+k'f(\boldsymbol{x}')+k''f(\boldsymbol{x}'')$. 特别地，线性映射把零向量映射到零向量：在定义中取常数 $k=0$，可以得到 $f(\boldsymbol{0}_n)=\boldsymbol{0}_m$，其中 $\boldsymbol{0}_n\in\mathbb{R}^n$ 是 $\mathbb{R}^n$ 中的零向量，而 $\boldsymbol{0}_m\in\mathbb{R}^m$ 是 $\mathbb{R}^m$ 中的零向量. 利用这个性质，可以对一些映射简单有效地做出其并非线性的判断.

例 1.1.8　线性映射在生活中随处可见.

① 注意，这里 $\boldsymbol{x}'$ 表示不同于 $\boldsymbol{x}$ 的另一个向量，而不是对向量 $\boldsymbol{x}$ 做了某种运算的结果. 类似地，后文出现的 $\widetilde{\boldsymbol{x}},\widehat{\boldsymbol{x}}$ 等都表示不同的向量.

1. 去超市买水果, 在苹果、香蕉、樱桃之间选择. 潜在的购买空间是由这三种水果组成的一个线性空间 $\mathbb{R}^3$, 设苹果 $\boldsymbol{a}=\begin{bmatrix}1\\0\\0\end{bmatrix}$, 香蕉 $\boldsymbol{b}=\begin{bmatrix}0\\1\\0\end{bmatrix}$, 和樱桃 $\boldsymbol{c}=\begin{bmatrix}0\\0\\1\end{bmatrix}$. 而结账这个过程就是一个映射, 它将水果组合, 映射到对应的总价上. 如果苹果、香蕉和樱桃的每千克价格分别是 12.98 元、9.98 元和 129.98 元, 那么“结账”这个映射可以表达为:

$$\begin{aligned}\text{结账:}\quad \mathbb{R}^3 &\to \mathbb{R},\\ \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} &\mapsto 12.98x_1+9.98x_2+129.98x_3.\end{aligned}$$

容易看出, 这是一个线性映射. 例如, $\boldsymbol{x},\boldsymbol{x}'$ 分别代表两种水果组合, 对 $\boldsymbol{x}+2\boldsymbol{x}'$ 结账时, 总价等于对 $\boldsymbol{x}$ 结账的总价加上对 $\boldsymbol{x}'$ 结账的总价的 2 倍. 即,

$$\text{结账}(\boldsymbol{x}+2\boldsymbol{x}')=\text{结账}(\boldsymbol{x})+2\,\text{结账}(\boldsymbol{x}').$$

2. 假设每千克苹果含有 12 克纤维素和 135 克糖, 每千克香蕉含有 12 克纤维素和 208 克糖, 每千克樱桃含有 3 克纤维素和 99 克糖. 所有纤维素和糖的含量的组合组成一个二维线性空间 $\mathbb{R}^2$, 设纤维素 $\boldsymbol{f}=\begin{bmatrix}1\\0\end{bmatrix}$, 糖 $\boldsymbol{s}=\begin{bmatrix}0\\1\end{bmatrix}$. 把每种水果组合映射到对应的纤维素和糖的含量, 是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}^2$ 的映射, 称为“营养”:

$$\begin{aligned}\text{营养:}\quad \mathbb{R}^3 &\to \mathbb{R}^2,\\ \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} &\mapsto \begin{bmatrix}12x_1+12x_2+3x_3\\135x_1+208x_2+99x_3\end{bmatrix}.\end{aligned}$$

容易看出, 这是一个线性映射.

注意, 这两个例子中的线性映射都具有形如 (1.1.1) 式的表达式. ☺

定义 1.1.9 (线性变换) 从 $\mathbb{R}^n$ 到自身的线性映射称为 $\mathbb{R}^n$ 上的**线性变换**.

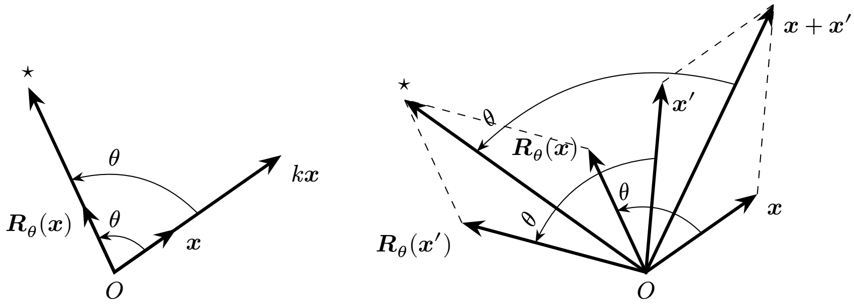
特别地, $\mathbb{R}^n$ 上的 **恒同变换**

$$\begin{aligned}\boldsymbol{I}=\mathrm{id}:\ \mathbb{R}^n &\to \mathbb{R}^n\\ \boldsymbol{x} &\mapsto \boldsymbol{x}\end{aligned} \tag{1.1.2}$$

是线性变换.

例 1.1.10 考虑例 1.1.6 中由平面向量构成的线性空间 $\mathbb{R}^2$，下面讨论平面①上的几类线性变换.

1. **旋转变换**: 对任意实数 θ, 将所有向量绕原点逆时针旋转 θ 大小的角. 这定义了一个 $\mathbb{R}^2$ 上的变换, 记为 $\boldsymbol{R}_\theta$. 图 1.1.3 证明了 $\boldsymbol{R}_\theta(k\boldsymbol{x}) = k\boldsymbol{R}_\theta(\boldsymbol{x})$, $\boldsymbol{R}_\theta(\boldsymbol{x}+\boldsymbol{x}') = \boldsymbol{R}_\theta(\boldsymbol{x}) + \boldsymbol{R}_\theta(\boldsymbol{x}')$.



(a) 数乘 $\boldsymbol{R}_\theta(k\boldsymbol{x}) = k\boldsymbol{R}_\theta(\boldsymbol{x})$ (b) 加法 $\boldsymbol{R}_\theta(\boldsymbol{x}+\boldsymbol{x}') = \boldsymbol{R}_\theta(\boldsymbol{x}) + \boldsymbol{R}_\theta(\boldsymbol{x}')$

图 1.1.3 旋转变换

因此, 旋转变换 $\boldsymbol{R}_\theta$ 是线性变换. 那么, 它是否具有形如 (1.1.1) 式的表达式呢? 由于 $\boldsymbol{R}_\theta$ 是线性的，我们有

$$\boldsymbol{R}_\theta\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \boldsymbol{R}_\theta\left(x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1\boldsymbol{R}_\theta\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + x_2\boldsymbol{R}_\theta\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right).$$

容易知道

$$\boldsymbol{R}_\theta\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad \boldsymbol{R}_\theta\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{bmatrix},$$

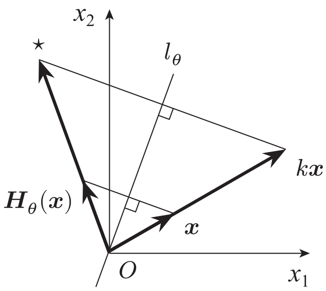
因此

$$\boldsymbol{R}_\theta\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + x_2\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{bmatrix}.$$

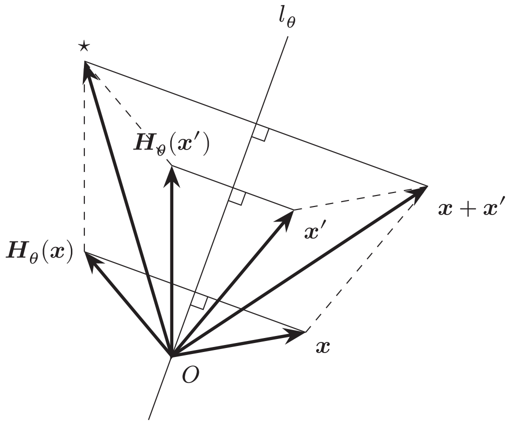
注意，线性映射 $\boldsymbol{R}_\theta$ 仅仅由它在 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 这两个向量上的取值决定．这是线性映射最特殊的地方．我们称 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 这两个向量为 $\mathbb{R}^2$ 的**标准坐标向量**.

① 我们以后经常会直接把 $\mathbb{R}^2$ 称为平面，但要注意多数情形中 $\mathbb{R}^2$ 中的元素是向量，而不是点.

2. **反射变换**：给定过原点的直线 l_θ：$x_2\cos\theta - x_1\sin\theta = 0$，其中 θ 是直线与 x_1 坐标轴的夹角．任意向量关于 l_θ 的反射，定义了一个 $\mathbb{R}^2$ 上的变换，记为 $\boldsymbol{H}_\theta$．图 1.1.4 展示了 $\boldsymbol{H}_\theta(k\boldsymbol{x}) = k\boldsymbol{H}_\theta(\boldsymbol{x})$，$\boldsymbol{H}_\theta(\boldsymbol{x}+\boldsymbol{x}') = \boldsymbol{H}_\theta(\boldsymbol{x}) + \boldsymbol{H}_\theta(\boldsymbol{x}')$．



(a) 数乘 $\boldsymbol{H}_\theta(k\boldsymbol{x}) = k\boldsymbol{H}_\theta(\boldsymbol{x})$



(b) 加法 $\boldsymbol{H}_\theta(\boldsymbol{x}+\boldsymbol{x}') = \boldsymbol{H}_\theta(\boldsymbol{x}) + \boldsymbol{H}_\theta(\boldsymbol{x}')$

图 1.1.4　反射变换

因此，反射变换 $\boldsymbol{H}_\theta$ 是线性变换．类似地，由于

$$\boldsymbol{H}_\theta\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\cos 2\theta\\ \sin 2\theta\end{bmatrix},\quad \boldsymbol{H}_\theta\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}\cos\left(2\theta - \dfrac{\pi}{2}\right)\\ \sin\left(2\theta - \dfrac{\pi}{2}\right)\end{bmatrix},$$

因此

$$\boldsymbol{H}_\theta\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = x_1\begin{bmatrix}\cos 2\theta\\ \sin 2\theta\end{bmatrix} + x_2\begin{bmatrix}\sin 2\theta\\ -\cos 2\theta\end{bmatrix} = \begin{bmatrix}x_1\cos 2\theta + x_2\sin 2\theta\\ x_1\sin 2\theta - x_2\cos 2\theta\end{bmatrix}.$$

特别地，如果直线就是 x_1 坐标轴，即 $\theta = 0$，则 $\boldsymbol{H}_\theta\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}x_1\\-x_2\end{bmatrix}$.

3. **对换变换**：对换 $\mathbb{R}^2$ 中向量的两个分量 x_1, x_2 也构成一个线性变换：

$$\begin{aligned}\boldsymbol{P}:\quad \mathbb{R}^2 &\to \mathbb{R}^2\\ \begin{bmatrix}x_1\\x_2\end{bmatrix} &\mapsto \begin{bmatrix}x_2\\x_1\end{bmatrix}.\end{aligned}$$

可以看到，其实 $\boldsymbol{P}$ 也是关于直线 $x_2 - x_1 = 0$ 的反射.

4. **伸缩变换**：设 $k\in\mathbb{R}$，定义一个 $\mathbb{R}^2$ 上的变换 $\boldsymbol{C}_k$，它把向量在 x_1 方向拉伸 k

倍，x_2 方向保持不变，其表达式为

$$\begin{aligned} \boldsymbol{C}_k\colon\quad \mathbb{R}^2 &\to \mathbb{R}^2 \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\mapsto \begin{bmatrix} kx_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

容易验证 $\boldsymbol{C}_k$ 是线性变换. 若 $k>0$，它把单位正方形 $\{0\leqslant x_1\leqslant 1, 0\leqslant x_2\leqslant 1\}$ 映射成矩形 $\{0\leqslant x_1\leqslant k, 0\leqslant x_2\leqslant 1\}$.

5. **投影变换**：当 $k=0$ 时，伸缩变换 $\boldsymbol{C}_0$ 是对 x_2 轴的垂直投影.
6. **错切变换**：设 $k\in\mathbb{R}$，定义 $\mathbb{R}^2$ 上的一个变换 $\boldsymbol{S}_k$，它把 x_1 方向的 k 倍加到 x_2 方向上，并保持 x_1 方向不变，其表达式为

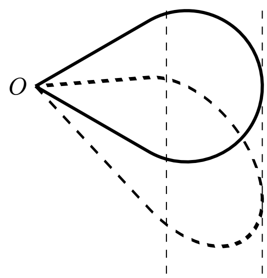


图 1.1.5　错切变换

$$\begin{aligned} \boldsymbol{S}_k\colon\quad \mathbb{R}^2 &\to \mathbb{R}^2 \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\mapsto \begin{bmatrix} x_1 \\ kx_1+x_2 \end{bmatrix}. \end{aligned}$$

容易验证 $\boldsymbol{S}_k$ 是线性变换. 它把单位正方形 $\{0\leqslant x_1\leqslant 1, 0\leqslant x_2\leqslant 1\}$ 映射成平行四边形 $\{0\leqslant x_1\leqslant 1, kx_1\leqslant x_2\leqslant kx_1+1\}$，而面积保持不变（参见图 1.1.5).　☺

例 1.1.11　容易验证，以下这几个映射都**不是**线性映射.

1. 平面上的平移变换：设 $\boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ 是平面向量，$\boldsymbol{T_a}\colon\mathbb{R}^2\to\mathbb{R}^2$ 是关于向量 $\boldsymbol{a}$ 的平移，其表达式为

$$\begin{aligned} \boldsymbol{T_a}\colon\quad \mathbb{R}^2 &\to \mathbb{R}^2 \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\mapsto \begin{bmatrix} x_1+a_1 \\ x_2+a_2 \end{bmatrix}. \end{aligned}$$

当 $\boldsymbol{a}\neq\boldsymbol{0}$ 时，$\boldsymbol{T_a}(\boldsymbol{0})=\boldsymbol{0}+\boldsymbol{a}=\boldsymbol{a}$，即 $\boldsymbol{T_a}$ 不保持零向量，因此 $\boldsymbol{T_a}$ **不是**线性映射.

类似地，如果直线 l 不经过原点，那么关于 l 的反射变换就不是线性映射.

2. 取长度：定义映射 $l\colon\mathbb{R}^2\to\mathbb{R}$，它将平面上的任意向量 $\boldsymbol{x}$ 映射到这个向量的长度. 其表达式为

$$\begin{aligned} l\colon\quad \mathbb{R}^2 &\to \mathbb{R} \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\mapsto \sqrt{x_1^2+x_2^2}. \end{aligned}$$

3. 绕圆旋转：定义映射 $c\colon\mathbb{R}\to\mathbb{R}^2$，它将输入的角度映射到单位圆上对应的位置.

如果将输入看作时间，那么这个映射就可以看作是一个点匀速绕定点旋转 (参见图 1.1.6). 这个映射的表达式为

$$\begin{array}{rccc} c\colon & \mathbb{R} & \to & \mathbb{R}^2 \\ & \theta & \mapsto & \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}. \end{array}$$

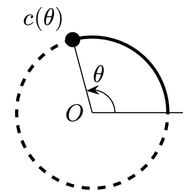


图 1.1.6 匀速圆周运动

4. 齐次非线性：定义映射 $f\colon \mathbb{R}^2 \to \mathbb{R}$. 其表达式为

$$\begin{array}{rccc} f\colon & \mathbb{R}^2 & \to & \mathbb{R} \\ & \begin{bmatrix}x_1 \\ x_2\end{bmatrix} & \mapsto & (x_1^3+x_2^3)^{\frac{1}{3}}. \end{array}$$

注意，这个映射虽然满足 $f(k\boldsymbol{a}) = kf(\boldsymbol{a})$，但是 $f(\boldsymbol{a}+\boldsymbol{b}) \neq f(\boldsymbol{a})+f(\boldsymbol{b})$. ☺

映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$, 单独看 $\mathbb{R}^m$ 中任意一个分量的表达式, 都是自变量 $x_1, x_2, \cdots, x_n$ 的多元函数. 可以观察到，例 1.1.10 中的线性映射，每一个分量都是自变量的**线性函数**，即常数项为零的一次函数，形如 $a_1x_1 + a_2x_2 + \cdots + a_nx_n$；而例 1.1.11 中的反例，映射的分量并不是线性函数. 那么，线性映射的分量是不是总是线性函数呢?

习题

练习 1.1.1 如图 1.1.7 所示，钟表表盘上对应整点有 12 个向量，其中 12 点对应向量为 $\begin{bmatrix}0\\1\end{bmatrix}$.

1. 计算 12 个向量之和.
2. 不计 2 点方向向量，计算其他 11 个向量之和.
3. 假设这 12 个向量的起点从表盘中心移到 6 点，则 12 点对应向量变为 $\begin{bmatrix}0\\2\end{bmatrix}$. 计算此时 12 个向量之和.

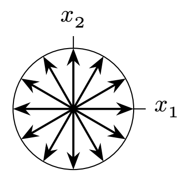


图 1.1.7

练习 1.1.2 如果平面上的向量 $\begin{bmatrix}a\\b\end{bmatrix}$ 与 $\begin{bmatrix}c\\d\end{bmatrix}$ 共线，那么 $\begin{bmatrix}a\\c\end{bmatrix}$ 与 $\begin{bmatrix}b\\d\end{bmatrix}$ 是否共线?

练习 1.1.3 证明命题 1.1.5.

练习 1.1.4

1. 如果只用加法交换律，不用加法结合律，那么 $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$ 有多少种与之相等的表达式?
2. 如果只用加法结合律，不用加法交换律，那么 $((\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c})+\boldsymbol{d}$ 有多少种与之相等的表达式?

练习 1.1.5 判断下列映射是否为线性映射：

1. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto x+1$.
2. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 2x$.
3. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 0$.
4. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 1$.
5. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto x^2$.
6. $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto 2^x$.
7. $f\colon \mathbb{R}^2\to\mathbb{R}^3$, $\begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}x+y\\y-x\\2x\end{bmatrix}$.
8. $f\colon \mathbb{R}^2\to\mathbb{R}^3$, $\begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}x+1\\y-x\\2x\end{bmatrix}$.

练习 1.1.6 设 $f\colon\mathbb{R}\to\mathbb{R}$ 是线性映射，证明存在实数 k，使得 $f(x)=kx$.

练习 1.1.7 设映射 $f\colon\mathbb{R}^3\to\mathbb{R}^2, f\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)=\begin{bmatrix}g(x,y,z)\\h(x,y,z)\end{bmatrix}$，其中 $g,h\colon\mathbb{R}^3\to\mathbb{R}$. 证明 f 是线性映射当且仅当 g,h 都是线性映射.

练习 1.1.8 设 $\boldsymbol{x}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{x}_2=\begin{bmatrix}1\\1\end{bmatrix},\boldsymbol{x}_3=\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}1\\1\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\0\\1\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}0\\1\\1\end{bmatrix}$，是否存在线性映射 $f\colon\mathbb{R}^2\to\mathbb{R}^3$，满足 $f(\boldsymbol{x}_i)=\boldsymbol{b}_i, i=1,2,3$?

练习 1.1.9 判断下列映射是否为线性映射:

1. 给定 $\boldsymbol{a}\in\mathbb{R}^m$，$f\colon\mathbb{R}\to\mathbb{R}^m, f(k)=k\boldsymbol{a}$.
2. 给定实数 k，$f\colon\mathbb{R}^m\to\mathbb{R}^m, f(\boldsymbol{a})=k\boldsymbol{a}$.
3. $f\colon\mathbb{R}^{m+1}\to\mathbb{R}^m, f\left(\begin{bmatrix}a_1\\\vdots\\a_m\\a_{m+1}\end{bmatrix}\right)=a_{m+1}\begin{bmatrix}a_1\\\vdots\\a_m\end{bmatrix}$.

练习 1.1.10 给定三维向量 $\boldsymbol{a}=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$，定义:

1. 二者**点积**为 $\boldsymbol{a}\cdot\boldsymbol{b}:=a_1b_1+a_2b_2+a_3b_3$.
2. 二者**叉积**为 $\boldsymbol{a}\times\boldsymbol{b}:=\begin{bmatrix}a_2b_3-a_3b_2\\a_3b_1-a_1b_3\\a_1b_2-a_2b_1\end{bmatrix}$.

那么给定 $\boldsymbol{a}\in\mathbb{R}^3$，映射 $f\colon\mathbb{R}^3\to\mathbb{R},\boldsymbol{b}\mapsto\boldsymbol{a}\cdot\boldsymbol{b}$ 和 $g\colon\mathbb{R}^3\to\mathbb{R}^3,\boldsymbol{b}\mapsto\boldsymbol{a}\times\boldsymbol{b}$ 是否为线性映射?

练习 1.1.11 给定平面上任意面积为 1 的三角形，经过下列变换之后，其面积是否确定？如果是，面积为多少？（不需严谨证明，猜测答案即可.）

1. 旋转变换.
2. 反射变换.
3. 对 x_2 投影的投影变换.
4. x_1 方向拉伸 3 倍，x_2 方向不变的伸缩变换.
5. 把 x_1 方向的 3 倍加到 x_2 方向上，保持 x_1 方向不变的错切变换.

练习 1.1.12 设 $\mathbb{R}^2$ 上的变换 $f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y+1 \\ x+2 \end{bmatrix}$.

1. 证明 f 不是线性变换.
2. 构造分解 $f = g \circ h$，其中 g 是 $\mathbb{R}^2$ 上的平移变换，h 是 $\mathbb{R}^2$ 上的线性变换.（这种平移与线性变换的复合称为仿射变换. 注意，平移变换并不是线性变换.）

练习 1.1.13 ☕ 设连续函数 $f\colon \mathbb{R} \to \mathbb{R}$ 满足 $f(a+b) = f(a) + f(b)$，f 是否为线性映射？（需要微积分知识）

1.2 线性映射的表示矩阵

本节主要讨论线性映射是不是总具有 (1.1.1) 式的形式.

我们称线性空间中的一组向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 为一个向量组. 注意，$\boldsymbol{a}_j$ 中 a 用黑体，表示它是一组向量中的第 j 个向量，而不是某个向量的第 j 个分量. 首先考虑定义域 $\mathbb{R}^n$ 中一组特殊的向量

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

这组向量称为 $\mathbb{R}^n$ 的**标准坐标向量组**，其中 $\boldsymbol{e}_i$ 称为第 i 个**标准坐标向量**. 这组向量的一个特殊之处在于，$\mathbb{R}^n$ 中的任意向量 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ 都可以由它们做线性运算得到：

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n.$$

对线性映射 f，由于线性映射保持线性运算，因此有

$$f(\boldsymbol{x}) = x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n). \tag{1.2.1}$$

定义 1.2.1 (线性组合与线性表示) 给定 $\mathbb{R}^m$ 中向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 和一组数 $k_1, k_2, \cdots, k_n \in \mathbb{R}$，称向量 $k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_n\boldsymbol{a}_n$ 是向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 的一个**线性组合**.

设 $\boldsymbol{b}$ 是 $\mathbb{R}^m$ 中的向量，如果存在一组数 $k_1, k_2, \cdots, k_n \in \mathbb{R}$，使得 $\boldsymbol{b} = k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_n\boldsymbol{a}_n$，则称 $\boldsymbol{b}$ 可以被向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ **线性表示**.

任意向量 $\boldsymbol{x} \in \mathbb{R}^n$ 都可以被标准坐标向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性表示，而 $f(\boldsymbol{x})$ 也以同样方式被它们在 f 下的像 $f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \cdots, f(\boldsymbol{e}_n)$ 线性表示. 可以说明，线性映射 f 由 n 个特殊的像 $f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \cdots, f(\boldsymbol{e}_n)$ 所决定.

命题 1.2.2 设 $f, g\colon \mathbb{R}^n \to \mathbb{R}^m$ 是两个线性映射，如果 $f(\boldsymbol{e}_i) = g(\boldsymbol{e}_i), i = 1, 2, \cdots, n$，则 $f = g$.

证 如果 $f(\boldsymbol{e}_i) = g(\boldsymbol{e}_i), i = 1, 2, \cdots, n$，则对任意 $\boldsymbol{x} \in \mathbb{R}^n$，都有

$$f(\boldsymbol{x}) = x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n) = x_1 g(\boldsymbol{e}_1) + x_2 g(\boldsymbol{e}_2) + \cdots + x_n g(\boldsymbol{e}_n) = g(\boldsymbol{x}).$$

这意味着两个映射相等. □

这个命题解决了线性映射的唯一性问题, 那么存在性问题呢? 即, 任取 $\mathbb{R}^m$ 中的 n 个向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$, 能否找到一个线性映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$, 满足 $f(\boldsymbol{e}_i) = \boldsymbol{a}_i, i = 1, 2, \cdots, n$? 答案是肯定的. 根据 (1.2.1) 式，首先定义映射

$$\begin{aligned} f\colon \quad \mathbb{R}^n \quad &\to \quad \mathbb{R}^m \\ \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad &\mapsto \quad x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \cdots + x_n \boldsymbol{a}_n. \end{aligned}$$

下面根据定义验证它是线性映射:

$$\begin{aligned} f(\boldsymbol{x} + \boldsymbol{x}') &= (x_1 + x_1')\boldsymbol{a}_1 + (x_2 + x_2')\boldsymbol{a}_2 + \cdots + (x_n + x_n')\boldsymbol{a}_n \\ &= (x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \cdots + x_n \boldsymbol{a}_n) + (x_1' \boldsymbol{a}_1 + x_2' \boldsymbol{a}_2 + \cdots + x_n' \boldsymbol{a}_n) \\ &= f(\boldsymbol{x}) + f(\boldsymbol{x}'), \end{aligned}$$

$$f(k\boldsymbol{x}) = kx_1 \boldsymbol{a}_1 + kx_2 \boldsymbol{a}_2 + \cdots + kx_n \boldsymbol{a}_n = k(x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \cdots + x_n \boldsymbol{a}_n) = kf(\boldsymbol{x}),$$

其中，$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{x}' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} \in \mathbb{R}^n, k \in \mathbb{R}$. 因此我们有如下结果，其中的唯一性由命题 1.2.2 保证.

命题 1.2.3 任取 $\mathbb{R}^m$ 中的 n 个向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$，都存在唯一的线性映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$，满足 $f(\boldsymbol{e}_i) = \boldsymbol{a}_i, i = 1, 2, \cdots, n$.

这个命题从理论上完整地描述了线性映射. 一方面, 尽管定义域和陪域都包含无穷多个元素，但保持线性运算这一性质使得线性映射仅由它在标准坐标向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 上的取值唯一决定；另一方面，这 n 个取值在陪域中又是完全自由的，可以任意选取.

下面从计算的角度，具体地描述线性映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$. 设 $f(\boldsymbol{e}_i) = \boldsymbol{a}_i, i = 1, 2, \cdots, n$, 其中

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \cdots, \boldsymbol{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

因为 $f(\boldsymbol{x}) = x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n)$，所以线性映射 f 恰好具有 (1.1.1) 式中的形式. 把向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 并列排在一起，得到一个矩形的数表：

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

数学上称之为**矩阵**. 我们用大写字母表示矩阵，如 A. 上面的矩阵中，每个元素带有两个下角标，第一个下角标代表元素所在的行，第二个下角标代表元素所在的列，如 a_{ij} 就表示该元素位于第 i 行第 j 列的交叉点处，称为矩阵 A 的 (i,j) 元[①]. 竖排的 n 个 $\mathbb{R}^m$ 中的向量

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \cdots, \boldsymbol{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

称为矩阵的**列向量**. 横排的 m 个 $\mathbb{R}^n$ 中的向量

$$\tilde{\boldsymbol{a}}_1^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}, \tilde{\boldsymbol{a}}_2^{\mathrm{T}} = \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix}, \cdots, \tilde{\boldsymbol{a}}_m^{\mathrm{T}} = \begin{bmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为矩阵的**行向量**. 有 m 行 n 列的矩阵简称为 $m \times n$ 矩阵. 矩阵既可以写成列向量组的横向排列 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$，又可以写成行向量组的纵向排列 $A = \begin{bmatrix} \tilde{\boldsymbol{a}}_1^{\mathrm{T}} \\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}} \\ \vdots \\ \tilde{\boldsymbol{a}}_m^{\mathrm{T}} \end{bmatrix}$. 而 $m \times n$ 矩阵 A 的 (i,j) 元如果是 a_{ij}，那么可以记作 $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{m \times n}$. 两个矩阵**相等**是指二者的每个元素都相等. 根据定义，n 维行向量也是 $1 \times n$ 矩阵，m 维列向量也是 $m \times 1$ 矩阵.

根据命题 1.2.2，这个矩阵 A 完全决定一个线性映射 f；根据命题 1.2.3，任何矩阵 A 都唯一确定一个线性映射 f.

① 具体计算或书写中，为避免单下标和双下标中可能的歧义，可以用逗号把双下标隔开，如 $a_{i,j-1}$.

定义 1.2.4 (线性映射的表示矩阵) 设线性映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$，$\boldsymbol{e}_i\,(i=1,2,\cdots,n)$ 为 $\mathbb{R}^n$ 的标准坐标向量，若 $\boldsymbol{a}_i = f(\boldsymbol{e}_i)$，则称矩阵 $A = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}$ 为线性映射 f 在标准坐标向量下的**表示矩阵**.

以后我们用 $\boldsymbol{A}$ 表示这个线性映射[①]. 因此，$\boldsymbol{A}(\boldsymbol{x}) = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$，其中 $\boldsymbol{x} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} \in \mathbb{R}^n$，$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n \in \mathbb{R}^m$ 是矩阵 A 的列向量组. 下面我们定义一种新的运算来表示像 $\boldsymbol{A}(\boldsymbol{x})$.

定义 1.2.5 (矩阵与向量的乘积) 定义 $m\times n$ 矩阵 A 和 n 维列向量 $\boldsymbol{x}$ 的乘积：

$$A\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & \cdots & a_{2n}\\ \vdots & \vdots & & & \vdots\\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\\vdots\\x_n\end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\\ \vdots\\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n\end{bmatrix}.$$

注意，根据定义，当 $n \neq k$ 时，$m\times n$ 矩阵和 k 维列向量不能相乘.

把 A 写成列向量组的横向排列 $A = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}$，则 $A\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$，即矩阵与向量的乘积是矩阵的列向量组的一个线性组合，系数由 $\boldsymbol{x}$ 的分量给出.

例 1.2.6 当行向量 $\boldsymbol{a}^{\mathrm{T}} = \begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}$ 与列向量 $\boldsymbol{x} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$ 进行矩阵乘法时，结果为 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$. 这恰好是我们中学学过的二维向量的内积的推广，我们称其为向量 $\boldsymbol{a}$ 和向量 $\boldsymbol{x}$ 的**内积**. 可见，矩阵与列向量的乘积，就是将这个矩阵的每一行对应的向量分别与这个列向量做内积而得到的向量，即，把 A 写成行向量组的纵向排列 $A = \begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\\vdots\\\tilde{\boldsymbol{a}}_m^{\mathrm{T}}\end{bmatrix}$，则

$$A\boldsymbol{x} = \begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\\vdots\\\tilde{\boldsymbol{a}}_m^{\mathrm{T}}\end{bmatrix}\boldsymbol{x} = \begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{x}\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{x}\\\vdots\\\tilde{\boldsymbol{a}}_m^{\mathrm{T}}\boldsymbol{x}\end{bmatrix}, \tag{1.2.2}$$

① 黑体大写字母表示的映射在本书中常暗示其矩阵用相应的大写字母表示.

其中 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}, i=1,2,\cdots,m$，是 $1\times n$ 矩阵，而矩阵与列向量的乘积 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{x}$ 是 1×1 矩阵，也就是一个实数. ☺

于是，线性映射 $\boldsymbol{A}$ 就可以写成矩阵与向量的乘积的形式：$\boldsymbol{A}(\boldsymbol{x})=A\boldsymbol{x}$，而线性映射保持线性运算的性质就对应于乘积的运算法则.

命题 1.2.7　*矩阵和向量的乘积对向量的线性运算满足分配律：*

$$A(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2)=k_1(A\boldsymbol{x}_1)+k_2(A\boldsymbol{x}_2).$$

例 1.2.8　对例 1.1.8 中的线性映射，我们计算其表示矩阵.

1. “结账”是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的线性映射,其中定义域 $\mathbb{R}^3$ 的标准坐标向量分别是一千克苹果、一千克香蕉和一千克樱桃，它们对应的单价分别是 12.98 元、9.98 元和 129.98 元. 因此结账作为线性映射，它的表示矩阵是

$$\begin{aligned}A&=\begin{bmatrix}\text{结账(一千克苹果)} & \text{结账(一千克香蕉)} & \text{结账(一千克樱桃)}\end{bmatrix}\\&=\begin{bmatrix}12.98 & 9.98 & 129.98\end{bmatrix}.\end{aligned}$$

其表达式可以由矩阵和向量乘法给出

$$\boldsymbol{A}(\boldsymbol{x})=A\boldsymbol{x}=\begin{bmatrix}12.98 & 9.98 & 129.98\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=12.98x_1+9.98x_2+129.98x_3.$$

2. “营养”是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}^2$ 的线性映射. 它的表示矩阵为

$$\begin{aligned}A&=\begin{bmatrix}\text{营养(一千克苹果)} & \text{营养(一千克香蕉)} & \text{营养(一千克樱桃)}\end{bmatrix}\\&=\begin{bmatrix}12 & 12 & 3\\135 & 208 & 99\end{bmatrix}.\end{aligned}$$

其表达式为

$$\boldsymbol{A}(\boldsymbol{x})=A\boldsymbol{x}=\begin{bmatrix}12 & 12 & 3\\135 & 208 & 99\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}12x_1+12x_2+3x_3\\135x_1+208x_2+99x_3\end{bmatrix}.$$ ☺

特别地，$\mathbb{R}^n$ 上的线性变换的表示矩阵是 $n\times n$ 矩阵，又称为 n 阶**方阵**.

例 1.2.9　考虑例 1.1.10 中平面向量构成的线性空间 $\mathbb{R}^2$ 上的线性变换的表示矩阵.

1. 旋转变换 $\boldsymbol{R}_\theta$ 的表示矩阵为

$$R_\theta=\begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}.$$

其表达式为

$$\boldsymbol{R}_\theta(\boldsymbol{x}) = R_\theta \boldsymbol{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{bmatrix}.$$

2. 反射变换 $\boldsymbol{H}_\theta$ 的表示矩阵为

$$H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

其表达式为

$$\boldsymbol{H}_\theta(\boldsymbol{x}) = H_\theta \boldsymbol{x} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1\cos 2\theta + x_2\sin 2\theta \\ x_1\sin 2\theta - x_2\cos 2\theta \end{bmatrix}.$$

3. 对换变换 $\boldsymbol{P}$ 的表示矩阵为

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

其表达式为

$$\boldsymbol{P}(\boldsymbol{x}) = P\boldsymbol{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}.$$

4. 伸缩变换 $\boldsymbol{C}_k$ 的表示矩阵为

$$C_k = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}.$$

其表达式为

$$\boldsymbol{C}_k(\boldsymbol{x}) = C_k\boldsymbol{x} = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} kx_1 \\ x_2 \end{bmatrix}.$$

5. 错切变换 $\boldsymbol{S}_k$ 的表示矩阵为

$$S_k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}.$$

其表达式为

$$\boldsymbol{S}_k(\boldsymbol{x}) = S_k\boldsymbol{x} = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ kx_1 + x_2 \end{bmatrix}.$$ ☺

下面介绍几类特殊的方阵，并讨论其对应线性变换的求原像问题.

容易得到，(1.1.2) 式中 $\mathbb{R}^n$ 上的恒同变换的表示矩阵为

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

即对角元素 (从左上到右下的对角线上的元素, 或者所在行数和所在列数相等的元素) 都是 1, 非对角元素都是 0. 这个矩阵称为 n 阶**恒同矩阵**或 n 阶**单位矩阵**, 记为 I_n. 它还可以用标准坐标向量组表示：$I_n = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1^{\mathrm{T}} \\ \boldsymbol{e}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{e}_n^{\mathrm{T}} \end{bmatrix}$. 对 $\mathbb{R}^n$ 中任意向量 $\boldsymbol{x}$，容易验证 $I_n\boldsymbol{x} = \boldsymbol{x}$. 这意味着它对应的线性映射 $\boldsymbol{I}$ 能够很容易地由像求出原像（二者相等).

一般地, 非对角元素全为零的方阵称为**对角矩阵**, 而其中为零的元素往往省略不写:

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{bmatrix} = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} =: \operatorname{diag}(d_1, d_2, \cdots, d_n).$$

如果对角元素依次为 $d_1, d_2, \cdots, d_n$，那么这个对角矩阵常用 $\operatorname{diag}(d_1, d_2, \cdots, d_n)$ 表示.

如果对角矩阵的对角元素都不为零，则向量 $\boldsymbol{b} = \begin{bmatrix} b_i \end{bmatrix}$ 在对应线性映射 $\boldsymbol{D}$ 下的原像就是向量 $\boldsymbol{x} = \begin{bmatrix} \dfrac{b_i}{d_i} \end{bmatrix}$. 事实上，由

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1x_1 \\ d_2x_2 \\ \vdots \\ d_nx_n \end{bmatrix},$$

就能直接得到 $\boldsymbol{x}$ 的所有元素.

形如 $U = \begin{bmatrix} u_{11} & \cdots & u_{1n} \\ & \ddots & \vdots \\ & & u_{nn} \end{bmatrix}$ 的方阵称为 n 阶**上三角矩阵**，形如 $L = \begin{bmatrix} l_{11} & & \\ \vdots & \ddots & \\ l_{n1} & \cdots & l_{nn} \end{bmatrix}$ 的方阵称为 n 阶**下三角矩阵**. 如果上（下）三角矩阵的对角元素都是 0，则称为**严格上（下）三角矩阵**. 如果上（下）三角矩阵的对角元素都是 1, 则称为**单位上（下）三角矩阵**.

如果上三角矩阵的对角元素都不为零, 那么给定向量 $\boldsymbol{b}$，它在对应线性映射 $\boldsymbol{U}$ 下的

原像 $\boldsymbol{x}$ 也容易求得. 事实上, 观察

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n \\ u_{22}x_2 + \cdots + u_{2n}x_n \\ \vdots \\ u_{nn}x_n \end{bmatrix},$$

由第 n 个分量相等 $b_n = u_{nn}x_n$, 立刻得到 $x_n = \dfrac{b_n}{u_{nn}}$. 把 x_n 代入第 $n-1$ 个分量相等的关系式 $b_{n-1} = u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n$, 立得 $x_{n-1} = \dfrac{b_{n-1} - u_{n-1,n}x_n}{u_{n-1,n-1}}$. 类似进行下去, 可得 $\boldsymbol{x}$ 的所有元素. 这种自下而上逐个代入从而求出每一个分量的方法, 称为**回代法**.

类似地, 如果下三角矩阵的对角元素都不为零, 那么给定向量 $\boldsymbol{b}$, 它在对应线性映射 $\boldsymbol{L}$ 下的原像 $\boldsymbol{x}$ 也容易求得. 事实上, 观察

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} l_{11}x_1 \\ l_{21}x_1 + l_{22}x_2 \\ \vdots \\ l_{n1}x_1 + l_{n2}x_2 + \cdots + l_{nn}x_n \end{bmatrix},$$

即可从第 1 个分量相等的关系开始计算, 从 x_1 到 x_n 逐个得出 $\boldsymbol{x}$ 的所有元素. 这种自上而下逐个代入从而求出每一个分量的方法, 称为**前代法**.

利用回代法或前代法可以求出向量在表示矩阵是对角元素都不为零的三角矩阵的线性映射下的原像. 那么对任意线性映射, 有何方法来求解向量的原像呢? 这将在下节详细阐述.

习题

练习 1.2.1 将下列向量 $\boldsymbol{b}$ 写成矩阵和向量乘积的形式:

1. $\boldsymbol{b} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

2. $\boldsymbol{b} = 5\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} + 4\begin{bmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$.

3. $\boldsymbol{b} = \begin{bmatrix} 2b+a+c \\ c-b \\ a+b+c \\ a+b \end{bmatrix}$, 其中 a, b, c 为常数.

4. $\boldsymbol{b} = f\left(\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}\right)$, 其中 $f\colon \mathbb{R}^5 \to \mathbb{R}^5$ 是线性变换, 满足 $f(\boldsymbol{e}_k) = k\boldsymbol{e}_{6-k}, k = 1, 2, \cdots, 5$.

5. 假设如果某天下雨，则第二天下雨概率为 0.8；如果当天不下雨，则第二天下雨概率为 0.3. 已知当天有一半的概率会下雨，令 $\boldsymbol{b} \in \mathbb{R}^2$，其两个分量分别是明天下雨和不下雨的概率.

练习 1.2.2 判断下列矩阵和向量的乘积是否良定义①. 在可以计算时，先将其写成列向量的线性组合，再进行计算：

1. $\begin{bmatrix}1&0\\1&1\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}$.
2. $\begin{bmatrix}1&0\\1&1\\1&1\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}$.
3. $\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}$.
4. $\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\end{bmatrix}$.
5. $\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}\begin{bmatrix}1\\3\\5\\7\end{bmatrix}$.
6. $\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}$.
7. $\begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}$.
8. $\begin{bmatrix}1&1&0\\3&2&1\\7&4&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}$.
9. $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}1\\-1\\-1\end{bmatrix}$.
10. $\begin{bmatrix}1&3&7\\1&2&4\\0&1&3\end{bmatrix}\begin{bmatrix}2\\-3\\1\end{bmatrix}$.

练习 1.2.3 设 $A = \begin{bmatrix}0.8&0.3\\0.2&0.7\end{bmatrix}, \boldsymbol{u}_0 = \begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{v}_0 = \begin{bmatrix}1\\3\end{bmatrix}$. 对任意自然数 i，令 $\boldsymbol{u}_{i+1} = A\boldsymbol{u}_i, \boldsymbol{v}_{i+1} = A\boldsymbol{v}_i$.

1. 对 $i = 1, 2, 3, 4$，计算 $\boldsymbol{u}_i, \boldsymbol{v}_i$.
2. 猜测 $\lim\limits_{i\to\infty} \boldsymbol{u}_i, \lim\limits_{i\to\infty} \boldsymbol{v}_i$.
3. 任取初始向量 $\boldsymbol{w}_0$，猜测极限 $\lim\limits_{i\to\infty} \boldsymbol{w}_i$，其中 $\boldsymbol{w}_{i+1} = A\boldsymbol{w}_i$；不同初始向量得到的极限在同一条直线上吗?

练习 1.2.4 设线性变换 f 的表示矩阵为 $A = \begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}$.

1. 令 $\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix} = \begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$，将 y_1, y_2, y_3 分别用 x_1, x_2, x_3 表达出来.
2. 将 x_1, x_2, x_3 分别用 y_1, y_2, y_3 表达出来.
3. 找到矩阵 B，使得 $\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = B\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$.
4. 设 g 是由矩阵 B 决定的线性变换，证明 f, g 互为逆变换.

练习 1.2.5 计算下列线性映射 f 的表示矩阵：

① 所谓良定义，就是说定义不产生矛盾. 这里就是指可以按定义计算.

1. f 为 xy 平面向 y 轴的投影变换.

2. $f\colon \mathbb{R}^3 \to \mathbb{R}$，$f(\boldsymbol{x}) = \begin{bmatrix}1\\2\\3\end{bmatrix} \cdot \boldsymbol{x}$，其中点积的定义见练习 1.1.10.

3. $f\colon \mathbb{R}^3 \to \mathbb{R}^3$，$f(\boldsymbol{x}) = \begin{bmatrix}1\\2\\3\end{bmatrix} \times \boldsymbol{x}$，其中叉积的定义见练习 1.1.10.

4. $f\colon \mathbb{R}^4 \to \mathbb{R}^4$，$f\left(\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}\right) = \begin{bmatrix}x_1\\x_2\\x_3\\x_4+x_2\end{bmatrix}$.

5. $f\colon \mathbb{R}^4 \to \mathbb{R}^4$，$f\left(\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}\right) = \begin{bmatrix}x_4\\x_2\\x_3\\x_1\end{bmatrix}$.

6. $f\colon \mathbb{R}^m \to \mathbb{R}^m$，$f\left(\begin{bmatrix}x_1\\x_2\\\vdots\\x_m\end{bmatrix}\right) = \begin{bmatrix}x_m\\x_{m-1}\\\vdots\\x_1\end{bmatrix}$.

练习 1.2.6 考虑桥墩载荷问题，其中 l_1, l_2, d 作为常数.

1. 映射 f 输入 F_1, F_2 得到输出 f_1, f_2，写出 f 的表示矩阵.
2. 以桥梁的左端为支点，逆时针方向的力矩为 $df_2 - l_1F_1 - l_2F_2$. 以桥梁的右端为支点或者以桥梁的中点作为支点，都能类似地得到逆时针方向的力矩. 假设映射 f 的输入为 F_1, F_2, f_1, f_2，而输出为桥梁的左端、中点和右端的逆时针方向的力矩，写出 f 的表示矩阵.

练习 1.2.7 设 xy 平面 $\mathbb{R}^2$ 上的变换 f 是下列三个变换的复合：先绕原点逆时针旋转 $\dfrac{\pi}{6}$；然后进行一个保持 y 坐标不变，同时将 y 坐标的两倍加到 x 坐标上的错切；最后再沿着直线 $x + y = 0$ 反射.

1. 证明 f 是线性变换.
2. 计算 $f(\boldsymbol{e}_1)$ 和 $f(\boldsymbol{e}_2)$.
3. 写出 f 的表示矩阵.
4. 计算 $f\left(\begin{bmatrix}3\\4\end{bmatrix}\right)$.

练习 1.2.8

1. 幻方矩阵是指元素分别是 $1, 2, \cdots, 9$ 的三阶矩阵 M，且每行、每列以及两条对角线上的三个元素之和都相同. 求 $M\begin{bmatrix}1\\1\\1\end{bmatrix}$ 的所有可能值.

2. 数独矩阵是指 9 阶矩阵 M，从上到下、从左到右依次分成九个 3×3 的子矩阵，且每行、每列以及九个子矩阵中的元素都是 $1, 2, \cdots, 9$. 求 $M\begin{bmatrix}1\\\vdots\\1\end{bmatrix}$ 的所有可能值.

练习 1.2.9 ☕ 证明，如果 n 阶方阵 A 对任意 n 维列向量 $\boldsymbol{x}$，都有 $A\boldsymbol{x}=\boldsymbol{0}$，则 A 是元素全为 0 的矩阵.

练习 1.2.10 ☕ 证明，如果 n 阶方阵 A 对任意 n 维列向量 $\boldsymbol{x}$，都存在依赖于 $\boldsymbol{x}$ 的常数 $c(\boldsymbol{x})$，满足 $A\boldsymbol{x}=c(\boldsymbol{x})\boldsymbol{x}$，则存在常数 c，使得 $A=cI_n$.

1.3 线性方程组

对线性映射 $\boldsymbol{A}\colon \mathbb{R}^n \to \mathbb{R}^m$，考虑求解原像的问题. 这等价于求解如下线性方程组：

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases}$$

例 1.3.1 (鸡兔同笼) 一个笼子中有鸡和兔共 8 只，二者共有 22 条腿，那么鸡和兔各有多少只?

利用二元一次方程组来求解. 设鸡有 x_1 只，兔有 x_2 只，那么：

$$\begin{cases} x_1+x_2=8, \\ 2x_1+4x_2=22. \end{cases}$$

通过消元法，把第一式的 -2 倍加到第二式上，就有

$$\begin{cases} x_1+x_2=8, \\ 2x_2=6. \end{cases}$$

再给第二式乘以 $\dfrac{1}{2}$，得

$$\begin{cases} x_1+x_2=8, \\ x_2=3. \end{cases}$$

然后把第二式代入第一式（也可以解释为把第二式的 -1 倍加到第一式上），即得结果

$$\begin{cases} x_1=5, \\ x_2=3. \end{cases}$$

☺

通过观察可知，在解线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的过程中，未知数没有参与任何计算. 因此，我们完全可以不写出未知数 $\boldsymbol{x}$，直接在**系数矩阵** A 和常数项 $\boldsymbol{b}$ 上计算，也就是在如下矩阵

$$\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} = \left[\begin{array}{c|c} A & \boldsymbol{b} \end{array}\right] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

上计算，其中矩阵 $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$ 称为线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的**增广矩阵**.

我们在增广矩阵上计算例 1.3.1：

$$\left[\begin{array}{cc|c} 1 & 1 & 8 \\ 2 & 4 & 22 \end{array}\right] \xrightarrow[\text{加到第二行}]{\text{第一行的 } -2 \text{ 倍}} \left[\begin{array}{cc|c} 1 & 1 & 8 \\ 0 & 2 & 6 \end{array}\right] \xrightarrow[\text{乘以 } \frac{1}{2}]{\text{第二行}} \left[\begin{array}{cc|c} 1 & 1 & 8 \\ 0 & 1 & 3 \end{array}\right] \xrightarrow[\text{加到第一行}]{\text{第二行的 } -1 \text{ 倍}} \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 3 \end{array}\right].$$

上述解法中对线性方程组有如下两种变换：(1) 把某一方程的 k 倍加到另一方程上，用以消去某个变元；(2) 把某一方程乘以非零常数 k，使得某个变元的系数为 1. 对增广矩阵有相应的两种行变换. 反复利用二者，当系数矩阵变为恒同矩阵 I_2 时，得到方程组有唯一解，且解就由此时的常数项给出.

无论未知数有多少个，如果条件适合，那么我们都可以用同样的办法逐个消去等式中的未知数，直到得到线性方程组的解. 我们再来看另一个例子.

例 1.3.2 求解

$$\begin{cases} \phantom{x_1-{}}2x_2 - x_3 = 1, \\ x_1 - x_2 + x_3 = 0, \\ 2x_1 - 2x_2 + 2x_3 = 1. \end{cases}$$

在增广矩阵上计算：

$$\left[\begin{array}{ccc|c} 0 & 2 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 2 & -2 & 2 & 1 \end{array}\right] \xrightarrow{\text{对换第一、二行}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 2 & -2 & 2 & 1 \end{array}\right] \xrightarrow[\text{加到第三行}]{\text{第一行的 } -2 \text{ 倍}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

注意到，第三个方程 $0x_1+0x_2+0x_3=1$ 无解. 因此，原线性方程组无解. ☺

在这个例子中，出于形式美的需要，我们依然希望用一个方程消去其他方程中的 x_1. 由于第一个方程中 x_1 的系数为零，也就是系数矩阵左上角的元素 $a_{11}=0$，我们需要首先对换第一行和第二行，才能再用新得到的第一行做行变换消去变元 x_1.

定义 1.3.3 (初等变换) 对线性方程组施加的如下三类变换的每一类都称为线性方程组的**初等变换**：

1. **对换变换**：互换两个方程的位置；
2. **倍乘变换**：把某个方程两边同乘一个非零常数 k；
3. **倍加变换**：把某个方程的 k 倍加到另一个方程上.

容易看到，初等变换可逆：

1. 对换变换的逆变换是它本身；
2. 参数为 k 的倍乘变换的逆变换也是倍乘变换，参数为 $\dfrac{1}{k}$；

3. 参数为 k 的倍加变换的逆变换也是倍加变换，参数为 $-k$.

由此得到如下重要的结论.

定理 1.3.4 线性方程组经某个初等变换后得到的新方程组与原方程组同解.

证 设线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有一组解 $x_1=k_1, x_2=k_2, \cdots, x_n=k_n$，则有恒等式 $\boldsymbol{A}\boldsymbol{k}=\boldsymbol{b}$. 容易验证, 如果对换两个方程的位置, 那么再代入 $x_1=k_1, x_2=k_2, \cdots, x_n=k_n$，依旧得到恒等式，因此 $x_1=k_1, x_2=k_2, \cdots, x_n=k_n$ 仍然是新方程组的一组解. 类似地，方程组施加了倍加变换或者倍乘变换后，代入 $x_1=k_1, x_2=k_2, \cdots, x_n=k_n$ 后也得到恒等式，即它依旧是新方程组的一组解. 于是，原方程组的任意一组解都是新方程组的解.

反过来，由于初等变换可逆，新方程组也可以通过某个初等变换得到原方程组，因此新方程组的任意一组解都是原方程组的解. 于是两个方程组同解. □

如前解法中，在线性方程组上计算和在增广矩阵上计算完全相同. 因此，对应于线性方程组的初等变换，我们可以定义矩阵的初等变换.

定义 1.3.5 (初等行（列）变换) 对矩阵施加的如下三类变换的每一类都称为矩阵的**初等行（列）变换**:

1. **对换变换**: 互换两行（列）的位置;
2. **倍乘变换**: 某一行（列）乘以非零常数 k;
3. **倍加变换**: 把某个行（列）的 k 倍加到另一个行（列）上.

例 1.3.6 求解

$$\begin{cases} \qquad\qquad 2x_3+x_4=1, \\ x_1-x_2+4x_3 \qquad =0. \end{cases}$$

在增广矩阵上计算:

$$\left[\begin{array}{cccc|c} 0 & 0 & 2 & 1 & 1 \\ 1 & -1 & 4 & 0 & 0 \end{array}\right] \xrightarrow[\text{第一、二行}]{\text{对换变换}} \left[\begin{array}{cccc|c} 1 & -1 & 4 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{array}\right]$$

$$\xrightarrow[\text{第二行}]{\text{倍乘变换}} \left[\begin{array}{cccc|c} 1 & -1 & 4 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right] \xrightarrow[\text{第一、二行}]{\text{倍加变换}} \left[\begin{array}{cccc|c} 1 & -1 & 0 & -2 & -2 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right].$$

原方程组变换为

$$\begin{cases} x_1-x_2 \qquad -2x_4=-2, \\ \qquad\qquad x_3+\dfrac{1}{2}x_4=\dfrac{1}{2}. \end{cases}$$

把 x_2, x_4 移到右边，得

$$x_1=-2+x_2+2x_4, \quad x_3=\frac{1}{2}-\frac{1}{2}x_4. \tag{1.3.1}$$

任取 x_2, x_4 的一组值，就得到方程组的一组解．注意，在变换后的系数矩阵中，x_1, x_3 所在的两列恰好构成恒同矩阵 I_2．☺

如例 1.3.6 中的 x_2, x_4 这类可以任意取值的未知数，称为**自由变量**，而取值依赖于右端项和自由变量的未知数，称为**主变量**．而如 (1.3.1) 式这样的表达式，称为**通解**或一**般解**．通解就是用含自由变量的表达式表示主变量．注意，在变换后的系数矩阵中，主变量的系数组成的列恰好构成恒同矩阵，这使得我们可以通过移项直接写出通解．

在消元的过程中，首先选择的是主变量，其余的变量即为自由变量．而主变量的选取并不是唯一的．如在例 1.3.6 中，可以选择 x_2, x_4 为主变量进行消元，其对应的增广矩阵计算为

$$\left[\begin{array}{cccc|c} 0 & 0 & 2 & 1 & 1 \\ 1 & -1 & 4 & 0 & 0 \end{array}\right] \xrightarrow[\text{第一、二行}]{\text{对换变换}} \left[\begin{array}{cccc|c} 1 & -1 & 4 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{array}\right] \xrightarrow[\text{第一行}]{\text{倍乘变换}} \left[\begin{array}{cccc|c} -1 & 1 & -4 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{array}\right].$$

通解为 $x_2 = x_1 + 4x_3, \quad x_4 = 1 - 2x_3$．

在实际计算中，我们通常选择角标小的变量（对应矩阵中靠左的列）作为主变量进行消元．观察上面的例子，对矩阵的计算总体上分两步．

第一步：从角标最小的变量开始，必要时通过换行把该变量的系数非零的方程换到上面，然后利用此方程向下消元．变换后的矩阵满足如下特点：

1. 元素全为 0 的行（称为**零行**），只可能存在于下方；
2. 元素不全为 0 的行（称为**非零行**），从左数第一个不为 0 的元素（称为**主元**）的列指标随着行指标的增加而严格增加．

这样的矩阵称为**阶梯形矩阵**．阶梯形矩阵具有如下形式：

$$\begin{bmatrix} & \bullet & * & \cdots & * & * & \cdots & * & * & \cdots & * & * & \cdots & * \\ & & & & \bullet & * & \cdots & * & * & \cdots & * & * & \cdots & * \\ & & & & & & & \bullet & * & \cdots & * & * & \cdots & * \\ & & & & & & & & & \ddots & \vdots & \vdots & & \vdots \\ & & & & & & & & & & \bullet & * & \cdots & * \\ & & & & & & & & & & & & & \end{bmatrix},$$

其中“$*$”表示可能不为 0 的数，“$\bullet$”表示一定不为 0 的数，即“$\bullet$”处元素是主元．一个阶梯形矩阵的非零行数称为它的**阶梯数**．当求解线性方程组时，把增广矩阵化成阶梯形矩阵后，就可以判断方程组解的情形．

在线性方程组的系数矩阵中，化成阶梯形后主元所在的列，对应着主变量的系数，称为**主列**；其他列，对应着自由变量的系数，称为**自由列**．

第二步：先通过倍乘变换将主元化成 1，然后从角标最大的主变量（对应矩阵中靠右的列）开始，依次向上消元．变换后的矩阵仍是阶梯形矩阵，并满足如下特点：

1. 每个非零行的主元都是 1;
2. 每个主列除主元外的其他元素都是 0.

这样的阶梯形矩阵称为**行简化阶梯形矩阵**. 行简化阶梯形矩阵具有如下形式:

$$\begin{bmatrix} 1 & * & \cdots & 0 & * & \cdots & 0 & * & \cdots & 0 & * & \cdots & * \\ & & & 1 & * & \cdots & 0 & * & \cdots & 0 & * & \cdots & * \\ & & & & & & 1 & * & \cdots & 0 & * & \cdots & * \\ & & & & & & & & \ddots & \vdots & \vdots & & \vdots \\ & & & & & & & & & 1 & * & \cdots & * \\ & & & & & & & & & & & & \end{bmatrix},$$

注意, 在行简化阶梯形矩阵中, 主元所在的行和列的交叉点上的元素组成的矩阵是恒同矩阵. 当求解方程组时, 把增广矩阵化成行简化阶梯形矩阵后, 就可以写出它的通解.

下面, 我们来解决理论问题, 即说明使用初等变换就可以把矩阵化为阶梯形和行简化阶梯形.

定理 1.3.7

1. *任意矩阵都可以用对换行变换和倍加行变换化为阶梯形;*
2. *任意矩阵都可以用初等行变换化为行简化阶梯形.*

证 先证明第 1 条. 对 $m \times n$ 矩阵, 对 m 用数学归纳法. $m = 1$ 时, 矩阵只有一行, 自然是阶梯形. 注意保持不变也是一种倍加变换, 故 $m = 1$ 时, 线性方程组可以通过倍加变换化为阶梯形. 现假设对任意 n, $(m-1) \times n$ 矩阵都可以通过对换变换和倍加变换化为阶梯形, 考察 $m \times n$ 矩阵, 分如下情形讨论.

1. 如果 $a_{11} \neq 0$, 那么把第 1 行的 $-\dfrac{a_{i1}}{a_{11}}$ 倍加到第 i 个方程上 (倍加变换), 就可以把其他行中的第一个元素化成 0. 那么第 2 到 m 行和第 2 到 n 列交叉点上的元素组成的矩阵有 $m-1$ 行, 根据归纳假设, 可以化为阶梯形. 从而原矩阵也可以化为阶梯形.
2. 否则 $a_{11} = 0$. 如果 $a_{21}, \cdots, a_{m1}$ 中有某个不为 0, 记为 a_{i1}. 则把第 $1, i$ 两方程互换位置 (对换变换), 问题归于第一种情形.
3. 否则 $a_{11}, \cdots, a_{m1}$ 全为 0, 即矩阵第 1 列元素全为 0, 则类似前面情形考察矩阵的其他列. 如果存在 $j \leqslant n$ 使得矩阵的第 $2, \cdots, j-1$ 列全为 0, 而第 j 列的元素不全为 0, 类似第一、二种情形, 可以将原方程组化为阶梯形.
4. 否则矩阵的所有元素全为 0, 则自然是阶梯形.

第 2 条留给读者. □

我们将在推论 2.3.4 中证明, 矩阵化成的行简化阶梯形唯一. 矩阵 A 的行简化阶梯形记为 $\operatorname{rref}(A)$.

阶梯形矩阵可以定性地判断方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 是否有解. 具体来说，用初等行变换把增广矩阵 $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 化成阶梯形，系数矩阵 A 也同时化成了阶梯形. 此时就可以确定解的情形.

定理 1.3.8 (判定定理)

1. 如果 $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 对应的阶梯数比 A 对应的阶梯数多 1 (此时方程出现了矛盾)，则方程组无解.
2. 如果 $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 和 A 对应的阶梯数相等，则方程组有解. 其中，
   (a) 如果阶梯形的阶梯数和未知数个数相等，则方程组有唯一解；
   (b) 如果阶梯形的阶梯数小于未知数个数，则方程组有无穷多组解.

例 1.3.2 中, $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 的阶梯数是 3, 而 A 的阶梯数是 2, 此时方程组无解. 例 1.3.1 中, $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 和 A 的阶梯数都是 2，未知数个数也是 2，此时方程组有唯一解. 例 1.3.6 中, $\begin{bmatrix}A & \boldsymbol{b}\end{bmatrix}$ 和 A 的阶梯数都是 2，而未知数个数是 4，此时方程组有无穷多组解.

线性方程组有解的情形下, 将增广矩阵化成阶梯形进行消元的办法, 称为 **Gauss 消元法**. 化为阶梯形后，可以自下而上逐个代入从而求出主变量的表达式，这种方法仍称为**回代法**；也可以进一步将增广矩阵从阶梯形化成行简化阶梯形，通过移项直接写出通解公式，这种方法称为 **Gauss-Jordan 消元法**. 由于计算在矩阵上进行，Gauss 消元法和 Gauss-Jordan 消元法又称为矩阵消元法.

例 1.3.9　求解

$$\begin{cases} x_1+x_2-3x_3=-1, \\ 2x_1+x_2-2x_3=1, \\ x_1+x_2+x_3=3, \\ x_1+2x_2-3x_3=3. \end{cases}$$

利用第一个方程对 x_1 消元，在增广矩阵上计算：

$$\left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & -3 & 3 \end{array}\right] \xrightarrow[\text{第一、二行}]{\text{倍加变换}} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 1 & 1 & 1 & 3 \\ 1 & 2 & -3 & 3 \end{array}\right]$$

$$\xrightarrow[\text{第一、三行}]{\text{倍加变换}} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 0 & 0 & 4 & 4 \\ 1 & 2 & -3 & 3 \end{array}\right] \xrightarrow[\text{第一、四行}]{\text{倍加变换}} \left[\begin{array}{ccc|c} 1 & 1 & -3 & -1 \\ 0 & -1 & 4 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 4 \end{array}\right].$$

可以看到，对矩阵上面的行计算时下面的行不变，因此只要按照从上到下的次序对行逐个计算，一次计算若干行也不会有干扰．因此如果从头开始计算，可以写成

$$\left[\begin{array}{ccc|c}1&1&-3&-1\\2&1&-2&1\\1&1&1&3\\1&2&-3&3\end{array}\right]\to\left[\begin{array}{ccc|c}1&1&-3&-1\\0&-1&4&3\\0&0&4&4\\0&1&0&4\end{array}\right]\to\left[\begin{array}{ccc|c}1&1&-3&-1\\0&-1&4&3\\0&0&1&1\\0&0&4&7\end{array}\right]\to\left[\begin{array}{ccc|c}1&1&-3&-1\\0&1&-4&-3\\0&0&1&1\\0&0&0&3\end{array}\right].$$

增广矩阵的阶梯数大于系数矩阵的阶梯数，故原方程组无解． ☺

例 1.3.10 求解

$$\begin{cases}x_1+x_2-3x_4-x_5=0,\\x_1-x_2+2x_3-x_4=0,\\4x_1-2x_2+6x_3+3x_4-4x_5=0,\\2x_1+4x_2-2x_3+4x_4-7x_5=0.\end{cases}$$

这个线性方程组的常数项全是 0，因此不论何种初等行变换，对应的计算结果全为 0．所以我们可以只对系数矩阵 A 做初等行变换：

$$\begin{bmatrix}1&1&0&-3&-1\\1&-1&2&-1&0\\4&-2&6&3&-4\\2&4&-2&4&-7\end{bmatrix}\to\begin{bmatrix}1&1&0&-3&-1\\0&-2&2&2&1\\0&-6&6&15&0\\0&2&-2&10&-5\end{bmatrix}$$

$$\to\begin{bmatrix}1&1&0&-3&-1\\0&-2&2&2&1\\0&0&0&9&-3\\0&0&0&12&-4\end{bmatrix}\to\begin{bmatrix}1&1&0&-3&-1\\0&-2&2&2&1\\0&0&0&9&-3\\0&0&0&0&0\end{bmatrix}.$$

化为阶梯形后，我们知道方程组有无穷多组解．继续化为行简化阶梯形：

$$\cdots\to\begin{bmatrix}1&1&0&-3&-1\\0&1&-1&-1&-\frac{1}{2}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}\to\begin{bmatrix}1&0&1&-2&-\frac{1}{2}\\0&1&-1&-1&-\frac{1}{2}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}\to\begin{bmatrix}1&0&1&0&-\frac{7}{6}\\0&1&-1&0&-\frac{5}{6}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}.$$

于是通解为

$$x_1=-x_3+\frac{7}{6}x_5,\quad x_2=x_3+\frac{5}{6}x_5,\quad x_4=\frac{1}{3}x_5,$$

其中 x_1,x_2,x_4 是主变量，x_3,x_5 是自由变量． ☺

例 1.3.10 中方程组的常数项 $\boldsymbol{b}=\boldsymbol{0}$. 方程组 $A\boldsymbol{x}=\boldsymbol{0}$, 称为**齐次**线性方程组. 齐次线性方程组显然有一组解 $x_1=0, x_2=0, \cdots, x_n=0$, 称为**零解**或**平凡解**. 除此之外其他的解（如果存在）称为**非零解**或**非平凡解**. 对应地, 不是齐次的线性方程组, 称为**非齐次**线性方程组.

如果 $A\boldsymbol{x}=\boldsymbol{0}$ 成立, 则 $A(k\boldsymbol{x})=k(A\boldsymbol{x})=k\boldsymbol{0}=\boldsymbol{0}$. 因此, 齐次线性方程组如果有一个非零解, 则有无穷多组解. 阶梯形可以用来判断齐次线性方程组是否有非零解.

命题 1.3.11　一个齐次线性方程组只有零解当且仅当其系数矩阵的（行简化）阶梯形的阶梯数等于未知数个数. 等价地, 一个齐次线性方程组有无穷多组解当且仅当它的（行简化）阶梯形的阶梯数小于未知数个数.

证　由于常数项全为零, 只需化简系数矩阵. 由定理 1.3.8 立得结论. □

由此容易得到如下推论.

命题 1.3.12　一个齐次线性方程组中的方程个数小于未知数个数, 则一定存在非零解.

证　齐次线性方程组对应的阶梯形矩阵的阶梯数小于等于方程的个数, 因此小于未知数的个数, 所以一定存在非零解. □

命题 1.3.12 对齐次线性方程组解的定性判断, 不用通过计算化为阶梯形再得到, 因此常会带来一些便利.

习题

练习 1.3.1　把下列矩阵化为行简化阶梯形, 并回答问题:

1. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$, 化简后第一列是否为主列?

2. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$, 化简后第二列是否为主列?

3. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$, 化简后第三列是否为主列?

4. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$, 化简后第四列是否为主列?

5. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$, 化简后第五列是否为主列?

练习 1.3.2 下列线性方程组有解时，找到所有解；线性方程组无解时，对线性方程组做初等变换，得到矛盾表达式 $0=1$：

1. $\begin{bmatrix}1&2\\3&4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3\\7\end{bmatrix}$.

2. $\begin{bmatrix}1&3&5\\2&4&6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$.

3. $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

4. $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

5. $\begin{bmatrix}1&0&1\\2&3&5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$.

6. $\begin{bmatrix}1&2\\0&3\\1&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$.

7. $\begin{bmatrix}1&2\\0&3\\1&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

8. $\begin{bmatrix}2&-1&0&0\\-1&2&-1&0\\0&-1&2&-1\\0&0&-1&2\end{bmatrix}\begin{bmatrix}x\\y\\z\\t\end{bmatrix}=\begin{bmatrix}0\\0\\0\\5\end{bmatrix}$.

9. $\begin{bmatrix}2&1&0&0\\1&2&1&0\\0&1&2&1\\0&0&1&2\end{bmatrix}\begin{bmatrix}x\\y\\z\\t\end{bmatrix}=\begin{bmatrix}0\\0\\0\\5\end{bmatrix}$.

练习 1.3.3 将下列问题首先化成 $A\boldsymbol{x}=\boldsymbol{b}$ 的形式，然后求所有解：

1. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第三列写成前两列的线性组合？

2. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第四列写成前两列的线性组合？

3. 对 $\begin{bmatrix}1&2&3&4&5\\6&7&8&9&10\\11&12&13&14&15\end{bmatrix}$，如何将其第五列写成前两列的线性组合？

练习 1.3.4 求证：齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2=0,\\a_{21}x_1+a_{22}x_2=0\end{cases}$$

有非零解当且仅当 $a_{11}a_{22}-a_{12}a_{21}=0$.

练习 1.3.5 求满足下列条件的常数 b,c：

1. 方程组 $\begin{bmatrix}1&2\\3&b\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3\\7\end{bmatrix}$ 无解.

2. 方程组 $\begin{bmatrix}3&2\\6&4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}10\\c\end{bmatrix}$ 无解.

3. 方程组 $\begin{bmatrix}2&5&1\\4&b&1\\0&1&-1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\2\\3\end{bmatrix}$ 无解.

4. 方程组 $\begin{bmatrix}b&3\\3&b\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}6\\-6\end{bmatrix}$ 有无穷多组解.

5. 方程组 $\begin{bmatrix}2&b\\4&8\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}16\\c\end{bmatrix}$ 有无穷多组解.

6. 方程组 $\begin{bmatrix}1&b&0\\1&-2&-1\\0&1&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$ 有非零解.

7. 方程组 $\begin{bmatrix}b&2&3\\b&b&4\\b&b&b\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$ 有非零解（求三个不同的 b).

练习 1.3.6 设齐次线性方程组

$$\begin{cases}2x_1+x_2+\ x_3=0,\\ ax_1\qquad -\ x_3=0,\\ -x_1\qquad +3x_3=0\end{cases}$$

有非零解，求 a 的值，并求出所有的解.

练习 1.3.7 在如下关于 x_1,x_2,x_3 的线性方程组中，讨论在 p 取不同值时方程组是否有解，并在有解时，求出所有的解.

$$\begin{cases}px_1+\ x_2+\ x_3=1,\\ x_1+px_2+\ x_3=p,\\ x_1+\ x_2+px_3=p^2.\end{cases}$$

练习 1.3.8 设 $A=\begin{bmatrix}1&1&0\\-1&0&1\\0&-1&-1\end{bmatrix}, \boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$. 求证：

1. $A\boldsymbol{x}=\boldsymbol{b}$ 有解当且仅当 $b_1+b_2+b_3=0$.
2. 齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集是 $\{k\boldsymbol{x}_1 \mid k\in\mathbb{R}\}, \boldsymbol{x}_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}$.
3. 当 $A\boldsymbol{x}=\boldsymbol{b}$ 有解时，设 $\boldsymbol{x}_0$ 是一个解，则解集是 $\{\boldsymbol{x}_0+k\boldsymbol{x}_1 \mid k\in\mathbb{R}\}$.

练习 1.3.9 求三阶方阵 A, 使得线性方程组 $A\boldsymbol{x}=\begin{bmatrix}2\\4\\2\end{bmatrix}$ 的解集是 $\left\{\begin{bmatrix}2\\0\\0\end{bmatrix}+c\begin{bmatrix}1\\1\\0\end{bmatrix}+d\begin{bmatrix}0\\0\\1\end{bmatrix}\,\middle|\, c,d\in\mathbb{R}\right\}$.

练习 1.3.10

1. 构造三阶方阵，其元素各不相同，且行简化阶梯形有且只有一个主元.
2. ☕ 构造 100 阶方阵 A，所有元素非零，且行简化阶梯形恰有 99 个主元. 试描述 $A\boldsymbol{x}=0$ 的解集.

练习 1.3.11 给定线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$，其中 A 是 100 阶方阵. 假设 Gauss 消元法计算到最后一行得到 $0=0$.

1. 消元法是在计算 A 的行的线性组合. 因此 A 的 100 行的某个线性组合是?
2. 计算出 $0 = 0$ 说明方程组有无穷多组解. 因此 A 的 100 列的某个线性组合是?
3. 试说明 Gauss 消元法计算出的零行的个数和自由变量的个数相等.

练习 1.3.12 仅用从上往下的倍加变换 (即把上面的行的若干倍加到下面的行上), 将下列矩阵化为阶梯形, 并分析其主元的规律:

$$1.\ \begin{bmatrix}1&1&&\\1&2&1&\\&1&2&1\\&&1&2\end{bmatrix}.\qquad 2.\ \begin{bmatrix}2&1&&\\1&2&1&\\&1&2&1\\&&1&2\end{bmatrix}.\qquad 3.\ \begin{bmatrix}2&-1&&\\-1&2&-1&\\&-1&2&-1\\&&-1&2\end{bmatrix}.$$

这些矩阵都是**三对角矩阵**, 即除对角元素及与其相邻的元素外其余元素都是零的方阵.

练习 1.3.13 求解

$$\begin{cases} x_1 & & & +x_n &= 0,\\ x_1+x_2 & & & &= 0,\\ & x_2+x_3 & & &= 0,\\ & & \ddots & &\vdots\\ & & & x_{n-1}+x_n &= 0.\end{cases}$$

练习 1.3.14 ☕ 证明用倍加变换与倍乘变换可以实现对换变换.

练习 1.3.15 证明定理 1.3.7 的第二部分.

练习 1.3.16 练习 1.2.8 中的数独矩阵经历哪些行变换或列变换后还是数独矩阵?

练习 1.3.17 (初等列变换在线性方程组上的含义) 考虑线性方程组 $\begin{bmatrix}1&2\\4&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3\\6\end{bmatrix}$. 进行下列换元, 写出 x', y' 满足的方程组. 对比原方程组, 其系数矩阵做了哪种初等变换?

1. $x' = y, y' = x$.　　2. $x' = 2x, y' = y$.　　3. $x' = x, y' = x + y$.　　4. $x' = x + 1, y' = y$.

练习 1.3.18 (方程、法向量与超平面) 给定原点 O, 则空间中的点 P 与向量 $\overrightarrow{OP}$ 一一对应. 空间中所有点构成的集合称为**点空间**, 上述一一对应是点空间到线性空间 $\mathbb{R}^3$ 的双射. 注意, 点空间与线性空间并不相同. 特别地, 空间中的点 P 也可用向量 $\overrightarrow{OP} \in \mathbb{R}^3$ 表示.

两个 $\mathbb{R}^3$ 中的向量垂直当且仅当其内积为零 (练习 1.1.10). 与空间中某个平面垂直的非零向量称为该平面的**法向量**. 考虑空间中以非零向量 $\begin{bmatrix}a\\b\\c\end{bmatrix}$ 为法向量的平面.

1. 给定 $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^3$, 分别对应该平面上的两个点, 证明, $\begin{bmatrix}a&b&c\end{bmatrix}\boldsymbol{p} = \begin{bmatrix}a&b&c\end{bmatrix}\boldsymbol{q}$;
2. 设该平面经过对应于 $\boldsymbol{p} \in \mathbb{R}^3$ 的点, 令 $d = \begin{bmatrix}a&b&c\end{bmatrix}\boldsymbol{p}$, 证明, 平面是方程 $ax + by + cz = d$ 的解集.
3. 设 $a_1x + b_1y + c_1z = d_1$ 和 $a_2x + b_2y + c_2z = d_2$ 的解集平面平行, 试分析两个方程的系数间的关系.

4. 设 $\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}, \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$ 为两个不共线的非零向量，则 $a_1x+b_1y+c_1z=d_1$ 与 $a_2x+b_2y+c_2z=d_2$ 的解集平面不平行，因此必相交于一条直线 l. 对方程做倍加变换，证明，$a_1x+b_1y+c_1z=d_1$ 与 $(a_1+a_2)x+(b_1+b_2)y+(c_1+c_2)z=d_1+d_2$ 的解集平面的交集还是直线 l.（从几何上看，倍加变换将一个平面沿相交直线旋转，因此不改变解集的交集.）

一个 m 元线性方程的解集是对应于 $\mathbb{R}^m$ 的点空间中的一个**超平面**，因此求解线性方程组等价于求若干超平面的交集.

练习 1.3.19 对下列问题，将其化成 $A\boldsymbol{x}=\boldsymbol{b}$，并找到所有解：

1. 考虑例 1.1.1 中的桥墩载荷问题. 假设重物的重力分别为 $F_1=2, F_2=3$，放置的位置 l_1, l_2 为未知变量，桥梁长度为 5. 如果两个桥墩的载荷分别为 $f_1=3, f_2=2$，求所有可能的 $\begin{bmatrix} l_1 \\ l_2 \end{bmatrix}$.
2. 笼中有若干鸡兔，每只鸡有一个头两条腿两只翅膀，每只兔有一个头四条腿没有翅膀. 笼中现有 4 个头、12 条腿、8 只翅膀，求鸡兔数目.
3. 设 $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 满足 $A\begin{bmatrix} 1 \\ 1 \end{bmatrix}=\begin{bmatrix} 4 \\ 8 \end{bmatrix}$，且 A 第一列的元素之和为 2，求所有可能的 A.
4. 求平面上直线 $y=2x+3, y=-x+5$ 的交点.
5. 空间中有三个平面，分别经过点 $\begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$，具有法向量 $\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix}$，求这三个平面的交点（练习 1.3.18）.
6. 空间中有一条经过原点的直线，并且与向量 $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ 垂直，求所有直线上的点.
7. 空间中有一个平面经过点 $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$，求所有与该平面垂直的向量.

练习 1.3.20 ☕ 如果线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $C\boldsymbol{x}=\boldsymbol{b}$，对任意 $\boldsymbol{b}$ 都有相同的解集，那么 $A=C$ 成立吗？

提示：取特殊的 $\boldsymbol{b}$.

练习 1.3.21 ☕ 设线性映射 $\boldsymbol{F}\colon \mathbb{R}^n \to \mathbb{R}^m$. 当 $n>m$ 时，$\boldsymbol{F}$ 是否可能是单射？当 $n<m$ 时，$\boldsymbol{F}$ 是否可能是满射？

提示：把 $\boldsymbol{F}$ 的表示矩阵化为阶梯形.

1.4 线性映射的运算

本节主要讨论有关线性映射的两种运算，即线性运算和复合.

1.4.1 线性运算

定义 1.4.1 (映射的线性运算) 设 $\boldsymbol{A},\boldsymbol{B}\colon \mathbb{R}^n\to\mathbb{R}^m$ 是两个线性映射，$k\in\mathbb{R}$，定义两个新的映射：

$$\begin{aligned}\boldsymbol{A}+\boldsymbol{B}\colon\ \mathbb{R}^n &\to \mathbb{R}^m\\ \boldsymbol{x} &\mapsto \boldsymbol{A}(\boldsymbol{x})+\boldsymbol{B}(\boldsymbol{x}),\\ k\boldsymbol{A}\colon\ \mathbb{R}^n &\to \mathbb{R}^m\\ \boldsymbol{x} &\mapsto k\boldsymbol{A}(\boldsymbol{x}).\end{aligned}$$

称 $\boldsymbol{A}+\boldsymbol{B}$ 为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和，$k\boldsymbol{A}$ 为数 k 与 $\boldsymbol{A}$ 的数乘.

注意，定义中涉及的运算是陪域 $\mathbb{R}^m$ 中的线性运算. 特别地，由于向量之间不能做乘除法，两个映射之间也没有乘除运算.

命题 1.4.2 映射 $\boldsymbol{A}+\boldsymbol{B}$ 和 $k\boldsymbol{A}$ 都是线性映射.

证 这里只验证 $\boldsymbol{A}+\boldsymbol{B}$ 的情形. 根据定义，

$$\begin{aligned}(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{x}+\boldsymbol{x}') &= \boldsymbol{A}(\boldsymbol{x}+\boldsymbol{x}')+\boldsymbol{B}(\boldsymbol{x}+\boldsymbol{x}') = \boldsymbol{A}(\boldsymbol{x})+\boldsymbol{A}(\boldsymbol{x}')+\boldsymbol{B}(\boldsymbol{x})+\boldsymbol{B}(\boldsymbol{x}')\\ &= \boldsymbol{A}(\boldsymbol{x})+\boldsymbol{B}(\boldsymbol{x})+\boldsymbol{A}(\boldsymbol{x}')+\boldsymbol{B}(\boldsymbol{x}') = (\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{x})+(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{x}'),\end{aligned}$$

$$(\boldsymbol{A}+\boldsymbol{B})(k\boldsymbol{x}) = \boldsymbol{A}(k\boldsymbol{x})+\boldsymbol{B}(k\boldsymbol{x}) = k\boldsymbol{A}(\boldsymbol{x})+k\boldsymbol{B}(\boldsymbol{x}) = k(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{x}),$$

其中，$\boldsymbol{x},\boldsymbol{x}'\in\mathbb{R}^n, k\in\mathbb{R}$. □

在 $k\boldsymbol{A}$ 中取 $k=0$，得到一个特殊的线性映射，**零映射**，也用符号 $\boldsymbol{O}\colon\mathbb{R}^n\to\mathbb{R}^m$ 表示，满足 $\boldsymbol{O}(\boldsymbol{x})=\boldsymbol{0}$，对任意 $\boldsymbol{x}\in\mathbb{R}^n$ 都成立. 与零向量在线性空间 $\mathbb{R}^n$ 中的性质类似，零映射在映射的加法之下满足：$\boldsymbol{O}+\boldsymbol{A}=\boldsymbol{A}=\boldsymbol{A}+\boldsymbol{O}$. 我们也可以定义线性映射的**减法**：$\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-\boldsymbol{B})$，其中 $-\boldsymbol{B}=(-1)\boldsymbol{B}$.

所有 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射构成一个集合，这个集合在映射加法和数乘两个运算下封闭. 这两个线性运算也满足类似于命题 1.1.5 中的八条运算法则.

由于线性映射与矩阵的对应，对矩阵也可以定义类似的线性运算.

定义 1.4.3 (矩阵的线性运算) 设 $A=\begin{bmatrix}a_{ij}\end{bmatrix}_{m\times n}, B=\begin{bmatrix}b_{ij}\end{bmatrix}_{m\times n}$ 是两个 $m\times n$ 矩阵，$k\in\mathbb{R}$，定义

$$A+B=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}+\begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{m1} & b_{m2} & \cdots & b_{mn}\end{bmatrix}=\begin{bmatrix}a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n}\\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn}\end{bmatrix};$$

$$kA = k\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix}.$$

命题 1.4.4 矩阵加法和数乘满足如下八条运算法则:

1. 加法结合律: $(A+B)+C = A+(B+C)$;
2. 加法交换律: $A+B = B+A$;
3. 零矩阵: $A+O_{m\times n} = A$, 其中 $O_{m\times n}$ 是所有元素全为 0 的矩阵, 称为 $m\times n$ **零矩阵**, 简记为 O;
4. 负矩阵: 对任意 $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{bmatrix}$, 记 $-A = \begin{bmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{1n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & -a_{m1} & \cdots & -a_{mn} \end{bmatrix}$,

   它满足 $A+(-A) = O$, 称它为 A 的**负矩阵**;
5. 单位数: $1A = A$;
6. 数乘结合律: $(kl)A = k(lA)$;
7. 数乘对数的分配律: $(k+l)A = kA+lA$;
8. 数乘对矩阵的分配律: $k(A+B) = kA+kB$.

根据负矩阵性质, 我们可以定义矩阵的**减法**: $A-B = A+(-B)$. 根据加法结合律, 我们不需要区分加法的计算次序, 因此我们可以写 $A+B+C = (A+B)+C = A+(B+C)$.

矩阵和向量的乘积对向量的线性运算满足分配律 (命题 1.2.7), 这里有如下类似结果.

命题 1.4.5 矩阵和向量的乘积对矩阵的线性运算满足分配律:

$$(k_1A_1+k_2A_2)\boldsymbol{x} = k_1(A_1\boldsymbol{x})+k_2(A_2\boldsymbol{x}).$$

例 1.4.6 (差分矩阵) $\mathbb{R}^n$ 中的向量 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ 可以看成一个数列 $x_1, x_2, \cdots, x_n$. 在数列的相关计算中, 经常需要计算相邻两项的差. 由此我们定义 $\mathbb{R}^n$ 上的一个线性变换, 称为**向前差分**:

$$\begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} & \mapsto & \begin{bmatrix} x_1 \\ x_2-x_1 \\ \vdots \\ x_n-x_{n-1} \end{bmatrix}. \end{array}$$

它的表示矩阵是

$$D_{nn} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix},$$

其中对角元素全是 1，其下方相邻的副对角线上的元素全是 -1，其余元素全是 0，这个矩阵称为**向前差分矩阵**．它因对应变换取每一项与前一项的差而得名．根据矩阵的线性运算，差分矩阵可以分解为

$$D_{nn} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} - \begin{bmatrix} & & & \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix} =: I_n - J_n.$$

而 $D_{nn}\boldsymbol{x} = I_n\boldsymbol{x} - J_n\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ \vdots \\ x_n - x_{n-1} \end{bmatrix}$.

我们给出这个线性变换在一些数列上的作用．从数列 $a_n = n^2$ 开始，做一次向前差分得到等差数列，再做一次得到常数数列（不考虑第一项），即

$$\begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \\ 25 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix};$$

从 Fibonacci 数列 $a_1 = 1, a_2 = 1, a_n = a_{n-1} + a_{n-2}$ 开始，做一次向前差分后，不考虑前面若干项，得到的还是 Fibonacci 数列，即

$$\begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 5 \\ 8 \\ 13 \\ 21 \\ 34 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 2 \\ 3 \\ 5 \\ 8 \\ 13 \end{bmatrix} \mapsto \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}.$$

类似地，也可以取每一项与后一项的差，得到的是**向后差分**和**向后差分矩阵**：

$$D_{nn}^{\mathrm{T}}=\begin{bmatrix}1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1\end{bmatrix}.$$

这里，上角标 $^{\mathrm{T}}$ 表示 D_{nn}^{T} 是由 D_{nn} 对调行和列的关系得到的. ☺

定义 1.4.7 (转置) 定义 $m\times n$ 矩阵 A 的**转置**：

$$A^{\mathrm{T}}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n}\\ \vdots & \vdots & & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{m,n-1} & a_{mn}\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}a_{11} & a_{21} & \cdots & a_{m1}\\ a_{12} & a_{22} & \cdots & a_{m2}\\ \vdots & \vdots & & \vdots\\ a_{1,n-1} & a_{2,n-1} & \cdots & a_{m,n-1}\\ a_{1n} & a_{2n} & \cdots & a_{mn}\end{bmatrix},$$

即，A^{T} 是由对调 A 的行和列得到的 $n\times m$ 矩阵.

特别地，对应分量相同的列向量与行向量互为转置. 之前我们用 $\boldsymbol{a}$ 表示列向量，而用 $\boldsymbol{a}^{\mathrm{T}}$ 表示行向量，就是为了与此保持一致.

从定义来看，显然有 $(A^{\mathrm{T}})^{\mathrm{T}}=A$. 另外

$$\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\ \boldsymbol{a}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{a}_n^{\mathrm{T}}\end{bmatrix}. \tag{1.4.1}$$

命题 1.4.8 矩阵的线性运算与转置满足如下运算法则：

$$(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}},\quad (kA)^{\mathrm{T}}=kA^{\mathrm{T}}.$$

定义 1.4.9 (对称矩阵) 如果矩阵 A 满足 $A=A^{\mathrm{T}}$，就称之为**对称矩阵**. 如果矩阵 A 满足 $A=-A^{\mathrm{T}}$，就称之为**反对称矩阵**或**斜对称矩阵**.

显然，对称矩阵和反对称矩阵一定是方阵.

1.4.2 复合

在实践中，常常需要考虑多个系统的叠加，例如

$$\text{输入}\xrightarrow{\text{系统 }1}\text{中间输出}\xrightarrow{\text{系统 }2}\text{输出}.$$

这显然对应于映射的复合. 我们先看一个简单的例子.

例 1.4.10 考虑例 1.2.9 中平面向量构成的线性空间 $\mathbb{R}^2$ 上的线性变换的复合.

1. 旋转变换的复合还是旋转变换：$\boldsymbol{R}_{\theta_1} \circ \boldsymbol{R}_{\theta_2} = \boldsymbol{R}_{\theta_1+\theta_2} = \boldsymbol{R}_{\theta_2} \circ \boldsymbol{R}_{\theta_1}$. 特别地，$\boldsymbol{R}_\theta \circ \boldsymbol{R}_{-\theta} = \boldsymbol{R}_{-\theta} \circ \boldsymbol{R}_\theta = \boldsymbol{R}_0 = \boldsymbol{I}$.
2. 反射变换与自己的复合是恒同变换：$\boldsymbol{H}_\theta \circ \boldsymbol{H}_\theta = \boldsymbol{I}$.
3. 对换变换与旋转变换的复合：两种复合 $\boldsymbol{P} \circ \boldsymbol{R}_\theta$ 和 $\boldsymbol{R}_\theta \circ \boldsymbol{P}$ 是否相等?
4. 伸缩变换的复合还是伸缩变换：$\boldsymbol{C}_{k_1} \circ \boldsymbol{C}_{k_2} = \boldsymbol{C}_{k_1 k_2} = \boldsymbol{C}_{k_2} \circ \boldsymbol{C}_{k_1}$. 特别地，投影变换与自己的复合保持不变：$\boldsymbol{C}_0 \circ \boldsymbol{C}_0 = \boldsymbol{C}_0$.
5. 错切变换的复合还是错切变换吗?
6. 我们还可以考虑不同方向的伸缩变换、错切变换的复合，留给读者思考.

注意，上述线性变换的复合都还是线性变换. ☺

一般地，多个线性映射的复合还是线性映射吗？如果是，其表示矩阵有何关系？

命题 1.4.11 设 $\boldsymbol{B}\colon \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{A}\colon \mathbb{R}^m \to \mathbb{R}^l$ 是两个线性映射，则二者的复合映射 $\boldsymbol{A} \circ \boldsymbol{B}\colon \mathbb{R}^n \to \mathbb{R}^l$ 也是一个线性映射.

证 根据定义，直接验证复合映射保持线性运算，即对任意 $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^n, k \in \mathbb{R}$,

$$(\boldsymbol{A} \circ \boldsymbol{B})(\boldsymbol{x} + \boldsymbol{x}') = \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{x} + \boldsymbol{x}')) = \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{x}) + \boldsymbol{B}(\boldsymbol{x}')) = (\boldsymbol{A} \circ \boldsymbol{B})(\boldsymbol{x}) + (\boldsymbol{A} \circ \boldsymbol{B})(\boldsymbol{x}'),$$

$$(\boldsymbol{A} \circ \boldsymbol{B})(k\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{B}(k\boldsymbol{x})) = \boldsymbol{A}(k\boldsymbol{B}(\boldsymbol{x})) = k\boldsymbol{A}(\boldsymbol{B}(\boldsymbol{x})) = k(\boldsymbol{A} \circ \boldsymbol{B})(\boldsymbol{x}).$$ □

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的表示矩阵分别为 $l \times m$ 矩阵 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{bmatrix}$ 和 $m \times n$ 矩阵 $B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix}$，那么根据定义，线性映射 $\boldsymbol{A} \circ \boldsymbol{B}$ 的表示矩阵 C 可以计算出来：

$$\begin{aligned} C &= \begin{bmatrix} \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{e}_1)) & \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{e}_2)) & \cdots & \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{e}_n)) \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{A}(\boldsymbol{b}_1) & \boldsymbol{A}(\boldsymbol{b}_2) & \cdots & \boldsymbol{A}(\boldsymbol{b}_n) \end{bmatrix} \\ &= \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix}. \end{aligned}$$

这启发我们给出如下对矩阵乘法的定义.

定义 1.4.12 (矩阵乘法) 给定 $l \times m$ 矩阵 A 和 $m \times n$ 矩阵 $B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix}$，定义二者乘积 AB 是如下 $l \times n$ 矩阵：

$$AB = A \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix} := \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix}. \tag{1.4.2}$$

特别地，对正整数 k，方阵 A 的 k 次幂定义为 $A^k = \underbrace{AA\cdots A}_{k\text{ 个}}$.

矩阵乘法是矩阵与列向量乘法（见定义 1.2.5）的一个自然推广，因为列向量是特殊的矩阵. 特别需要注意的是：

1. 两个矩阵 A, B 能够相乘，亦即乘积矩阵 AB 能够存在，必须要求 A 的列数与 B 的行数相同；此时，AB 的行数等于 A 的行数，AB 的列数等于 B 的列数. 这与映射的复合一致，而这对应着：如果映射 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 可以复合成 $\boldsymbol{A} \circ \boldsymbol{B}$，那么 $\boldsymbol{A}$ 的定义域必须等于 $\boldsymbol{B}$ 的陪域.
2. 乘积矩阵 AB 的第 k 列是 $A\boldsymbol{b}_k$，这说明，只有 B 的第 k 列 $\boldsymbol{b}_k$ 会影响 AB 的第 k 列，而 B 的其他列不会影响 AB 的第 k 列.

类似地，我们还可以得到 AB 的行向量的性质. 根据 (1.2.2) 式中矩阵与列向量的乘积，设 $A = \begin{bmatrix} \tilde{\boldsymbol{a}}_1^{\mathrm{T}} \\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}} \\ \vdots \\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}} \end{bmatrix}, B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix}$，我们得到：

$$AB = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_1 & \tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_2 & \cdots & \tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_n \\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_1 & \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_2 & \cdots & \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_n \\ \vdots & \vdots & & \vdots \\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_1 & \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_2 & \cdots & \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{a}}_1^{\mathrm{T}}B \\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}B \\ \vdots \\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}B \end{bmatrix}. \tag{1.4.3}$$

这说明, AB 的第 i 行是 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}B$, 仅由矩阵 A 的第 i 行 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}$ 和矩阵 B 决定. 具体到 $C = AB$ 的每个元素，设 $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{l \times m}, B = \begin{bmatrix} b_{jk} \end{bmatrix}_{m \times n}, C = \begin{bmatrix} c_{ik} \end{bmatrix}_{l \times n}$，我们有

$$c_{ik} = \tilde{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{b}_k = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{im}b_{mk} = \sum_{j=1}^{m} a_{ij}b_{jk}. \tag{1.4.4}$$

这里我们引入了求和号 $\sum$，其中求和号后面是一个带变动下角标 j 的项，求和号上下提示了变动下角标及其变化范围，即求和式表示变动下角标取遍变化范围对应的所有项的和. 注意，这里一共涉及三个下角标 i, j, k，求和仅仅针对变动下角标 j，而下角标 i, k 都固定不变. 因此，矩阵乘法可表示为

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} \sum\limits_{j=1}^{m} a_{1j}b_{j1} & \sum\limits_{j=1}^{m} a_{1j}b_{j2} & \cdots & \sum\limits_{j=1}^{m} a_{1j}b_{jn} \\ \sum\limits_{j=1}^{m} a_{2j}b_{j1} & \sum\limits_{j=1}^{m} a_{2j}b_{j2} & \cdots & \sum\limits_{j=1}^{m} a_{2j}b_{jn} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{j=1}^{m} a_{lj}b_{j1} & \sum\limits_{j=1}^{m} a_{lj}b_{j2} & \cdots & \sum\limits_{j=1}^{m} a_{lj}b_{jn} \end{bmatrix}.$$

注意，(1.4.2) 式、(1.4.3) 式和 (1.4.4) 式都可以作为矩阵乘法的定义，因为三者互相等价.

矩阵乘积的公式在计算中必不可少，而它和映射复合之间的对应关系，对我们的理解十分重要．根据定义 1.4.12，可以直接得出下面的结果．

命题 1.4.13 设 $\boldsymbol{A}\colon \mathbb{R}^m \to \mathbb{R}^l, \boldsymbol{B}\colon \mathbb{R}^n \to \mathbb{R}^m$ 是两个线性映射，记二者的表示矩阵分别为 A 和 B，则复合映射 $\boldsymbol{A}\circ\boldsymbol{B}\colon \mathbb{R}^n \to \mathbb{R}^l$ 的表示矩阵为 $C = AB$.

一个直接的推论是，对任意列向量 $\boldsymbol{x}\in\mathbb{R}^n$，都有

$$(AB)\boldsymbol{x} = (\boldsymbol{A}\circ\boldsymbol{B})(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{x})) = \boldsymbol{A}(B\boldsymbol{x}) = A(B\boldsymbol{x}), \tag{1.4.5}$$

即矩阵与列向量乘积和矩阵乘积有结合律．更一般地，我们有如下重要的运算法则．

命题 1.4.14 矩阵乘法满足结合律：$A(BC) = (AB)C$.

证 1 令 $C = \begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n\end{bmatrix}$，则

$$\begin{aligned}
A(BC) &= A\left(B\begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n\end{bmatrix}\right)\\
&= A\begin{bmatrix}B\boldsymbol{c}_1 & B\boldsymbol{c}_2 & \cdots & B\boldsymbol{c}_n\end{bmatrix}\\
&= \begin{bmatrix}A(B\boldsymbol{c}_1) & A(B\boldsymbol{c}_2) & \cdots & A(B\boldsymbol{c}_n)\end{bmatrix} \qquad (\text{由 (1.4.5) 式})\\
&= \begin{bmatrix}(AB)\boldsymbol{c}_1 & (AB)\boldsymbol{c}_2 & \cdots & (AB)\boldsymbol{c}_n\end{bmatrix}\\
&= (AB)\begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n\end{bmatrix} = (AB)C.
\end{aligned}$$

□

证 2 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 是三个线性映射，其表示矩阵分别为 A,B,C，则复合映射 $\boldsymbol{A}\circ\boldsymbol{B}$ 和 $\boldsymbol{B}\circ\boldsymbol{C}$ 的表示矩阵分别为 AB 和 BC．因此，$(\boldsymbol{A}\circ\boldsymbol{B})\circ\boldsymbol{C}$ 和 $\boldsymbol{A}\circ(\boldsymbol{B}\circ\boldsymbol{C})$ 的表示矩阵分别为 $(AB)C$ 和 $A(BC)$．映射的复合具有结合律 $(\boldsymbol{A}\circ\boldsymbol{B})\circ\boldsymbol{C} = \boldsymbol{A}\circ(\boldsymbol{B}\circ\boldsymbol{C})$，因此表示矩阵相等，即 $(AB)C = A(BC)$. □

证 3 还可以根据 (1.4.4) 式直接计算. 设 $A = \begin{bmatrix}a_{ij}\end{bmatrix}_{l\times m}, B = \begin{bmatrix}b_{jk}\end{bmatrix}_{m\times n}, C = \begin{bmatrix}c_{ks}\end{bmatrix}_{n\times p}$，则 $(AB)C$ 的 (i,s) 元是 $\displaystyle\sum_{k=1}^{n}\left(\sum_{j=1}^{m} a_{ij}b_{jk}\right)c_{ks}$，而 $A(BC)$ 的 (i,s) 元是 $\displaystyle\sum_{j=1}^{m} a_{ij}\left(\sum_{k=1}^{n} b_{jk}c_{ks}\right)$. 根据数的乘法结合律和求和顺序的交换性，二者都等于 $\displaystyle\sum_{j=1}^{m}\sum_{k=1}^{n} a_{ij}b_{jk}c_{ks}$. □

根据命题 1.4.14 的乘法结合律，我们不需要区分矩阵乘法的计算次序，因此可以写 $ABC = (AB)C = A(BC)$.

矩阵乘法与数的乘法最重要的区别在于，乘法交换律和消去律一般不成立：

1. AB, BA 不一定同时存在；即便都存在，仍然需要 A, B 是同阶方阵，才能保证 AB, BA 是同阶方阵.

2. 即使 A, B, AB, BA 都是同阶的方阵，也存在 A, B 使得 $AB \neq BA$，例如

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

3. 两个非零矩阵的乘积**可能**是零矩阵．例如

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

因此 $AB = O$ **不能**推出 $A = O$ 或 $B = O$．更一般地，消去律也**不一定**成立：$AB = AC, A \neq O$ 不能直接推出 $B = C$．

由于交换律一般不成立，矩阵乘法的左右顺序十分重要．乘积 AB，我们常读作 “A（左）乘 B” 或者 “B 右乘 A”．

如果 $AB = BA$，我们称 A, B **可交换**．此时 A, B 一定是同阶方阵．

注意，矩阵乘法的特殊之处，和映射复合的性质类似．例如，两个一般的映射不一定可以复合；即便是同一集合上的两个变换，两种不同顺序的复合得到的变换也不一定相同．类似地，如果两个映射 $\boldsymbol{A}, \boldsymbol{B}$ 满足 $\boldsymbol{B} \circ \boldsymbol{A} = \boldsymbol{A} \circ \boldsymbol{B}$，则称 $\boldsymbol{A}, \boldsymbol{B}$ **可交换**．

例 1.4.15 考虑例 1.4.10 中线性变换的复合对应的矩阵乘法．

1. 旋转变换的复合还是旋转变换：$\boldsymbol{R}_{\theta_1} \circ \boldsymbol{R}_{\theta_2} = \boldsymbol{R}_{\theta_1+\theta_2} = \boldsymbol{R}_{\theta_2} \circ \boldsymbol{R}_{\theta_1}$．对应的矩阵乘法为

$$\begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix}.$$

2. 反射变换与自己的复合是恒同变换：$\boldsymbol{H}_\theta \circ \boldsymbol{H}_\theta = \boldsymbol{I}$．对应的矩阵乘法为

$$\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

3. 对换变换与旋转变换的复合：先计算 $\boldsymbol{P} \circ \boldsymbol{R}_\theta$ 和 $\boldsymbol{R}_\theta \circ \boldsymbol{P}$．我们直接计算对应的矩阵乘法，有

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix},$$

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix}.$$

显然二者不相等．事实上，二者的关系是正弦的系数互为相反数而余弦的系数相等，而这恰好是 θ 换成 $-\theta$ 的效果．因此，我们有 $\boldsymbol{P} \circ \boldsymbol{R}_\theta = \boldsymbol{R}_{-\theta} \circ \boldsymbol{P}$．从

几何直观上看，旋转再翻折，等价于翻折再反向旋转．类似地，我们可以验证，反射变换与旋转变换之间满足 $\boldsymbol{H}_{\theta_2} \circ \boldsymbol{R}_{\theta_1} = \boldsymbol{R}_{-\theta_1} \circ \boldsymbol{H}_{\theta_2}$.

4. 伸缩变换的复合还是伸缩变换：$\boldsymbol{C}_{k_1} \circ \boldsymbol{C}_{k_2} = \boldsymbol{C}_{k_1 k_2} = \boldsymbol{C}_{k_2} \circ \boldsymbol{C}_{k_1}$，对应的矩阵乘法为
$$\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} k_1 k_2 & 0 \\ 0 & 1 \end{bmatrix}.$$
5. 错切变换的复合还是错切变换：$\boldsymbol{S}_{k_1} \circ \boldsymbol{S}_{k_2} = \boldsymbol{S}_{k_1+k_2} = \boldsymbol{S}_{k_2} \circ \boldsymbol{S}_{k_1}$. 对应的矩阵乘法为
$$\begin{bmatrix} 1 & 0 \\ k_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ k_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ k_1 + k_2 & 1 \end{bmatrix}.$$

需要指出的是，一般情形下，A 与 B 不可交换；而这里经常可交换的原因是矩阵具有特殊性. ☺

我们还关心映射的线性运算与复合之间的关系，以及矩阵的线性运算与乘法之间的关系.

命题 1.4.16 设 $\boldsymbol{B}, \boldsymbol{B}' \colon \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{A}, \boldsymbol{A}' \colon \mathbb{R}^m \to \mathbb{R}^l$ 是线性映射，$k \in \mathbb{R}$，则有如下运算法则：

1. 对加法的分配律：$(\boldsymbol{A}+\boldsymbol{A}')\circ\boldsymbol{B} = \boldsymbol{A}\circ\boldsymbol{B}+\boldsymbol{A}'\circ\boldsymbol{B}, \boldsymbol{A}\circ(\boldsymbol{B}+\boldsymbol{B}') = \boldsymbol{A}\circ\boldsymbol{B}+\boldsymbol{A}\circ\boldsymbol{B}'$;
2. 对数乘的交换律：$k(\boldsymbol{A} \circ \boldsymbol{B}) = (k\boldsymbol{A}) \circ \boldsymbol{B} = \boldsymbol{A} \circ (k\boldsymbol{B})$.

命题 1.4.17 矩阵的乘法还满足如下运算法则：

1. 对加法的分配律：$(A+B)C = AC + BC, A(B+C) = AB + AC$;
2. 对数乘的交换律：$k(AB) = (kA)B = A(kB)$.

上面两个命题的验证是直接的，留给读者．特别地，kI_n 称为**数量矩阵**，因为 kI_n 在矩阵乘法下的作用和数乘一样：$(kI_n)A = k(I_nA) = kA$.

矩阵的乘法和转置之间，有如下关系.

命题 1.4.18 设 A, B 分别是 $l \times m$ 矩阵和 $m \times n$ 矩阵，$\boldsymbol{a}, \boldsymbol{x}$ 是 m 维列向量，则

1. $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}$;
2. $(A\boldsymbol{x})^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}$;
3. $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

证

1. 设 $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$，计算可得
$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = a_1x_1 + a_2x_2 + \cdots + a_mx_m = x_1a_1 + x_2a_2 + \cdots + x_ma_m = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}.$$

2. 设 $A=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\ \vdots\\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\end{bmatrix}$，则 $A\boldsymbol{x}=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{x}\\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{x}\\ \vdots\\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{x}\end{bmatrix}$，其中，$\tilde{\boldsymbol{a}}_i\,(i=1,2,\cdots,l),\boldsymbol{x}$ 都是 m 维列向量，满足 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\tilde{\boldsymbol{a}}_i$. 根据 (1.4.1) 式，我们有

$$(A\boldsymbol{x})^{\mathrm{T}}=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{x}\\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{x}\\ \vdots\\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{x}\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{x}^{\mathrm{T}}\tilde{\boldsymbol{a}}_1 & \boldsymbol{x}^{\mathrm{T}}\tilde{\boldsymbol{a}}_2 & \cdots & \boldsymbol{x}^{\mathrm{T}}\tilde{\boldsymbol{a}}_l\end{bmatrix}=\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}\tilde{\boldsymbol{a}}_1 & \tilde{\boldsymbol{a}}_2 & \cdots & \tilde{\boldsymbol{a}}_l\end{bmatrix}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}.$$

3. 设 $B=\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n\end{bmatrix}$，则

$$(AB)^{\mathrm{T}}=\begin{bmatrix}A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}(A\boldsymbol{b}_1)^{\mathrm{T}}\\ (A\boldsymbol{b}_2)^{\mathrm{T}}\\ \vdots\\ (A\boldsymbol{b}_n)^{\mathrm{T}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{b}_1^{\mathrm{T}}A^{\mathrm{T}}\\ \boldsymbol{b}_2^{\mathrm{T}}A^{\mathrm{T}}\\ \vdots\\ \boldsymbol{b}_n^{\mathrm{T}}A^{\mathrm{T}}\end{bmatrix}=B^{\mathrm{T}}A^{\mathrm{T}}. \qquad \square$$

命题 1.4.14 和命题 1.4.18 的证明，都把矩阵写成行向量组或者列向量组的形式来进行计算，请读者仔细体会.

例 1.4.19 (弹簧振子) 读者在中学物理中学习过弹簧振子. 所谓弹簧振子，就是不考虑摩擦阻力、不考虑弹簧质量、不考虑振子的大小和形状的理想化模型. 假设三根弹簧和三个振子如图 1.4.1 所示依次交替地挂在天花板下方，弹簧的原始长度分别为 l_1,l_2,l_3，劲度系数分别为 k_1,k_2,k_3，振子的质量分别为 m_1,m_2,m_3. 系统处于静止状态，问三个振子在哪里？

这一问题并不难. 我们先求三根弹簧的拉伸长度，设每根弹簧的拉伸长度分别是 u_1，u_2，u_3. 对振子 m_3 受力分析，利用 Hooke 定律，易得 $k_3u_3=m_3g$，于是 $u_3=g\dfrac{m_3}{k_3}$.

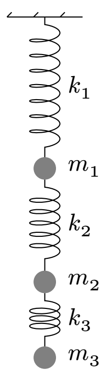


图 1.4.1 弹簧振子：自由系统

对振子 m_2，m_3 整体受力分析，易得 $k_2u_2=m_2g+m_3g$，于是 $u_2=g\dfrac{m_2+m_3}{k_2}$. 对振子 m_1,m_2,m_3 整体受力分析，易得 $k_1u_1=m_1g+m_2g+m_3g$，于是 $u_1=g\dfrac{m_1+m_2+m_3}{k_1}$. 这就能得到 m_1 到天花板的距离是 $u_1+l_1=g\dfrac{m_1+m_2+m_3}{k_1}+l_1$，$m_2$ 到天花板的距离是 $u_1+l_1+u_2+l_2=g\dfrac{m_1+m_2+m_3}{k_1}+l_1+g\dfrac{m_2+m_3}{k_2}+l_2$，$m_3$ 到天花板的距离是 $u_1+l_1+u_2+l_2+u_3+l_3=g\dfrac{m_1+m_2+m_3}{k_1}+l_1+g\dfrac{m_2+m_3}{k_2}+l_2+g\dfrac{m_3}{k_3}+l_3$.

可以看到，如果我们增加弹簧和振子的数目，解将越来越复杂，虽然很有规律，但是并不容易用简单的公式表示.

我们用线性映射和矩阵的框架来分析．首先我们可以看到，m_1 离天花板至少有 l_1 那么远，m_2 离天花板至少有 l_1+l_2 那么远，m_3 离天花板至少有 $l_1+l_2+l_3$ 那么远．因此我们没有必要去计算振子到天花板的距离．我们假设 m_1 到天花板的距离是 l_1+x_1，m_2 的是 $l_1+l_2+x_2$，m_3 的是 $l_1+l_2+l_3+x_3$.

把三个振子的位移 x_1,x_2,x_3 看作系统的输入，它们所受的拉力 f_1,f_2,f_3 看作系统的输出．这个系统可以描述成以下三个系统的叠加：

$$\text{输入} = \text{振子的位移} \to \text{弹簧的拉伸长度} \to \text{弹簧的拉力} \to \text{振子所受的拉力} = \text{输出}.$$

我们把输入写成向量 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$．那么弹簧的拉伸长度

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\end{bmatrix}=\begin{bmatrix}1&&\\-1&1&\\&-1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=D_{33}\boldsymbol{x}.$$

从 $\boldsymbol{x}$ 到 $\boldsymbol{u}$ 构成了一个线性映射，其表示矩阵是向前差分矩阵 D_{33}.

下一步，三根弹簧提供的拉力就是

$$\boldsymbol{y}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}k_1u_1\\k_2u_2\\k_3u_3\end{bmatrix}=\begin{bmatrix}k_1&&\\&k_2&\\&&k_3\end{bmatrix}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=K_3\boldsymbol{u}.$$

从 $\boldsymbol{u}$ 到 $\boldsymbol{y}$ 构成了一个线性映射，其表示矩阵为对角矩阵 K_3.

最后，振子所受拉力为

$$\boldsymbol{f}=\begin{bmatrix}y_1-y_2\\y_2-y_3\\y_3\end{bmatrix}=\begin{bmatrix}1&-1&\\&1&-1\\&&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=D_{33}^{\mathrm{T}}\boldsymbol{y}.$$

从 $\boldsymbol{y}$ 到 $\boldsymbol{f}$ 构成了一个线性映射，其表示矩阵为向后差分矩阵 D_{33}^{T}.

记振子的质量为 $\boldsymbol{m}=\begin{bmatrix}m_1\\m_2\\m_3\end{bmatrix}$，则振子所受重力为 $\boldsymbol{g}=g\boldsymbol{m}=\begin{bmatrix}m_1g\\m_2g\\m_3g\end{bmatrix}$．利用力的平衡有 $\boldsymbol{f}=\boldsymbol{g}$．注意 $D_{33},K_3,D_{33}^{\mathrm{T}}$ 都是对角元素都不为零的三角矩阵，因此，利用前代法和回代法，我们可以逐步倒推，由 $\boldsymbol{f}$ 反解出 $\boldsymbol{x}$.

系统对应的映射可以分解为三个映射的复合，其矩阵分别为 $D_{33}^{\mathrm{T}}, K_3, D_{33}$，因此复合映射的矩阵为

$$D_{33}^{\mathrm{T}}K_3D_{33} = \begin{bmatrix} 1 & -1 & \\ & 1 & -1 \\ & & 1 \end{bmatrix} \begin{bmatrix} k_1 & & \\ & k_2 & \\ & & k_3 \end{bmatrix} \begin{bmatrix} 1 & & \\ -1 & 1 & \\ & -1 & 1 \end{bmatrix} = \begin{bmatrix} k_1+k_2 & -k_2 & \\ -k_2 & k_2+k_3 & -k_3 \\ & -k_3 & k_3 \end{bmatrix},$$

这是一个对称矩阵. 可以看到，从映射和矩阵的观点考虑问题，不仅可以很容易地把所列方程推广到任意数目的弹簧和振子上，而且能够更清楚地表示物理量之间的关系. K_3 反映了弹簧的性质，$\boldsymbol{m}$ 反映了振子的性质. 因此，当我们对向量和矩阵熟练之后，很容易就可以直接列出方程 $g\boldsymbol{m} = D_{33}^{\mathrm{T}}K_3D_{33}\boldsymbol{x}$，而不再需要通过分量的计算，分量的数目也就变得不那么重要了. ☺

习题

练习 1.4.1 设 A, B, C 分别是 $3\times5, 5\times3, 3\times1$ 矩阵，则 $BA, AB, BC^{\mathrm{T}}, A(B+C)$ 中哪些定义良好?

练习 1.4.2 设 $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & 2 & -2 \end{bmatrix}$，求 $AB, BA, AB-BA$.

练习 1.4.3 计算:

1. $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -7 & 3 \\ 5 & -2 \end{bmatrix}$.
2. $\begin{bmatrix} -7 & 3 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$.
3. $\begin{bmatrix} 17 & -6 \\ 35 & -12 \end{bmatrix}^5$.
4. $\begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 2 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & -1 \\ 1 & 1 & -1 \\ -2 & -5 & 4 \end{bmatrix}$.
5. $\begin{bmatrix} 8 & 6 & -6 \\ 4 & 6 & -4 \\ 8 & 8 & -6 \end{bmatrix}^6$.

练习 1.4.4 考虑下列 $\mathbb{R}^2$ 上的线性变换.

1. 设 $\boldsymbol{R}_\theta(\boldsymbol{x}) = R_\theta\boldsymbol{x}$，其中 $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 求证:
   (a) $\boldsymbol{R}_\theta$ 是绕原点逆时针方向旋转角度 θ 的变换.
   (b) 分析当 θ 取何值时，存在常数 λ 和非零向量 $\boldsymbol{x}$，满足 $R_\theta\boldsymbol{x} = \lambda\boldsymbol{x}$.
   (c) 计算 $R_\theta^n, n>0, n\in\mathbb{N}$.
2. 设 $\boldsymbol{H}_\theta(\boldsymbol{x}) = H_\theta\boldsymbol{x}$，其中 $H_\theta = \begin{bmatrix} \cos2\theta & \sin2\theta \\ \sin2\theta & -\cos2\theta \end{bmatrix}$，而 $\boldsymbol{v} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$ 是 $\mathbb{R}^2$ 中的向量. 求证:
   (a) $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = 0, \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} = 1; H_\theta\boldsymbol{v} = \boldsymbol{v}, H_\theta\boldsymbol{w} = -\boldsymbol{w}$. 试分析变换 $\boldsymbol{H}_\theta$ 的几何意义.
   (b) $H_\theta^2 = I_2, H_\theta = I_2 - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}$.
   (c) $R_{-\theta}H_\phi R_\theta = H_{\phi-\theta}, H_\phi R_\theta H_\phi = R_{-\theta}$，并分析其几何意义.

3. 设 $\boldsymbol{S}(\boldsymbol{x}) = S\boldsymbol{x}$，其中 $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. 设 λ 是常数，求证：$S\boldsymbol{x} = \lambda\boldsymbol{x}$ 有非零解当且仅当 $\lambda = 1$，并求出所有的非零解；计算 $S^n, n > 0, n \in \mathbb{N}$.

练习 1.4.5 计算：

1. $\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}^2$.

2. $\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}^k$.

练习 1.4.6 设 E_{ij} 是 (i,j) 元为 1，其余元素为 0 的矩阵，而 n 阶方阵 $J_n = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$.

1. 计算 $A = I_2 + J_2, B = I_2 + J_2^{\mathrm{T}}, C = J_2 - J_2^{\mathrm{T}}$.
2. 先计算 AB，再计算 $(AB)C$，然后先计算 BC，再计算 $A(BC)$.
3. 先计算 AB, AC，再计算 $AB + AC$，然后先计算 $B + C$，再计算 $A(B + C)$.
4. 计算 $AB + BC, B(A + C), (A + C)B$，三者是否相等?
5. 计算 $(A + B)^2, A^2 + 2AB + B^2, A^2 + AB + BA + B^2$，三者是否相等?
6. 计算 $(AB)^2$ 与 A^2B^2，二者是否相等?
7. 计算 $E_{13} - \boldsymbol{e}_1\boldsymbol{e}_3^{\mathrm{T}}$，$\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}} + \cdots + \boldsymbol{e}_4\boldsymbol{e}_4^{\mathrm{T}}$，$E_{13}E_{32}$，$E_{32}E_{13}$，其中 $\boldsymbol{e}_i \in \mathbb{R}^4$，$E_{ij}$ 是 4 阶方阵.
8. 计算 $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$.
9. 计算 $J_4 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^{\mathrm{T}} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^2 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$，$J_4^4$.
10. 计算 J_n^k，其中 k 是正整数.
11. 设 $P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$，计算 $PJ_4 + J_4P, (P + I_4)(2P + 3I_4) - (2P + 3I_4)(P + I_4)$.
12. 设 $R = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, S = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, T = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix}$，计算 $SRT, SR^{\mathrm{T}}T$.

练习 1.4.7 设二阶矩阵 B 的 $(1,2)$ 元增加 1，对下列矩阵讨论该矩阵的哪行哪列一定不变，并举例说明所有其他行列确实可以变化：

1. $A + B$.
2. AB.
3. BA.
4. B^2.

练习 1.4.8 判断对错：

1. 如果 B 的第一列等于第三列，则 AB 的第一列等于第三列.
2. 如果 B 的第一列等于第三列，则 BA 的第一列等于第三列.
3. 如果 A 的第一行等于第三行，则 ABC 的第一行等于第三行.

练习 1.4.9 求所有满足条件的矩阵 B：

1. 对任意三阶方阵 A，$BA = 4A$.
2. 对任意三阶方阵 A，$BA = 4B$.
3. 对任意三阶方阵 A，BA 的每一行都是 A 的第一行.
4. 对任意三阶方阵 A，AB 的每一行的每一个元素都是 A 的对应行的平均值.
5. $B^2 \neq O, B^3 = O$，且 B 是三阶上三角矩阵.
6. $B\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}B$.

练习 1.4.10 给定矩阵 $A = \begin{bmatrix}p & 0\\q & r\end{bmatrix}, B = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}, C = \begin{bmatrix}0 & z\\0 & 0\end{bmatrix}$.

1. p, q, r 取何值时，有 $AB = BA$?
2. z 取何值时，有 $BC = CB$?
3. p, q, r, z 取何值时，有 $ABC = CAB$?

练习 1.4.11 证明：

1. 设 n 维向量 $\boldsymbol{x}$ 的每个分量都是 1，则 n 阶方阵 A 的各行元素之和为 1 当且仅当 $A\boldsymbol{x} = \boldsymbol{x}$.
2. 若 n 阶方阵 A, B 的各行元素之和均为 1，则 AB 的各行元素之和也均为 1.
3. 若 n 阶方阵 A, B 的各列元素之和均为 1，则 AB 的各列元素之和也均为 1.

练习 1.4.12 证明命题 1.4.17.

练习 1.4.13 证明上三角矩阵对加法、数乘、乘法封闭，即：设 U_1, U_2 是 n 阶上三角矩阵，k 是实数，则 $U_1 + U_2, kU_1, U_1U_2$ 都是上三角矩阵. 此外，U_1U_2 的对角元素是 U_1, U_2 对应的对角元素的乘积.

练习 1.4.14 设 A 是 n 阶上三角矩阵. 求证：$A^n = O$ 当且仅当 A 是严格上三角矩阵.

练习 1.4.15 证明：如果 n 阶方阵 A 满足 $A^n = O$，则 $(I_n - A)(I_n + A + A^2 + \cdots + A^{n-1}) = I_n$.

练习 1.4.16 对 n 阶方阵 A，考虑集合 $\mathrm{Comm}(A) = \{n \text{ 阶方阵 } B \mid AB = BA\}$.

1. 求证：$\mathrm{Comm}(A)$ 是所有 n 阶方阵的集合当且仅当 A 是数量矩阵 kI_n.
2. 设 $A = \mathrm{diag}(d_i)$，d_i 互不相同，求 $\mathrm{Comm}(A)$.
3. 求证：任取 $B, C \in \mathrm{Comm}(A)$，都有 $I_n, kB + lC, BC \in \mathrm{Comm}(A)$.
4. 设 $A = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix}$，求证：$J_3 = \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix} \in \mathrm{Comm}(A)$，而且

$$\mathrm{Comm}(A) = \{k_1I_3 + k_2J_3 + k_3J_3^2 \mid k_1, k_2, k_3 \in \mathbb{R}\}.$$

练习 1.4.17 设 $A=\begin{bmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{bmatrix}$，求 A^k，其中 k 是正整数.

练习 1.4.18 设 A 是 n 阶方阵，求证：$A+A^{\mathrm{T}}, AA^{\mathrm{T}}, A^{\mathrm{T}}A$ 都是对称矩阵，而 $A-A^{\mathrm{T}}$ 是反对称矩阵.

练习 1.4.19 求证：

1. 任意方阵 A 都可唯一地表示为 $A=B+C$，其中 B 是对称矩阵，C 是反对称矩阵.
2. n 阶方阵 A 是反对称矩阵当且仅当对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=0$.
3. 设 A,B 是 n 阶对称矩阵，则 $A=B$ 当且仅当对任意 n 维列向量 $\boldsymbol{x}$，都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}$.

练习 1.4.20 设 A,B 是同阶对称矩阵. 求证：AB 是对称矩阵当且仅当 $AB=BA$.

练习 1.4.21 设 A 是实对称矩阵，如果 $A^2=O$，求证：$A=O$.

练习 1.4.22 (矩阵的迹) 方阵 A 的对角元素的和称为它的**迹**，记作 $\operatorname{trace}(A)$. 验证下列性质：

1. 对任意同阶方阵 A,B，$\operatorname{trace}(A+B)=\operatorname{trace}(A)+\operatorname{trace}(B)$.
2. 对任意方阵 A 与实数 k，$\operatorname{trace}(kA)=k\operatorname{trace}(A)$.
3. 对 m 阶单位矩阵 I_m，$\operatorname{trace}(I_m)=m$.
4. 对任意方阵 A，$\operatorname{trace}(A)=\operatorname{trace}(A^{\mathrm{T}})$.
5. 设 $A=\begin{bmatrix}a_1 & a_2\\ a_3 & a_4\end{bmatrix}, B=\begin{bmatrix}b_1 & b_2\\ b_3 & b_4\end{bmatrix}$，则 $\operatorname{trace}(A^{\mathrm{T}}B)=\sum_{i=1}^{4}a_ib_i$. 如果 A 是 m 阶方阵呢？
6. 设 $\boldsymbol{v},\boldsymbol{w}$ 是 m 维向量，则 $\operatorname{trace}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w})=\operatorname{trace}(\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})$.
7. 设 A,B 分别是 $m\times n, n\times m$ 矩阵，则 $\operatorname{trace}(AB)=\operatorname{trace}(BA)$.
8. 设 A,B 是任意 m 阶方阵，则 $AB-BA\neq I_m$.

练习 1.4.23 (差分矩阵与求导) 在微积分中，$f(x)$ 的导数 $f'(x)\approx\dfrac{f(x+\delta)-f(x)}{\delta}$. 对数列 $\{a_n\}$ 的相邻两项做减法，可以看作是一种离散导数 $\dfrac{a_{n+1}-a_n}{1}$. 两者具有某种共性. 令 D 为 100 阶向前差分矩阵，而 $\boldsymbol{a}=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_{100}\end{bmatrix}$. 求证：

1. 若 a_k 是关于 k 的 3 次多项式，则除第 1 个分量外，$D\boldsymbol{a}$ 的第 k 个分量是关于 k 的 2 次多项式.
2. 若 $a_k=\mathrm{e}^k$，则除第 1 个分量外，$\dfrac{\mathrm{e}}{\mathrm{e}-1}D\boldsymbol{a}$ 的第 k 个分量是 e^k.

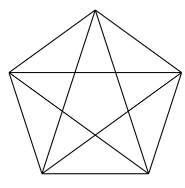


图 1.4.2

练习 1.4.24 如图 1.4.2 所示的电路包含 5 个顶点，电势分别为 $v_i, i=1,2,\cdots,5$，i,j 两点之间电阻为 $r_{ij}\neq 0$，从顶点 i 到 j 的电流为 c_{ij}，又记 $c_{ii}=0$.令电势矩阵为 $V=\operatorname{diag}(v_1,v_2,\cdots,v_5)$，电流矩阵为 $C=\left[c_{ij}\right]$，电导矩阵为 $G=\left[g_{ij}\right]$，其中 $g_{ii}=0, g_{ij}=\dfrac{1}{r_{ij}}, i\neq j$. 求证 $C=VG-GV$.

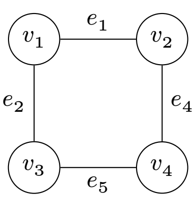


图 1.4.3

练习 1.4.25 图 1.4.3 中的图含有四个顶点v_1,v_2,v_3,v_4，顶点之间有边

连接. 对称矩阵 A 称为该图的**邻接矩阵**: 如果 v_i, v_j 之间有边, 则 A 的 (i,j) 元是 1; 如果 v_i, v_j 之间没有边, 则 A 的 (i,j) 元是 0.

1. 从 v_3 出发, 三次通过连线, 最后回到 v_3 的方法有几种?
2. 求矩阵 A^3 的 $(3,3)$ 元.
3. 分析 A^n 的 (i,j) 元的意义.

练习 1.4.26 设 $\boldsymbol{v}^{\mathrm{T}}, \boldsymbol{w}^{\mathrm{T}}$ 是 k 维行向量, A 是 $m \times n$ 矩阵.

1. 如果矩阵乘积 $\boldsymbol{v}^{\mathrm{T}} A$ 良定义, 需要满足什么条件?
2. 对任意常数 $a, b \in \mathbb{R}$, 证明 $(a\boldsymbol{v}^{\mathrm{T}} + b\boldsymbol{w}^{\mathrm{T}})A = a\boldsymbol{v}^{\mathrm{T}}A + b\boldsymbol{w}^{\mathrm{T}}A$.
3. 把 $2\boldsymbol{a}^{\mathrm{T}} + 3\boldsymbol{b}^{\mathrm{T}} + 4\boldsymbol{c}^{\mathrm{T}}$ 写成一个行向量和矩阵的乘积.

   注意: 类比于矩阵乘列向量是矩阵的列的线性组合, 行向量乘矩阵是矩阵的行的线性组合.
4. 求矩阵 A, 使得 $\begin{bmatrix} x_1 & x_2 \end{bmatrix} A \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2x_1y_1 + 3x_1y_2 + 4x_2y_1 + 5x_2y_2$.
5. 求矩阵 A, 使得 $\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4xy + 5y^2$. 这样的 A 是否唯一?
6. 求三阶对称矩阵 A, 使得对任意非零三维向量 $\boldsymbol{v}$, 都有 $\boldsymbol{v}^{\mathrm{T}} A \boldsymbol{v} > 0$.

练习 1.4.27 由标准坐标向量 $\boldsymbol{e}_i$ 定义 $E_{ij} = \boldsymbol{e}_i \boldsymbol{e}_j^{\mathrm{T}}$, 它是 (i,j) 元为 1, 其他元素为 0 的矩阵.

1. 证明, 当 $j \neq k$ 时, $E_{ij}E_{kl} = O$.
2. 对任意矩阵 A, 求向量 $\boldsymbol{v}$ 使得 $A\boldsymbol{v}$ 是 A 的第 i 列.
3. 对任意矩阵 A, 求向量 $\boldsymbol{v}$ 使得 $\boldsymbol{v}^{\mathrm{T}}A$ 是 A 的第 i 行.
4. 对任意矩阵 A, 证明, $\boldsymbol{e}_i^{\mathrm{T}} A \boldsymbol{e}_j$ 为 A 的 (i,j) 元.
5. 对 $\boldsymbol{e}_k \in \mathbb{R}^m$ 证明, $\displaystyle\sum_{k=1}^{m} \boldsymbol{e}_k \boldsymbol{e}_k^{\mathrm{T}} = I_m$.
6. (阅读) 计算矩阵乘积 AB 的 (i,j) 元的另一种方法: 设 A 有 m 列, B 有 m 行, 则
$$\boldsymbol{e}_i^{\mathrm{T}} AB \boldsymbol{e}_j = \boldsymbol{e}_i^{\mathrm{T}} A I_m B \boldsymbol{e}_j = \boldsymbol{e}_i^{\mathrm{T}} A \left(\sum_{k=1}^{m} \boldsymbol{e}_k \boldsymbol{e}_k^{\mathrm{T}} \right) B \boldsymbol{e}_j = \sum_{k=1}^{m} (\boldsymbol{e}_i^{\mathrm{T}} A \boldsymbol{e}_k)(\boldsymbol{e}_k^{\mathrm{T}} B \boldsymbol{e}_j).$$

练习 1.4.28 注意 m 维向量是 $m \times 1$ 矩阵.

1. 对 $\boldsymbol{v} \in \mathbb{R}^m, k \in \mathbb{R}$, $k\boldsymbol{v}$ 与 $\boldsymbol{v}k$ 是否可以看作矩阵乘法?
2. 对 $\boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{w} \in \mathbb{R}^n$, $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}}$ 是否良定义? 如果是, 乘积有几行几列?
3. 求 $\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 100 \end{bmatrix} \begin{bmatrix} 10 & 9 & \cdots & 1 \end{bmatrix}$ 的 $(12,7)$ 元.
4. 求 $\boldsymbol{v}, \boldsymbol{w}$, 使得 $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}} = \begin{bmatrix} (-1)^{i+j} \end{bmatrix}$.
5. 求 $\boldsymbol{v}, \boldsymbol{w}$, 使得 $\boldsymbol{v}\boldsymbol{w}^{\mathrm{T}} = \begin{bmatrix} \frac{i}{j} \end{bmatrix}$.
6. 令 $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}, B = \begin{bmatrix} \boldsymbol{b}_1^{\mathrm{T}} \\ \boldsymbol{b}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{b}_n^{\mathrm{T}} \end{bmatrix}$. 证明 $AB = \displaystyle\sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{b}_i^{\mathrm{T}}$.

1.5　矩阵的逆

1.5.1　初等矩阵

回顾 Gauss 消元法中矩阵的初等行变换. 给定矩阵 A, 对 A 的第 i,j 行做对换变换是互换 A 的第 i,j 行的位置. 首先考虑 A 是列向量的情形, 即 $A=\boldsymbol{x}\in\mathbb{R}^n$, 则对换变换定义了如下 $\mathbb{R}^n$ 上的变换:

$$\begin{aligned}\boldsymbol{P}_{ij}\colon\quad \mathbb{R}^n &\quad\to\quad \mathbb{R}^n\\ \begin{bmatrix}\vdots\\ x_i\\ \vdots\\ x_j\\ \vdots\end{bmatrix} &\quad\mapsto\quad \begin{bmatrix}\vdots\\ x_j\\ \vdots\\ x_i\\ \vdots\end{bmatrix}.\end{aligned}$$

容易验证, $\boldsymbol{P}_{ij}$ 是线性变换, 其表示矩阵为 $P_{ij}=\begin{bmatrix}\boldsymbol{P}_{ij}(\boldsymbol{e}_1) & \boldsymbol{P}_{ij}(\boldsymbol{e}_2) & \cdots & \boldsymbol{P}_{ij}(\boldsymbol{e}_n)\end{bmatrix}$. P_{ij} 可由恒同矩阵 $\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n\end{bmatrix}$ 的每一列的第 i,j 行位置互换得到, 即 P_{ij} 可由互换 I_n 的第 i,j 行位置得到. 根据表示矩阵的定义, 我们有 $\boldsymbol{P}_{ij}(\boldsymbol{x})=P_{ij}\boldsymbol{x}$, 这意味着列向量的对换行变换可以由左乘矩阵实现. 一般地, 对 $n\times p$ 矩阵 $A=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p\end{bmatrix}$, 则

$$P_{ij}A=P_{ij}\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p\end{bmatrix}=\begin{bmatrix}P_{ij}\boldsymbol{a}_1 & P_{ij}\boldsymbol{a}_2 & \cdots & P_{ij}\boldsymbol{a}_p\end{bmatrix}$$

是互换 A 的第 i,j 行的位置得到的矩阵. 换言之, 任意矩阵的对换行变换也可以由左乘矩阵实现.

容易验证, 以上讨论对其他两种初等行变换也适用. 由此给出如下定义.

定义 1.5.1 (初等矩阵)　对恒同矩阵 I_n 做一次初等变换, 得到的矩阵统称为**初等矩阵**:

1. 对换行变换: 把 I_n 的第 i,j 行位置互换, 得到**对换矩阵**

$$P_{ij}=\begin{array}{cc}\begin{bmatrix}\ddots &&&&&&&& \\ & 1 &&&&&&& \\ && 0 &&&& 1 && \\ &&& 1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& 1 &&& \\ && 1 &&&& 0 && \\ &&&&&&& 1 & \\ &&&&&&&& \ddots\end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \end{matrix} \\ \begin{matrix} \quad & \quad & i & \quad & \quad & \quad & j & \quad & \quad \end{matrix} & \end{array};$$

2. 倍乘行变换：第 i 行乘非零常数 k，得到**倍乘矩阵**

$$E_{ii;k}=\begin{array}{cl}\begin{bmatrix}\ddots&&&&\\&1&&&\\&&k&&\\&&&1&\\&&&&\ddots\end{bmatrix}&\begin{array}{l}\\ \\ i\\ \\ \\ \end{array}\\ \begin{array}{ccccc}\quad&\quad&i&\quad&\quad\end{array}&\end{array};$$

3. 倍加行变换：把第 i 行的 k 倍加到第 j 行上，得到**倍加矩阵**

$$E_{ji;k}=\begin{array}{cl}\begin{bmatrix}\ddots&&&&\\&1&&&\\&&\ddots&&\\&k&&1&\\&&&&\ddots\end{bmatrix}&\begin{array}{l}\\ i\\ \\ j\\ \\ \end{array}\\ \begin{array}{ccccc}\quad&i&\quad&j&\quad\end{array}&\end{array}(j>i),\quad E_{ji;k}=\begin{array}{cl}\begin{bmatrix}\ddots&&&&\\&1&&k&\\&&\ddots&&\\&&&1&\\&&&&\ddots\end{bmatrix}&\begin{array}{l}\\ j\\ \\ i\\ \\ \end{array}\\ \begin{array}{ccccc}\quad&j&\quad&i&\quad\end{array}&\end{array}(j<i).$$

我们还可以考虑初等列变换. 根据前面的分析，我们容易得到上述初等矩阵左（右）乘与初等行（列）变换之间的对应关系，即“左行右列”.

命题 1.5.2

1. 若矩阵 I_m 经过某一初等行变换得到矩阵 T，则任意 $m\times n$ 矩阵 A 经过相同初等行变换得到矩阵 TA.
2. 若矩阵 I_n 经过某一初等列变换得到矩阵 T，则任意 $m\times n$ 矩阵 A 经过相同初等列变换得到矩阵 AT.

因此，我们利用 Gauss 消元法或 Gauss-Jordan 消元法求解线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$，就是找到一系列初等矩阵 $E_1,E_2,\cdots,E_t$，使得矩阵 $E_t\cdots E_2E_1\begin{bmatrix}A&\boldsymbol{b}\end{bmatrix}$ 是阶梯形矩阵或行简化阶梯形矩阵. 而对任意初等矩阵 E，定理 1.3.4 说明了 $A\boldsymbol{x}=\boldsymbol{b}$ 与 $EA\boldsymbol{x}=E\boldsymbol{b}$ 同解.

我们知道，这些方法能够求解线性方程组，根本原因在于初等变换可逆. 那么初等变换诱导的线性变换，作为线性映射可逆吗？这又对应了初等矩阵的什么性质？

1.5.2 可逆矩阵

我们知道，$\mathbb{R}^n$ 上的任意两个线性变换都可以复合，得到的还是 $\mathbb{R}^n$ 上的线性变换. 其中，恒同变换 $\boldsymbol{I}$ 跟 $\mathbb{R}^n$ 上的任意线性变换 $\boldsymbol{A}$ 都可交换：$\boldsymbol{A}\circ\boldsymbol{I}=\boldsymbol{A}=\boldsymbol{I}\circ\boldsymbol{A}$. 设 $\boldsymbol{A},\boldsymbol{B}\colon\mathbb{R}^n\to\mathbb{R}^n$ 是两个线性变换，如果 $\boldsymbol{A}\circ\boldsymbol{B}=\boldsymbol{B}\circ\boldsymbol{A}=\boldsymbol{I}$，则 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是可逆变换，即 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆变换，即 $\boldsymbol{B}=\boldsymbol{A}^{-1}$；类似地，$\boldsymbol{A}$ 是 $\boldsymbol{B}$ 的逆变换，即 $\boldsymbol{A}=\boldsymbol{B}^{-1}$. 一个变换可逆当且仅当它既是单射又是满射. 细节参见 0.3 节. 注意，当 $\boldsymbol{A}$ 是可逆线性变换时，$\boldsymbol{A}^{-1}$ 必然是线性变换.

回顾例 1.4.15，$\boldsymbol{R}_\theta$，$\boldsymbol{H}_\theta$，$\boldsymbol{P}$，$\boldsymbol{C}_k(k\neq 0)$，$\boldsymbol{S}_k$ 都是可逆变换，其逆变换分别是 $\boldsymbol{R}_{-\theta}$，$\boldsymbol{H}_\theta$，$\boldsymbol{P}$，$\boldsymbol{C}_{k^{-1}}$，$\boldsymbol{S}_{-k}$. 而投影变换 $\boldsymbol{C}_0$ 不是可逆变换，因为它既不是单射，也不是满射.

考虑可逆变换的表示矩阵，自然地诱导出如下概念.

定义 1.5.3 (逆矩阵) 设 A 是 n 阶方阵, 如果存在 n 阶方阵 B, 使得 $AB = BA = I_n$, 则称 A 是**可逆矩阵**或**非奇异矩阵**, 并称 B 是 A 的一个**逆 (矩阵)**.

不可逆的矩阵称为**奇异矩阵**.

矩阵的逆如果存在，则一定唯一：设 $B, \widetilde{B}$ 都是 A 的逆，则有 $AB = I_n, \widetilde{B}A = I_n$，于是 $B = I_nB = \widetilde{B}AB = \widetilde{B}I_n = \widetilde{B}$. 因此，我们记 A 的（唯一的）逆矩阵为 A^{-1}. 另外，这一点也可以从矩阵和映射的关系得到，因为映射的逆如果存在则唯一.

根据定义，恒同矩阵 I_n 可逆，且 $I_n^{-1} = I_n$. 而零矩阵一定不可逆：对任意（可以相乘的）A, B，$AO = OB = O$. 例 1.2.9 中的大多数矩阵都可逆：

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} k^{-1} & 0 \\ 0 & 1 \end{bmatrix} (k \neq 0), \quad \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -k & 1 \end{bmatrix}.$$

命题 1.5.4 初等矩阵可逆，且 $P_{ij}^{-1} = P_{ij}, E_{ii;k}^{-1} = E_{ii;k^{-1}}, E_{ji;k}^{-1} = E_{ji;-k}$.

命题 1.5.5 对角矩阵 D 可逆当且仅当对角元素都不为零，且 D 可逆时有

$$D^{-1} = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}^{-1} = \begin{bmatrix} d_1^{-1} & & & \\ & d_2^{-1} & & \\ & & \ddots & \\ & & & d_n^{-1} \end{bmatrix}.$$

证 如果对角元素都不为零, 容易验证 $\operatorname{diag}(d_i)\operatorname{diag}(d_i^{-1}) = \operatorname{diag}(d_i^{-1})\operatorname{diag}(d_i) = I_n$.

如果对角元素至少有一个为零，设 $d_i = 0$，则 D 的第 i 行全是零. 任取 n 阶方阵 B，DB 的第 i 行也一定全是零，因此 $DB \neq I_n$，D 不可逆. □

命题 1.5.6 设 A, B 是可逆矩阵，则：

1. A^{-1} 可逆，且 $(A^{-1})^{-1} = A$;
2. A^{T} 可逆，且 $(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}$，记为 $A^{-\mathrm{T}}$;
3. AB 可逆，且 $(AB)^{-1} = B^{-1}A^{-1}$.

证 这里只证明第 3 条.

$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I,$$

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I.$$

因此 $B^{-1}A^{-1}$ 是 AB 的逆. □

下面我们就可以初步讨论矩阵可逆的条件了.

定理 1.5.7　设 A 是 n 阶方阵，以下叙述等价：

1.　A 可逆；
2.　任取 n 维向量 $\boldsymbol{b}$，线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的解唯一；
3.　齐次线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 只有零解；
4.　A 对应的阶梯形矩阵有 n 个主元；
5.　A 对应的行简化阶梯形矩阵一定是 I_n；
6.　A 是有限个初等矩阵的乘积.

证　采用轮转证法.

$1\Rightarrow 2$：考虑由 A 确定的线性映射 $\boldsymbol{A}\colon \mathbb{R}^n\to\mathbb{R}^n, \boldsymbol{x}\mapsto A\boldsymbol{x}$. 方阵 A 可逆当且仅当 $\boldsymbol{A}$ 是可逆映射. 根据命题 0.3.2，这等价于 $\boldsymbol{A}$ 既是单射又是满射. 映射 $\boldsymbol{A}$ 是满射等价于 $A\boldsymbol{x}=\boldsymbol{b}$ 总有解，而 $\boldsymbol{A}$ 是单射等价于解唯一.

$2\Rightarrow 3$：显然.

$3\Rightarrow 4$：求解此方程组，先用初等行变换将 A 化为阶梯形矩阵，方程组有唯一解等价于阶梯形矩阵的阶梯数与未知数的个数 n 相等，故主元数等于阶梯数 n.

$4\Rightarrow 5$：显然.

$5\Rightarrow 6$：由假设，我们可以用一系列初等行变换把 A 化成 I_n，即存在初等矩阵 $E_1,E_2,\cdots,E_t$ 使得 $E_t\cdots E_1A=I_n$. 注意到初等矩阵都可逆，因此 $A=E_1^{-1}\cdots E_t^{-1}$.

$6\Rightarrow 1$：初等矩阵都可逆，而可逆矩阵的乘积也可逆. □

注 1.5.8　可以证明，n 阶可逆矩阵可以表示成不多于 n^2 个初等矩阵的乘积.

设 A 可逆，定理 1.5.7 的证明提示了如何计算逆矩阵：求解 $AX=I_n$，利用 Gauss-Jordan 消元法，我们可以把 A 化成 I_n，记所需的初等行变换的乘积为 E，则有 $EA=I$，于是 $X=EAX=E$.（这说明，若 $AB=I$，则 $BA=I$.）由此得到计算逆矩阵的一种方法，把 A 和 I_n 并列排成一个 $n\times 2n$ 矩阵 $\begin{bmatrix}A & I_n\end{bmatrix}$，通过消元法对其做初等行变换，同样的初等行变换作用在 A 和 I_n 上，当 A 化成 I_n 时，I_n 就化成了 A^{-1}：

$$\begin{bmatrix}A & I_n\end{bmatrix}\xrightarrow{\text{Gauss-Jordan}}\begin{bmatrix}I_n & A^{-1}\end{bmatrix}.$$

同样的方法还可以用来求解矩阵方程 $AX=B$，其中 A 是 n 阶可逆矩阵. 对矩阵 $\begin{bmatrix}A & B\end{bmatrix}$ 做初等行变换，当 A 化成 I_n 时，B 就化成了 $A^{-1}B$.

例 1.5.9　回顾例 1.3.1 中的鸡兔同笼问题，一个笼子中有鸡和兔共 8 只，二者共有 22 条腿，设鸡有 x_1 只，兔有 x_2 只，有线性方程组

$$\begin{cases} x_1+x_2=8,\\ 2x_1+4x_2=22,\end{cases}$$

亦即

$$\begin{bmatrix}1 & 1\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}8\\22\end{bmatrix}.$$

x_1, x_2 的求解可以直接用消元法计算出来. 当然，根据定理 1.5.7，我们也可以通过计算系数矩阵的逆来求解.

记系数矩阵为 A，对 $\begin{bmatrix}A & I_2\end{bmatrix}$ 消元：

$$\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\2 & 4 & 0 & 1\end{array}\right]\xrightarrow{\text{倍加变换}}\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\0 & 2 & -2 & 1\end{array}\right]$$

$$\xrightarrow{\text{倍乘变换}}\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\0 & 1 & -1 & \frac{1}{2}\end{array}\right]\xrightarrow{\text{倍加变换}}\left[\begin{array}{cc|cc}1 & 0 & 2 & -\frac{1}{2}\\0 & 1 & -1 & \frac{1}{2}\end{array}\right].$$

因此 A 可逆，且 $A^{-1}=\begin{bmatrix}2 & -\frac{1}{2}\\-1 & \frac{1}{2}\end{bmatrix}$.

注意 A 的列向量组, 第一列表示一只鸡的头腿数目, 第二列表示一只兔的头腿数目. 那么 A^{-1} 的列向量组有什么实际含义？由 $AA^{-1}=I_2$ 可得

$$\begin{bmatrix}1 & 1\\2 & 4\end{bmatrix}\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix},\quad\begin{bmatrix}1 & 1\\2 & 4\end{bmatrix}\begin{bmatrix}-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}.$$

也就是说，两只鸡和负一只兔子的组合刚好有一个头且没有腿；而负半只鸡和半只兔子的组合刚好没有头但有一条腿. 而一个有 8 头 22 腿的动物组合，就应该是

$$8\begin{bmatrix}2\\-1\end{bmatrix}+22\begin{bmatrix}-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}=\begin{bmatrix}2 & -\frac{1}{2}\\-1 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}8\\22\end{bmatrix}=\begin{bmatrix}5\\3\end{bmatrix},$$

这是因为动物头和腿的数量与动物的数量之间满足线性叠加原理. ☺

定理 1.5.7 可以用来证明某些具有特殊形式的矩阵的可逆性. 先来看前面提及的三角矩阵，容易验证如下结论.

命题 1.5.10　上三角矩阵可逆当且仅当其对角元素都不为零. 此时，其逆矩阵也是上三角矩阵，逆矩阵的对角元素是该矩阵的对应对角元素的倒数.

下三角矩阵也有类似性质.

定义 1.5.11 (置换矩阵)　单位矩阵经一系列对换行变换得到的矩阵称为**置换矩阵**.

简单验证可以得到置换矩阵的一些性质.

命题 1.5.12

1. 单位矩阵经一系列对换列变换得到的矩阵也是置换矩阵;
2. 置换矩阵可以通过按某种次序排列单位矩阵的行来得到;
3. 置换矩阵可以通过按某种次序排列单位矩阵的列来得到;
4. 不同的 n 阶置换矩阵共有 $n!$ 个[1];
5. 置换矩阵的乘积也是置换矩阵;
6. 置换矩阵的逆是其转置，也是置换矩阵.

定义 1.5.13 (对角占优)　如果矩阵 $A=\left[a_{ij}\right]_{n\times n}$ 对 $i=1,2,\cdots,n$，都有 $|a_{ii}|>\displaystyle\sum_{j\neq i}|a_{ij}|$，则称其为（行）**对角占优矩阵**.

命题 1.5.14　对角占优矩阵一定可逆.

证　根据定理 1.5.7，只需证明 $A\boldsymbol{x}=\mathbf{0}$ 只有零解. 设 $\boldsymbol{x}$ 的分量中，绝对值最大的元素是 x_i. 考虑方程组的第 i 个方程:

$$a_{i1}x_1+\cdots+a_{i,i-1}x_{i-1}+a_{ii}x_i+a_{i,i+1}x_{i+1}+\cdots+a_{in}x_n=0.$$

若 $x_i\neq 0$，根据对角占优性质，有

$$|a_{ii}||x_i|=\left|\sum_{j\neq i}a_{ij}x_j\right|\leqslant\sum_{j\neq i}|a_{ij}||x_j|\leqslant\sum_{j\neq i}|a_{ij}||x_i|<|a_{ii}||x_i|.$$

矛盾. □

例 1.5.15　给定铁丝网如图 1.5.1 所示，铁丝粗细均匀，网格等距，格点温度如图所标. 假设格点的温度都已经稳定，亦即每个格点的温度等于与其相连的格点的温度的平均值，那么各格点上的温度是多少?

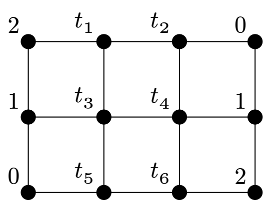


图 1.5.1　热传导

可以列出线性方程组:

$$\begin{cases}3t_1 &= t_2+t_3+2,\\ 3t_2 &= t_1+t_4+0,\\ 4t_3 &= t_1+t_4+t_5+1,\\ 4t_4 &= t_2+t_3+t_6+1,\\ 3t_5 &= t_3+t_6+0,\\ 3t_6 &= t_4+t_5+2,\end{cases}$$

[1] 符号 $n!$ 表示 n 的阶乘，即 $n!=1\cdot 2\cdot\cdots\cdot n$.

亦即

$$\begin{bmatrix} 3 & -1 & -1 & 0 & 0 & 0 \\ -1 & 3 & 0 & -1 & 0 & 0 \\ -1 & 0 & 4 & -1 & -1 & 0 \\ 0 & -1 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 0 & 3 & -1 \\ 0 & 0 & 0 & -1 & -1 & 3 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

不难看出，线性方程组的系数矩阵对角占优．根据命题 1.5.14，该矩阵可逆，因此方程组有唯一解. ☺

最后，利用矩阵的逆，我们可以定义矩阵的负整数幂．对正整数 k，如果 A 是可逆矩阵，则其负整数幂定义为 $A^{-k}=(A^{-1})^k$，其零次幂为 $A^0=I_n$．注意，$A^{-k}=(A^k)^{-1}$.

还可以定义方阵的多项式：给定 n 阶方阵 A 和多项式 $f(x)=a_0+a_1x+\cdots+a_px^p$，定义 $f(A)=a_0I_n+a_1A+\cdots+a_pA^p$.

1.5.3 等价关系

我们已经知道，一系列初等行变换可以把矩阵化成行简化阶梯形矩阵．为描述矩阵在初等行变换前后的关系，我们给出如下定义.

定义 1.5.16 (左相抵) 如果矩阵 A 可以经过一系列初等行变换化成矩阵 B，则称 A 和 B **左相抵**.

命题 1.5.17 给定两个 $m\times n$ 矩阵 A,B．那么二者左相抵，当且仅当存在 m 阶可逆矩阵 P，使得 $PA=B$.

证 A 与 B 左相抵 $\iff$ A 经一系列初等行变换得到 B $\iff$ 存在 m 阶初等矩阵 $P_1,P_2,\cdots,P_s$，使得 $P_s\cdots P_2P_1A=B \overset{\text{定理 1.5.7}}{\iff}$ 存在 m 阶可逆矩阵 P，使得 $PA=B$. □

可以看到在所有和矩阵 A 左相抵的矩阵中，形式最简单的应该就是其行简化阶梯形矩阵，因此我们也可以称此矩阵为 A 的**左相抵标准形**.

容易看到如果两个矩阵左相抵，那么以二者为系数矩阵的齐次线性方程组，或者以二者为增广矩阵的线性方程组，其解集相同.

不难看出，左相抵关系满足如下三条基本性质：

1. 反身性：每个矩阵和自身左相抵；
2. 对称性：如果 A 和 B 左相抵，那么 B 和 A 左相抵；
3. 传递性：如果 A 和 B 左相抵，B 和 C 左相抵，那么 A 和 C 左相抵.

事实上，很多数学对象之间都存在类似的关系．为此，我们抽象出一系列概念.

定义 1.5.18 (等价关系) 如果非空集合 S 的元素之间定义了一种二元关系“$\sim$”，满足：

1. 反身性：对任意 $a \in S$，$a \sim a$；
2. 对称性：如果 $a \sim b$，那么 $b \sim a$；
3. 传递性：如果 $a \sim b, b \sim c$，那么 $a \sim c$；

则称此关系为 S 上的一个**等价关系**.

S 中所有与其中某一元素 a 等价的元素组成的集合称为 a 所在的**等价类**. 由元素 a 变成与其等价的元素 b 的过程称为对应于这一等价关系的**等价变换**. 同一等价类中的元素共享的某个属性称为这一等价关系中 (或这一等价变换下) 的**不变量**. 同一等价类中的形式最简单、性质最好的元素往往称为这一等价关系中 (或这一等价变换下) 的**标准形**.

显然，S 等于所有等价类的无交并. 等价类的数目可以有限，也可以无限.

例 1.5.19 平面上所有三角形组成的集合上，相似是一个等价关系，三个角的角度是等价关系中的不变量.

平面上所有三角形组成的集合上，全等也是一个等价关系，三个角的角度和三条边的长度是等价关系中的不变量.

所有 $m \times n$ 矩阵组成的集合上，左相抵是一个等价关系，初等行变换是等价变换，作为系数矩阵得到的齐次线性方程组的解集是等价关系中的不变量，作为增广矩阵得到的线性方程组的解集也是等价关系中的不变量，行简化阶梯形矩阵是此等价关系中的标准形.

所有关于未知数 x 的一元方程组成的集合上，同解是一个等价关系，同解变换[1]是等价变换，解集是等价关系中的不变量，而解出的 $x = x_1, x_2, \cdots$ 是等价关系中的标准形. ☺

为了研究某个集合的某个问题，在此集合上引入等价关系，然后找到其标准形，往往能够使问题的研究得到简化：这是数学中的一个重要方法. 这一方法将在本书中反复体现. 我们将为矩阵定义更多的变换，并通过找到其标准形来研究矩阵的性质.

习题

练习 1.5.1 计算下列矩阵乘法：

1. $\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}$.

2. $\begin{bmatrix} 1 & & \\ -1 & 1 & \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 4 & 0 \end{bmatrix}$.

3. $\begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \end{bmatrix}^k$，$k$ 是正整数.

4. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

[1] 大家中学都学过，只有保持解不变的变换才能用来解方程或方程组，这种变换称为同解变换.

5. $\begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ -1 & & 1 & & \\ -1 & & & 1 & \\ -1 & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$.

练习 1.5.2 求下列矩阵的逆矩阵:

1. $\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$.
2. $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$.
3. $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{bmatrix}$.
4. $\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$.

练习 1.5.3 设 $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & a & -2 \\ 5 & -2 & 1 \end{bmatrix}$ 不可逆，求 a.

练习 1.5.4 设 $A = \begin{bmatrix} 1 & & \\ 2 & 3 & \\ 4 & 5 & 6 \end{bmatrix}$，解方程组 $A\boldsymbol{x} = \boldsymbol{e}_i, i = 1, 2, 3$，并求 A^{-1}.

练习 1.5.5 对 $\begin{bmatrix} A & I_n \end{bmatrix}$ 做初等行变换，求下列矩阵的逆矩阵:

1. $\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}$.
2. $\begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}$.
3. $\begin{bmatrix} 0 & 1 & & \\ & 0 & & 1 \\ 1 & & 0 & \\ & & 1 & 0 \end{bmatrix}$.
4. $\begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$.
5. $\begin{bmatrix} 1 & 2 & & \\ 3 & 4 & & \\ & & 1 & 2 \\ & & 3 & 4 \end{bmatrix}$.
6. $\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$.
7. $\begin{bmatrix} 1 & -a & & \\ & 1 & -b & \\ & & 1 & -c \\ & & & 1 \end{bmatrix}$.
8. $\begin{bmatrix} 1 & a & & \\ & 1 & \ddots & \\ & & \ddots & a \\ & & & 1 \end{bmatrix}_{n \times n}$.

练习 1.5.6 说明 $\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$ 可逆，并计算其逆.

练习 1.5.7 求 $A=\begin{bmatrix} & a_1 & & \\ & & \ddots & \\ & & & a_{n-1} \\ a_n & & & \end{bmatrix}$ 的逆矩阵，其中 $a_i \neq 0$.

练习 1.5.8 求下列矩阵方程的解：$\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix} X = \begin{bmatrix} 4 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix}$.

练习 1.5.9 证明二阶方阵 $A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 可逆当且仅当 $d=a_{11}a_{22}-a_{12}a_{21}\neq 0$. 此时，

$$A^{-1}=\frac{1}{d}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

练习 1.5.10 证明：有一列元素（或一行元素）全为零的方阵不可逆.

练习 1.5.11 证明：对角元素全非零的上三角矩阵 U 可逆，其逆矩阵 U^{-1} 也是上三角矩阵，且 U^{-1} 的对角元素是 U 的对角元素的倒数.

练习 1.5.12 是否只有方阵才有可能可逆？设 A 是 $m\times n$ 矩阵.

1. 设 $A=\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$，能否找到 B,C 使得 $AB=I_2, CA=I_3$？
2. 设 $A=\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$，能否找到 B,C 使得 $AB=I_3, CA=I_2$？
3. 设 $CA=I_n$，证明 $A\boldsymbol{x}=\mathbf{0}$ 只有零解. 此时 m,n 之间有何种关系？
4. 设 $AB=I_m$，证明 $A^{\mathrm{T}}\boldsymbol{x}=\mathbf{0}$ 只有零解. 此时 m,n 之间有何种关系？
5. 设 $AB=I_m, CA=I_n$，证明 $m=n$ 且 $B=C$；由此可知，可逆矩阵一定是方阵.
6. 如果 $m\neq n$，那么 $\mathbb{R}^m,\mathbb{R}^n$ 之间是否存在线性双射？

练习 1.5.13 如果 n 阶方阵 A,B 满足 $AB=I_n$，判断 A,B 是否可逆.

练习 1.5.14 设

$$A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad B=\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad C=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},\quad D=\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad P=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

1. 求上述矩阵对应的初等行、列变换.
2. 从行变换的角度看，是否一定有 $AB=BA$？如果否，$AB-BA$ 从行变换的角度意味着什么？
3. 从行变换的角度看，是否一定有 $AC=CA$？如果否，$AC-CA$ 从行变换的角度意味着什么？
4. 从行变换的角度看，是否一定有 $BC=CB$？如果否，$BC-CB$ 从行变换的角度意味着什么？
5. 从行变换的角度看，D 是否和 A,B,C 可交换？计算 $DAD^{-1}, DBD^{-1}, DCD^{-1}$.

6. 从行变换的角度看，P 是否和 A, B, C 可交换？计算 $PAP^{-1}, PBP^{-1}, PCP^{-1}$.
7. 对三阶方阵 X，$(AX)B$ 和 $A(XB)$ 对 X 分别做了何种行、列变换？
8. 对任意矩阵 X，先做初等行变换，再做初等列变换，其结果是否等于先做该初等列变换，再做该初等行变换？这对应着矩阵乘法的什么性质？

练习 1.5.15 设矩阵 A 和 B 左相抵. 求证：

1. 如果 A 的第一列全是零，则 B 的第一列全是零.
2. 如果 A 的所有列都相同，则 B 的所有列都相同.
3. 如果 A 的第一列是第二列与第三列的和，则 B 的第一列也是第二列与第三列的和.
4. 如果 A 的第一列和第二列不成比例，则 B 的第一列和第二列也不成比例.

练习 1.5.16 设有 M_1, M_2, M_3 三个城市，城市之间有人口迁移. 定义矩阵 A，其 (i, j) 元 a_{ij} 为人在一年中从 A_j 迁移到 A_i 的概率. 注意，A 的每个元素都在 $0, 1$ 之间，且 A 的每一列的元素之和都是 1.

1. 设今年城市 M_1, M_2, M_3 的人口分别是 x_1, x_2, x_3，证明明年它们的预期人口分别是 $A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 的三个分量.
2. 人口迁移满足什么条件时，A 为列对角占优矩阵（行对角占优矩阵的转置）？人口迁移满足什么条件时，A 为行对角占优矩阵？考虑现实生活中的情形，讨论这些假设是否合理.
3. 如果把矩阵对角占优定义中的大于号换成大于等于号，则称该矩阵为**弱对角占优矩阵**. 证明 $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ 弱对角占优，且可逆，并求其逆矩阵.
4. 对任意 n 阶方阵，设它有 $n-1$ 行都是对角占优，仅有 1 行弱对角占优，该矩阵可逆吗？

练习 1.5.17 设 A 是 n 阶方阵.

1. 对任意两个多项式 $p(x), q(x)$，是否一定有 $p(A)q(A) = q(A)p(A)$？
2. 证明所有形如 $\begin{bmatrix} a & b & c \\ & a & b \\ & & a \end{bmatrix}$ 的矩阵全都彼此交换.
3. 设 $f(x) = p(x) + q(x), g(x) = p(x)q(x)$，证明：$f(A) = p(A) + q(A), g(A) = p(A)q(A)$.
4. 求 A 使得 $A + I_n, A - I_n$ 均不为零矩阵，但是 $A^2 - I_n = O$.
5. 设 $A^2 - I_n = O$，证明：任意 n 维向量 $\boldsymbol{v}$ 都存在分解式 $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$，其中 $\boldsymbol{x}, \boldsymbol{y}$ 满足
$$(A - I_n)\boldsymbol{x} = (A + I_n)\boldsymbol{y} = \boldsymbol{0}.$$
6. 设 $A^3 = O$，证明 $A + I_n$ 与 $A - I_n$ 都可逆，并求 $p(x), q(x)$ 使得 $p(A), q(A)$ 分别为其逆.

练习 1.5.18 证明，如果 n 阶方阵 A 满足 $A^2 = A$，则 $I_n - 2A$ 可逆.

练习 1.5.19 如果一个 n 阶方阵从 $(1,n)$ 元到 $(n,1)$ 元的对角线下的所有元素均为零，则称为西北矩阵. 类似地，可以定义东南矩阵. 如果 B 是西北矩阵，那么 $B^{\mathrm{T}}, B^2, B^{-1}$ 是什么矩阵？西北矩阵和东南矩阵的乘积是什么矩阵？

练习 1.5.20

1. 求一对可逆矩阵 A, B，使得 $A+B$ 不可逆.
2. 求一组不可逆矩阵 A, B，使得 $A+B$ 可逆.
3. 求一个三阶不可逆矩阵 A，使得对任意 $k>0$，$A+kI_3$ 都对角占优.

练习 1.5.21 求所有三阶矩阵 A，满足 $A^2=I_3$，且 A 的每个元素只能是 0 或 1.

练习 1.5.22 设 $A=\begin{bmatrix}1&2&3\\2&4&5\\3&5&6\end{bmatrix}$ 是对称矩阵，通过“对称化简”求其逆矩阵：

1. 将 A 的第一行的 2 倍从第二行中减去，第一行的 3 倍从第三行中减去. 这对应哪个初等矩阵 E_1？E_1^{T} 对应的列变换是什么？计算 E_1A 和 $A_1=E_1AE_1^{\mathrm{T}}$.
2. 将 A_1 的第二行与第三行调换. 这对应哪个初等矩阵 E_2？E_2^{T} 对应的列变换是什么？计算 E_2A_1 和 $A_2=E_2A_1E_2^{\mathrm{T}}$.
3. 将 A_2 的第三行的 $\dfrac{1}{3}$ 倍从第三行中减去. 这对应哪个初等矩阵 E_3？E_3^{T} 对应的列变换是什么？计算 E_3A_2 和 $A_3=E_3A_2E_3^{\mathrm{T}}$.
4. 综上，$A_3=E_3E_2E_1AE_1^{\mathrm{T}}E_2^{\mathrm{T}}E_3^{\mathrm{T}}$，由此求 A 的逆.

练习 1.5.23 求证：可逆对称矩阵的逆矩阵也是对称矩阵；可逆反对称矩阵的逆矩阵也是反对称矩阵.

练习 1.5.24 ☕ 给定 n 阶实反对称矩阵 A，求证：I_n-A 可逆.

提示：反证法，考虑 $A\boldsymbol{x}=\boldsymbol{x}$.

1.6 分块矩阵

在矩阵乘法中，我们经常把矩阵写成列向量组或者行向量组并排的形式. 例如，设 $A=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\\vdots\\\tilde{\boldsymbol{a}}_l^{\mathrm{T}}\end{bmatrix}, B=\begin{bmatrix}\boldsymbol{b}_1&\boldsymbol{b}_2&\cdots&\boldsymbol{b}_n\end{bmatrix}$，其中 $\tilde{\boldsymbol{a}}_i^{\mathrm{T}}\,(i=1,2,\cdots,l)$ 是 m 维行向量，$\boldsymbol{b}_j\,(j=1,2,\cdots,n)$ 是 m 维列向量，则

$$AB=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\\vdots\\\tilde{\boldsymbol{a}}_l^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{b}_1&\boldsymbol{b}_2&\cdots&\boldsymbol{b}_n\end{bmatrix}=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_1&\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_2&\cdots&\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\boldsymbol{b}_n\\\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_1&\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_2&\cdots&\tilde{\boldsymbol{a}}_2^{\mathrm{T}}\boldsymbol{b}_n\\\vdots&\vdots&&\vdots\\\tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_1&\tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_2&\cdots&\tilde{\boldsymbol{a}}_l^{\mathrm{T}}\boldsymbol{b}_n\end{bmatrix}.$$

另一方面，若 $n=l$，我们有

$$BA=\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{a}}_1^{\mathrm{T}}\\ \tilde{\boldsymbol{a}}_2^{\mathrm{T}}\\ \vdots\\ \tilde{\boldsymbol{a}}_l^{\mathrm{T}}\end{bmatrix}=\boldsymbol{b}_1\tilde{\boldsymbol{a}}_1^{\mathrm{T}}+\boldsymbol{b}_2\tilde{\boldsymbol{a}}_2^{\mathrm{T}}+\cdots+\boldsymbol{b}_n\tilde{\boldsymbol{a}}_l^{\mathrm{T}}.$$

可以看到，这和把向量看成元素计算矩阵乘法的规则类似，最重要的区别是矩阵乘法**不能**颠倒左右顺序，而数的乘法满足交换律，因此与顺序无关.

事实上，我们可以把矩阵分成任意大小的块，如

$$A=\begin{bmatrix}A_{11} & A_{12} & \cdots & A_{1s}\\ A_{21} & A_{22} & \cdots & A_{2s}\\ \vdots & \vdots & & \vdots\\ A_{r1} & A_{r2} & \cdots & A_{rs}\end{bmatrix},\qquad B=\begin{bmatrix}B_{11} & B_{12} & \cdots & B_{1t}\\ B_{21} & B_{22} & \cdots & B_{2t}\\ \vdots & \vdots & & \vdots\\ B_{s1} & B_{s2} & \cdots & B_{st}\end{bmatrix},$$

其中，A_{ij} 为 $l_i\times m_j$ 矩阵，B_{jk} 为 $m_j\times n_k$ 矩阵，则分块矩阵的乘法和转置分别为

$$AB=\begin{bmatrix}\sum\limits_{p=1}^{s}A_{1p}B_{p1} & \sum\limits_{p=1}^{s}A_{1p}B_{p2} & \cdots & \sum\limits_{p=1}^{s}A_{1p}B_{pt}\\ \sum\limits_{p=1}^{s}A_{2p}B_{p1} & \sum\limits_{p=1}^{s}A_{2p}B_{p2} & \cdots & \sum\limits_{p=1}^{s}A_{2p}B_{pt}\\ \vdots & \vdots & & \vdots\\ \sum\limits_{p=1}^{s}A_{rp}B_{p1} & \sum\limits_{p=1}^{s}A_{rp}B_{p2} & \cdots & \sum\limits_{p=1}^{s}A_{rp}B_{pt}\end{bmatrix},\quad A^{\mathrm{T}}=\begin{bmatrix}A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}} & \cdots & A_{r1}^{\mathrm{T}}\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}} & \cdots & A_{r2}^{\mathrm{T}}\\ \vdots & \vdots & & \vdots\\ A_{1s}^{\mathrm{T}} & A_{2s}^{\mathrm{T}} & \cdots & A_{rs}^{\mathrm{T}}\end{bmatrix}.$$

运算法则和针对元素的法则完全一样，稍加计算即可证明其正确性，这里要注意的是:

1. 分块矩阵的运算并不是新的运算，而仅仅是矩阵运算的另一种书写形式；
2. 由于矩阵乘法没有交换律，分块矩阵的乘法运算中的每一个矩阵乘法，一定要左边矩阵的块去左乘右边矩阵的块；
3. 由于矩阵的转置和本身一般不相等，所以分块矩阵的转置中的块也要转置.

分块矩阵的线性运算是显然的.

例 1.6.1　具有如下形式的矩阵

$$A=\begin{bmatrix}A_1 & O & \cdots & O\\ O & A_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & O\\ O & \cdots & O & A_s\end{bmatrix}=\begin{bmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{bmatrix}, \tag{1.6.1}$$

称为**分块对角矩阵**，其中对角线上是可能非零的矩阵 $A_1, A_2, \cdots, A_s$，称为对角块. 考虑线性映射 $\boldsymbol{A}\colon \mathbb{R}^n \to \mathbb{R}^m$，把向量 $\boldsymbol{x} \in \mathbb{R}^n$ 也作对应的分块 $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_s \end{bmatrix}$，得到

$$\boldsymbol{A}(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_s \end{bmatrix} = \begin{bmatrix} A_1\boldsymbol{x}_1 \\ A_2\boldsymbol{x}_2 \\ \vdots \\ A_s\boldsymbol{x}_s \end{bmatrix}.$$

可以看到，在分块对角矩阵对应的线性映射之下，$\boldsymbol{x}$ 的各个分块 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 互相独立. 由此得到的线性映射 $\boldsymbol{A}$ 等价于 s 个独立的子线性映射 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_s$.

当 (1.6.1) 式中的 $A_1, A_2, \cdots, A_s$ 都可逆时，分块对角矩阵 A 可逆，其逆矩阵是

$$A^{-1} = \begin{bmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_s^{-1} \end{bmatrix}.$$

也可以类似地定义分块上（下）三角矩阵. ☺

由于分成多个块的情形可以通过反复进行 2×2 的分块得到，下面主要讨论 2×2 分块矩阵. 设 $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，其中 A_{ij} 是 $m_i \times n_j$ 矩阵，则 $\boldsymbol{A}\colon \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{m_1+m_2}$ 为

$$\boldsymbol{A}(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11}\boldsymbol{x}_1 + A_{12}\boldsymbol{x}_2 \\ A_{21}\boldsymbol{x}_1 + A_{22}\boldsymbol{x}_2 \end{bmatrix}.$$

如果记 $\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{x})$ 并分块，那么

$$\boldsymbol{y}_1 = A_{11}\boldsymbol{x}_1 + A_{12}\boldsymbol{x}_2,$$

$$\boldsymbol{y}_2 = A_{21}\boldsymbol{x}_1 + A_{22}\boldsymbol{x}_2.$$

可以看到，A_{ij} 表示了 $\boldsymbol{x}_j$ 和 $\boldsymbol{y}_i$ 的关系. 由此可知，$\boldsymbol{A}$ 等价于四个线性映射 $\boldsymbol{A}_{ij}\colon \mathbb{R}^{n_j} \to \mathbb{R}^{m_i}$，可以用图 1.6.1 表示.

$$\begin{tikzcd} \mathbb{R}^{n_1} \arrow[r] \arrow[dr] & \mathbb{R}^{m_1} \\ \mathbb{R}^{n_2} \arrow[r] \arrow[ur] & \mathbb{R}^{m_2} \end{tikzcd}$$

图 1.6.1 分块对角矩阵对应的线性映射

当 $A_{21} = O$ 时，A 是分块上三角矩阵，$\boldsymbol{A}(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} \begin{bmatrix} * \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} * \\ A_{22}\boldsymbol{x}_2 \end{bmatrix}$，这里 $*$ 表示我们不关心其具体值.

例 1.6.2 对例 1.1.8 中的线性映射，我们考虑其表示矩阵的分块.

1. 结账是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的线性映射，其中定义域 $\mathbb{R}^3$ 的标准坐标向量分别是一千克苹果、一千克香蕉和一千克樱桃，它们对应的单价分别是 12.98 元、9.98 元和 129.98 元. 苹果和香蕉组成廉价组 $\boldsymbol{x}_1$，樱桃组成高价组 $\boldsymbol{x}_2$. 那么廉价组和高价组的结账的表示矩阵分别是

$$A_1 = \begin{bmatrix}\text{结账(廉价组)}\end{bmatrix} = \begin{bmatrix}12.98 & 9.98\end{bmatrix},$$

$$A_2 = \begin{bmatrix}\text{结账(高价组)}\end{bmatrix} = \begin{bmatrix}129.98\end{bmatrix}.$$

结账的总表示矩阵是

$$A = \begin{bmatrix}\text{结账(廉价组)} & \text{结账(高价组)}\end{bmatrix} = \begin{bmatrix}A_1 & A_2\end{bmatrix} = \begin{bmatrix}12.98 & 9.98 & 129.98\end{bmatrix}.$$

2. 营养是一个从 $\mathbb{R}^3$ 到 $\mathbb{R}^2$ 的线性映射. 纤维素组成补充组 $\boldsymbol{y}_1$，糖组成供能组 $\boldsymbol{y}_2$. 那么廉价组和高价组的补充组营养与供能组营养的表示矩阵分别是

$$A_{11} = \begin{bmatrix}\text{补充(廉价组)}\end{bmatrix} = \begin{bmatrix}12 & 12\end{bmatrix},$$

$$A_{21} = \begin{bmatrix}\text{供能(廉价组)}\end{bmatrix} = \begin{bmatrix}135 & 208\end{bmatrix},$$

$$A_{12} = \begin{bmatrix}\text{补充(高价组)}\end{bmatrix} = \begin{bmatrix}3\end{bmatrix},$$

$$A_{22} = \begin{bmatrix}\text{供能(高价组)}\end{bmatrix} = \begin{bmatrix}99\end{bmatrix}.$$

营养的总表示矩阵是

$$A = \begin{bmatrix}\text{补充(廉价组)} & \text{补充(高价组)} \\ \text{供能(廉价组)} & \text{供能(高价组)}\end{bmatrix} = \begin{bmatrix}A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix} = \begin{bmatrix}12 & 12 & 3 \\ 135 & 208 & 99\end{bmatrix}. ☺$$

与定义 1.5.1 中的初等矩阵类似，我们可以定义分块初等矩阵. 为方便起见，仍仅考虑 2×2 分块矩阵：

1. 分块对换矩阵 $\begin{bmatrix} & I_{n_1} \\ I_{n_2} & \end{bmatrix}$：

$$\begin{bmatrix} & I_{n_1} \\ I_{n_2} & \end{bmatrix}\begin{bmatrix}A_1 \\ A_2\end{bmatrix} = \begin{bmatrix}A_2 \\ A_1\end{bmatrix}, \quad \begin{bmatrix}B_1 & B_2\end{bmatrix}\begin{bmatrix} & I_{n_1} \\ I_{n_2} & \end{bmatrix} = \begin{bmatrix}B_2 & B_1\end{bmatrix};$$

2. 分块倍乘矩阵 $\begin{bmatrix}I_{n_1} & \\ & P\end{bmatrix}$，其中 P 是可逆矩阵：

$$\begin{bmatrix}I_{n_1} & \\ & P\end{bmatrix}\begin{bmatrix}A_1 \\ A_2\end{bmatrix} = \begin{bmatrix}A_1 \\ PA_2\end{bmatrix}, \quad \begin{bmatrix}B_1 & B_2\end{bmatrix}\begin{bmatrix}I_{n_1} & \\ & P\end{bmatrix} = \begin{bmatrix}B_1 & B_2P\end{bmatrix};$$

3. 分块倍加矩阵 $\begin{bmatrix} I_{n_1} & M \\ & I_{n_2} \end{bmatrix}$，其中 M 是 $n_1 \times n_2$ 矩阵:

$$\begin{bmatrix} I_{n_1} & M \\ & I_{n_2} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1 + MA_2 \\ A_2 \end{bmatrix}, \ \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} I_{n_1} & M \\ & I_{n_2} \end{bmatrix} = \begin{bmatrix} B_1 & B_1 M + B_2 \end{bmatrix}.$$

容易验证，三类分块初等矩阵都可逆，它们的逆矩阵分别为

$$\begin{bmatrix} & I_{n_2} \\ I_{n_1} & \end{bmatrix}, \quad \begin{bmatrix} I_{n_1} & \\ & P^{-1} \end{bmatrix}, \quad \begin{bmatrix} I_{n_1} & -M \\ & I_{n_2} \end{bmatrix}.$$

现在我们来考虑对分块矩阵 $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ 做 Gauss 消元法，只是现在数变成了矩阵. 如我们前面所说，分块矩阵的运算法则和数的基本相同，但是尤其要注意乘法的运算次序. 设 A_{11} 可逆. 直接计算，得

$$\begin{bmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \tag{1.6.2}$$

类似于命题 1.5.10，我们知道矩阵 $A_{22} - A_{21}A_{11}^{-1}A_{12}$ 可逆，等价于等式右边矩阵可逆，又等价于矩阵 A 可逆.

命题 1.6.3 若矩阵 A_{11} 和 $A_{22} - A_{21}A_{11}^{-1}A_{12}$ 可逆，则矩阵 $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ 可逆，且

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ O & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & O \\ O & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \begin{bmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}. \tag{1.6.3}$$

证 对 (1.6.2) 式右边消去右上角矩阵，即

$$\begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ O & I \end{bmatrix} = \begin{bmatrix} A_{11} & O \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \tag{1.6.4}$$

综合 (1.6.2) 式和 (1.6.4) 式，我们有

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & O \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ O & I \end{bmatrix}. \tag{1.6.5}$$

两边取逆，即得结论. □

矩阵 $A_{22}-A_{21}A_{11}^{-1}A_{12}$ 称为 A_{11} 关于 A 的 **Schur 补**.

类似地，我们有如下推论.

推论 1.6.4 若矩阵 A_{22} 和 $A_{11}-A_{12}A_{22}^{-1}A_{21}$ 可逆，则矩阵 $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ 可逆，且

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I & O \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \begin{bmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & O \\ O & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ O & I \end{bmatrix}. \tag{1.6.6}$$

证 利用命题 1.6.3 和分块对换矩阵. □

矩阵 $A_{11}-A_{12}A_{22}^{-1}A_{21}$ 称为 A_{22} 关于 A 的 Schur 补.

下面再看两个稍微复杂的例子，来体会矩阵分块的用处.

例 1.6.5 (Sherman-Morrison-Woodbury 公式) 将 (1.6.3) 式和 (1.6.6) 式右端乘出来，有

$$\begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & -(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \end{bmatrix}.$$

比较对应左上角块，我们可以得到

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} = A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1},$$

这给出了一定条件下一个可逆矩阵加上某个矩阵乘积后的逆的公式，常称为 **Sherman-Morrison-Woodbury 公式**，它在矩阵分析、系统和控制论等领域中有着广泛应用.

我们下面给出一个常用的简化版本，涉及的分块矩阵为 $\begin{bmatrix} A & \boldsymbol{u} \\ -\boldsymbol{v}^{\mathrm{T}} & 1 \end{bmatrix}$，其中 A 是 n 阶方阵，$\boldsymbol{u},\boldsymbol{v}$ 是 n 维向量.

Sherman-Morrison 公式：若 A 可逆，则 $A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆当且仅当 $1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}\neq 0$，且 $A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆时，

$$(A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = A^{-1}-\frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}A^{-1}}{1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}}. \tag{1.6.7}$$

特别地，$I+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆当且仅当 $1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\neq 0$，且 $I+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆时，

$$(I+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = I-\frac{\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}}{1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}}. \tag{1.6.8}$$

事实上，(1.6.7) 式也可以由 (1.6.8) 式来证明，留给读者思考.

来看一个简单例子. 给定可逆矩阵 A, 若其 $(1,1)$ 元 a_{11} 有一微小变化 δ, 变为 $a_{11}+\delta$, 则新得到的矩阵是否可逆? 如何计算?

不难看出新得到的矩阵为 $A+\delta \boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}$. 根据 Sherman-Morrison 公式, 该矩阵可逆当且仅当 $1+\delta\boldsymbol{e}_1^{\mathrm{T}}A^{-1}\boldsymbol{e}_1\neq 0$. 因此只要 $|\delta|<\dfrac{1}{|\boldsymbol{e}_1^{\mathrm{T}}A^{-1}\boldsymbol{e}_1|}$, 即 δ 足够小时, 该矩阵就可逆, 且其逆为 $A^{-1}-\dfrac{1}{1+\boldsymbol{e}_1^{\mathrm{T}}A^{-1}\boldsymbol{e}_1}A^{-1}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}A^{-1}$. 可以看到逆的变化只和矩阵的逆的首行首列有关. 类似地, 如果矩阵的 (i,j) 元发生微小改变, 则逆的变化只需逆的第 j 行和第 i 列就可以得到. ☺

例 1.6.6 (弹簧振子) 假设四根弹簧和三个振子如图 1.6.2 所示依次交替地挂在天花板和地板之间, 弹簧的原始长度分别为 l_1,l_2,l_3,l_4, 劲度系数分别为 k_1,k_2,k_3,k_4, 振子的质量分别为 m_1,m_2,m_3, 天花板和地板的间距是 d, 而 $d>l_1+l_2+l_3+l_4$. 系统处于静止状态, 问三个振子距离天花板有多远?

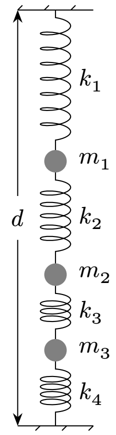


图 1.6.2　弹簧振子: 约束系统

和例 1.4.19 相比, 这一问题如果不列方程就很难直接得到解了. 我们仍然假设三个振子到天花板的距离分别是 $l_1+x_1, l_1+l_2+x_2, l_1+l_2+l_3+x_3$, 欲求 x_1,x_2,x_3. 记 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$. 令 $\delta=d-l_1-l_2-l_3-l_4$, 那么四根弹簧的拉伸长度是

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\u_3\\u_4\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\\\delta-x_3\end{bmatrix}=\delta\begin{bmatrix}0\\0\\0\\1\end{bmatrix}+\begin{bmatrix}1&&\\-1&1&\\&-1&1\\&&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$
$$=:\delta\boldsymbol{e}_4+D_{43}\boldsymbol{x}.$$

四根弹簧提供的拉力就是 $\boldsymbol{y}=\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}=K_4\boldsymbol{u}$, 其中 $K_4=\begin{bmatrix}k_1&&&\\&k_2&&\\&&k_3&\\&&&k_4\end{bmatrix}$. 振子所受拉力为 $\boldsymbol{f}=\begin{bmatrix}y_1-y_2\\y_2-y_3\\y_3-y_4\end{bmatrix}=\begin{bmatrix}1&-1&&\\&1&-1&\\&&1&-1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=D_{43}^{\mathrm{T}}\boldsymbol{y}$. 因此

$$\boldsymbol{f}=D_{43}^{\mathrm{T}}\boldsymbol{y}=D_{43}^{\mathrm{T}}K_4\boldsymbol{u}=D_{43}^{\mathrm{T}}K_4(D_{43}\boldsymbol{x}+\delta\boldsymbol{e}_4)=D_{43}^{\mathrm{T}}K_4D_{43}\boldsymbol{x}+\delta D_{43}^{\mathrm{T}}K_4\boldsymbol{e}_4.$$

注意, 和例 1.4.19 不同, 因为存在常数项 $\delta D_{43}^{\mathrm{T}}K_4\boldsymbol{e}_4$, 所以从 $\boldsymbol{x}$ 到 $\boldsymbol{f}$ 不构成线性映射.

记振子的质量向量为 $\boldsymbol{m}=\begin{bmatrix}m_1\\m_2\\m_3\end{bmatrix}$，则振子所受重力为 $\boldsymbol{g}=g\boldsymbol{m}=\begin{bmatrix}m_1g\\m_2g\\m_3g\end{bmatrix}$. 利用力的平衡原理，我们有 $\boldsymbol{g}=\boldsymbol{f}$. 因此，得到方程 $(D_{43}^{\mathrm{T}}K_4D_{43})\boldsymbol{x}=g\boldsymbol{m}-\delta D_{43}^{\mathrm{T}}K_4\boldsymbol{e}_4$，其中

$$D_{43}^{\mathrm{T}}K_4D_{43}=\begin{bmatrix}1&-1&&\\&1&-1&\\&&1&-1\end{bmatrix}\begin{bmatrix}k_1&&&\\&k_2&&\\&&k_3&\\&&&k_4\end{bmatrix}\begin{bmatrix}1&&\\-1&1&\\&-1&1\\&&-1\end{bmatrix}$$
$$=\begin{bmatrix}k_1+k_2&-k_2&\\-k_2&k_2+k_3&-k_3\\&-k_3&k_3+k_4\end{bmatrix}.$$

可以看到，这个解法可以很容易推广到任意多个弹簧和振子上.

另一方面，由于在例 1.4.19 中我们已经考虑过三根弹簧和三个振子组成的系统，因此我们也可以把上面三个弹簧和三个振子作为一个整体来考虑. 我们记这个子系统的弹簧拉伸长度、拉力、振子拉力为 $\widehat{\boldsymbol{u}},\hat{\boldsymbol{y}},\widehat{\boldsymbol{f}}$. 那么四根弹簧的拉伸长度就是

$$\boldsymbol{u}=\begin{bmatrix}\widehat{\boldsymbol{u}}\\\delta-x_3\end{bmatrix}=\begin{bmatrix}D_{33}\boldsymbol{x}\\\delta-\boldsymbol{e}_3^{\mathrm{T}}\boldsymbol{x}\end{bmatrix}=\delta\boldsymbol{e}_4+\begin{bmatrix}D_{33}\\-\boldsymbol{e}_3^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}.$$

四根弹簧提供的拉力就是

$$\boldsymbol{y}=\begin{bmatrix}\hat{\boldsymbol{y}}\\k_4u_4\end{bmatrix}=\begin{bmatrix}K_3\widehat{\boldsymbol{u}}\\k_4u_4\end{bmatrix}=\begin{bmatrix}K_3&\\&k_4\end{bmatrix}\begin{bmatrix}\widehat{\boldsymbol{u}}\\u_4\end{bmatrix}=\begin{bmatrix}K_3&\\&k_4\end{bmatrix}\boldsymbol{u}.$$

振子所受拉力为

$$\boldsymbol{f}=\widehat{\boldsymbol{f}}-\begin{bmatrix}0\\0\\y_4\end{bmatrix}=D_{33}^{\mathrm{T}}\hat{\boldsymbol{y}}-\boldsymbol{e}_3y_4=\begin{bmatrix}D_{33}^{\mathrm{T}}&-\boldsymbol{e}_3\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{y}}\\y_4\end{bmatrix}=\begin{bmatrix}D_{33}^{\mathrm{T}}&-\boldsymbol{e}_3\end{bmatrix}\boldsymbol{y}.$$

因此

$$\boldsymbol{f}=\begin{bmatrix}D_{33}^{\mathrm{T}}&-\boldsymbol{e}_3\end{bmatrix}\begin{bmatrix}K_3&\\&k_4\end{bmatrix}\begin{bmatrix}D_{33}\\-\boldsymbol{e}_3^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}+\delta\begin{bmatrix}D_{33}^{\mathrm{T}}&-\boldsymbol{e}_3\end{bmatrix}\begin{bmatrix}K_3&\\&k_4\end{bmatrix}\boldsymbol{e}_4.$$

此方程和前述方程一致. 特别地，系数矩阵

$$D_{43}^{\mathrm{T}}K_4D_{43}=\begin{bmatrix}D_{33}^{\mathrm{T}}&-\boldsymbol{e}_3\end{bmatrix}\begin{bmatrix}K_3&\\&k_4\end{bmatrix}\begin{bmatrix}D_{33}\\-\boldsymbol{e}_3^{\mathrm{T}}\end{bmatrix}=D_{33}^{\mathrm{T}}K_3D_{33}+k_4\boldsymbol{e}_3\boldsymbol{e}_3^{\mathrm{T}}.$$

我们已经知道 $D_{33}^{\mathrm{T}}K_3D_{33}$ 可逆. 利用例 1.6.5，首先 $1+k_4\boldsymbol{e}_3^{\mathrm{T}}(D_{33}^{\mathrm{T}}K_3D_{33})^{-1}\boldsymbol{e}_3 = 1+k_4(D_{33}^{-\mathrm{T}}\boldsymbol{e}_3)^{\mathrm{T}}K_3^{-1}(D_{33}^{-\mathrm{T}}\boldsymbol{e}_3)>0$，这是因为 $D_{33}^{-\mathrm{T}}\boldsymbol{e}_3=\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3$ 的每个元素都是正数. 因此 $D_{43}^{\mathrm{T}}K_4D_{43}$ 可逆.

于是这个方程有唯一解 $\boldsymbol{x}=(D_{43}^{\mathrm{T}}K_4D_{43})^{-1}(g\boldsymbol{m}-\delta D_{43}^{\mathrm{T}}K_4\boldsymbol{e}_4)$. ☺

观察例 1.6.6 和例 1.4.19，我们可以发现共同的模式，总结在图 1.6.3 中.

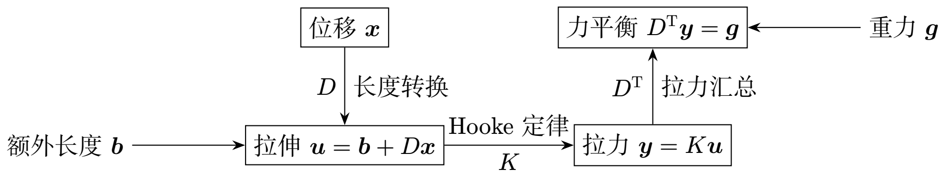


图 1.6.3 弹簧振子的基本框架

习题

练习 1.6.1 设 A 的行简化阶梯形矩阵为 R，行变换对应的可逆矩阵是 P.

1. 求 $\begin{bmatrix} A & 2A \end{bmatrix}$ 的行简化阶梯形矩阵，以及行变换对应的可逆矩阵.
2. 求 $\begin{bmatrix} A \\ 2A \end{bmatrix}$ 的行简化阶梯形矩阵，以及行变换对应的可逆矩阵.

练习 1.6.2 求下列矩阵的逆矩阵：

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

练习 1.6.3 计算下列分块矩阵：

1. $\begin{bmatrix} I_n & O \\ A & I_m \end{bmatrix}^{-1}$.
2. $\begin{bmatrix} O & I_m \\ I_n & A \end{bmatrix}^{-1}$.

练习 1.6.4 设分块矩阵 $X=\begin{bmatrix} O & A \\ B & O \end{bmatrix}$，其中 A,B 可逆. 试证 X 可逆，并求其逆.

练习 1.6.5 设分块矩阵 $U=\begin{bmatrix} A & C \\ O & B \end{bmatrix}$，其中 A,B 可逆. 试证 U 可逆，并求其逆.

练习 1.6.6 设 A_1,A_2 分别为 m,n 阶方阵，且存在可逆方阵 T_1,T_2，使得 $T_1^{-1}A_1T_1$ 和 $T_2^{-1}A_2T_2$ 都是对角矩阵. 求证：存在 $m+n$ 阶可逆矩阵 T，使得 $T^{-1}\begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}T$ 是对角矩阵.

练习 1.6.7

1. 任取 $m\times n$ 矩阵 X，分块矩阵 $\begin{bmatrix}A & B\\ C & D\end{bmatrix}$，计算 $\begin{bmatrix}I_m & X\\ O & I_n\end{bmatrix}\begin{bmatrix}A & B\\ C & D\end{bmatrix}$.

2. 由此判断 $\begin{bmatrix}1 & & & & c_1\\ & 1 & & & c_2\\ & & \ddots & & \vdots\\ & & & 1 & c_{n-1}\\ b_1 & b_2 & \cdots & b_{n-1} & a\end{bmatrix}$ 何时可逆，并在可逆时求其逆.

练习 1.6.8 利用 Sherman-Morrison 公式判断下列矩阵何时可逆，并在可逆时求其逆.

$$\begin{bmatrix}a_1+1 & 1 & 1 & \cdots & 1\\ 1 & a_2+1 & 1 & \cdots & 1\\ 1 & 1 & a_3+1 & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 1\\ 1 & 1 & \cdots & 1 & a_n+1\end{bmatrix}.$$

练习 1.6.9

1. 求一个矩阵 A 使得 $\begin{bmatrix}O & A\\ A & O\end{bmatrix}$ 为反对称矩阵.
2. 求一个 5 阶置换矩阵 P，满足 $P^5\neq I_5, P^6=I_5$.
3. 求一个对称矩阵 A，使得不存在 B，满足 $A=BB^{\mathrm{T}}$.

练习 1.6.10 ☕ 给定 $m\times n, n\times m$ 矩阵 A, B，求证：I_m+AB 可逆当且仅当 I_n+BA 可逆.

提示：考虑分块矩阵.

练习 1.6.11 ☕☕ 设实分块方阵 $X=\begin{bmatrix}A & C\\ O & B\end{bmatrix}$, 其中 A 是方阵. 如果 X 与 X^{T} 可交换, 求证: $C=O$.

提示：利用矩阵的迹.

练习 1.6.12 设分块对角矩阵 $J=\begin{bmatrix}J_1 & & \\ & J_2 & \\ & & J_3\end{bmatrix}$，其中 $J_1=\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, J_2=\begin{bmatrix}2 & 1 & \\ & 2 & 1\\ & & 2\end{bmatrix}, J_3=\begin{bmatrix}1 & 1\\ & 1\end{bmatrix}$. 试找出一个五次多项式 $f(x)$，使得 $f(J)=O$.

练习 1.6.13 构造 $2n$ 阶实方阵 A，满足 $A^2=-I_{2n}$.

练习 1.6.14 设 A, B 为两个左相抵的行简化阶梯形矩阵.

1. 证明：$A\boldsymbol{x}=\mathbf{0}$ 与 $B\boldsymbol{x}=\mathbf{0}$ 同解.
2. 证明：如果 A 的最后一列不是主列，则存在 $\boldsymbol{x}$ 使得 $A\begin{bmatrix}\boldsymbol{x}\\ 1\end{bmatrix}=\mathbf{0}$.
3. 证明：如果 A 最后一列是主列，则对任意 $\boldsymbol{x}$，$A\begin{bmatrix}\boldsymbol{x}\\ 1\end{bmatrix}\neq\mathbf{0}$.

4. 设 A, B 除了最后一列之外，其他列均相等，且 A 的最后一列不是主列，证明 $A = B$.
5. 设 A, B 除了最后一列之外，其他列均相等，且 A 的最后一列是主列，证明 $A = B$.
6. 用数学归纳法证明，左相抵的行简化阶梯形矩阵必然相等．由此证明，一个矩阵 A 的行简化阶梯形矩阵唯一.

练习 1.6.15 (置换的不动点) 设 P 是置换矩阵．如果 P 对应的行变换保持第 i 行不变，则称 i 为 P 的不动点.

1. 证明 i 是 P 的不动点当且仅当 P 的第 i 个对角元为 1.
2. 证明 P 的不动点个数等于 $\mathrm{trace}(P)$（练习 1.4.22）.
3. 对任意置换矩阵 P_1, P_2，证明 P_1P_2 和 P_2P_1 的不动点个数相等.

练习 1.6.16 考虑 $\mathbb{R}^n$ 上的变换 $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$，具有该形式的变换称为**仿射变换**．仿射变换通常不是线性变换.

1. 证明 $\begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = \begin{bmatrix} f(\boldsymbol{x}) \\ 1 \end{bmatrix}$.
2. 对仿射变换 $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$, 记 $M_f := \begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{0}^{\mathrm{T}} & 1 \end{bmatrix}$．证明对任意仿射变换 $f(\boldsymbol{x}), g(\boldsymbol{x})$, $M_fM_g = M_{f\circ g}$.
3. 证明当仿射变换 f 可逆时，矩阵 M_f 可逆，且 $M_{f^{-1}} = (M_f)^{-1}$.

练习 1.6.17 图 1.6.4 中有四个弹簧振子．弹簧的初始长度分别为 l_1, l_2, l_3, l_4，振子在稳态的最终长度为 $l_1 + x_1, l_2 + x_2, l_3 + x_3, l_4 + x_4$.

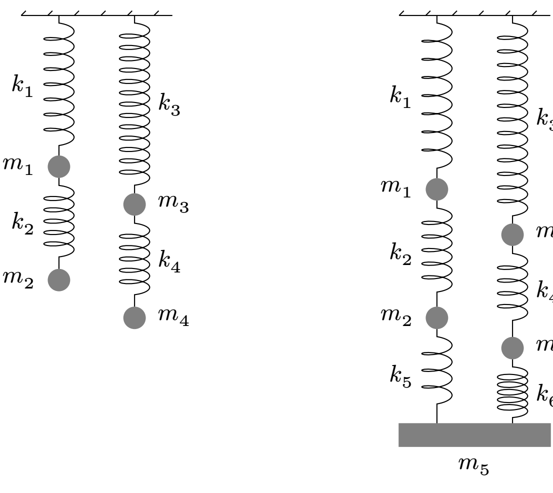


图 1.6.4

1. 找到矩阵 A 使得 $A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = g\begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix}$．此时 A 的分块结构有什么特点？你能否从系统中看出原因？

2. 图 1.6.4 中振子 m_2, m_4 下面，由劲度系数分别为 k_5, k_6 的弹簧挂住了一个质量为 m_5 的振子. 找到矩阵 A 使得 $A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = g\begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \end{bmatrix}$. 此时 A 的分块结构是否还有之前的特点？为什么？

练习 1.6.18 考虑一个元素皆为复数的矩阵. 将每个元素分成实部和虚部，不难发现对任意复数矩阵，都存在实数矩阵 A, B，使得该复数矩阵为 $A + \mathrm{i}B$，其中 i 是虚数单位. 同理，任意复数向量，都存在实数向量 $\boldsymbol{v}, \boldsymbol{w}$，使得该复数向量为 $\boldsymbol{v} + \mathrm{i}\boldsymbol{w}$. 下面用 $\mathrm{Re}(A), \mathrm{Re}(\boldsymbol{v}), \mathrm{Im}(A), \mathrm{Im}(\boldsymbol{v})$ 来表达复矩阵 A 和复向量 $\boldsymbol{v}$ 的实部和虚部.

1. 用实数矩阵和向量 $A, B, \boldsymbol{v}, \boldsymbol{w}$ 来表达 $(A + \mathrm{i}B)(\boldsymbol{v} + \mathrm{i}\boldsymbol{w})$ 的实部和虚部；
2. 对任意实数矩阵 A, B，求矩阵 X，使得 $X\begin{bmatrix} \mathrm{Re}(\boldsymbol{v}+\mathrm{i}\boldsymbol{w}) \\ \mathrm{Im}(\boldsymbol{v}+\mathrm{i}\boldsymbol{w}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}\left((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})\right) \\ \mathrm{Im}\left((A+\mathrm{i}B)(\boldsymbol{v}+\mathrm{i}\boldsymbol{w})\right) \end{bmatrix}$ 对任意实数向量 $\boldsymbol{v}, \boldsymbol{w}$ 都成立.
3. 考虑映射 $f(a + \mathrm{i}b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. 验证 $f\left((a + \mathrm{i}b)(c + \mathrm{i}d)\right) = f(a + \mathrm{i}b)f(c + \mathrm{i}d)$.

练习 1.6.19 (错位分块对角矩阵) 定义

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \triangle \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & b_{11} & 0 & b_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & b_{21} & 0 & b_{22} \end{bmatrix}.$$

1. 证明 $(A_1 \triangle B_1)(A_2 \triangle B_2) = (A_1A_2)\triangle(B_1B_2)$.
2. 证明 $A \triangle B$ 可逆当且仅当 A, B 都可逆；此时有 $(A\triangle B)^{-1} = A^{-1}\triangle B^{-1}$.
3. 求 X，使得对任意二阶方阵 A, B，都有 $X(A\triangle B)X^{-1} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$.

1.7 LU 分解

在 1.4 节中可以看到，以三角矩阵为系数矩阵的线性方程组不需要消元，只需要逐个代入就可以求出通解，即使用回代法或前代法计算. 例 1.4.19 中的线性方程组是 $D_{33}^{\mathrm{T}}K_3D_{33}\boldsymbol{x} = g\boldsymbol{m}$，其中 D_{33}^{T} 是单位上三角矩阵，K_3 是可逆对角矩阵，D_{33} 是单位下三角矩阵，因此我们可以逐次求解：先用回代法求解 $D_{33}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{f}$；然后简单求解 $K_3\boldsymbol{u} = \boldsymbol{y}$；最后用前代法求解 $D_{33}\boldsymbol{x} = \boldsymbol{u}$. 这比先把矩阵 $D_{33}^{\mathrm{T}}K_3D_{33}$ 计算出来再通过消元法计算方程组的解要快得多.（为什么?）

注意，阶梯形方阵一定是上三角矩阵. 而我们在将矩阵化成阶梯形矩阵时只用到了对换矩阵 P_{ij} 和倍加矩阵 $E_{ji;k}(j > i)$. 在对矩阵 A 化简的过程中，如果没有使用对换矩阵，那么我们就相当于对矩阵 A 的下三角元素按照 $(2,1), (3,1), \cdots, (n,1), (3,2), \cdots, (n,2)$,

$\cdots, (n, n-1)$ 的次序逐一消元. 这类倍加矩阵是下三角矩阵. 记这些初等矩阵的乘积为 L^{-1} (我们知道它可逆), 那么就有 $L^{-1}A = U$, 其中 L 是下三角矩阵, U 是上三角矩阵. 于是, 我们有如下定理.

定理 1.7.1 (LU 分解) 如果 n 阶方阵 A 只使用倍加矩阵 $E_{ji;k}(j > i)$ 做行变换就可以化成阶梯形矩阵, 那么存在 n 阶单位下三角矩阵 L 和 n 阶上三角矩阵 U, 使得 $A = LU$.

分解 $A = LU$ 称为矩阵 A 的 **LU 分解**.

如果 A 有 LU 分解 $A = LU$, 那么求解线性方程组 $A\boldsymbol{x} = \boldsymbol{b}$, 就化成了求解 $L\boldsymbol{y} = \boldsymbol{b}$ 和 $U\boldsymbol{x} = \boldsymbol{y}$ 这两个线性方程组, 而二者的系数矩阵都是三角矩阵, 易于求解.

下面主要考虑 A 是可逆矩阵的情形.

定义 1.7.2 (顺序主子阵) 方阵 A 的左上角 $k \times k$ 块, 称为 A 的第 k 个**顺序主子阵**.

显然, n 阶方阵共有 n 个顺序主子阵.

定理 1.7.3 (可逆矩阵的 LU 分解) 对 n 阶可逆矩阵 A, A 有 LU 分解 $A = LU$, 当且仅当 A 的所有顺序主子阵可逆. 此时, A 的 LU 分解唯一.

证 充分性: 已知 $A = LU$. A 可逆, L 为单位下三角阵也可逆, 因此 $U = L^{-1}A$ 也可逆, 于是 L, U 的所有对角元都不为零. 考虑分块矩阵 $L = \begin{bmatrix} L_{11} & O \\ L_{21} & L_{22} \end{bmatrix}, U = \begin{bmatrix} U_{11} & U_{12} \\ O & U_{22} \end{bmatrix}$, 其中 L_{11}, U_{11} 是 $k \times k$ 矩阵. 显然 L_{11}, U_{11} 的对角元不为零, 二者都可逆. 注意 $A = LU = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} + L_{22}U_{22} \end{bmatrix}$, 即有 A 的第 k 个顺序主子阵是 $L_{11}U_{11}$, 可逆.

必要性: 采用数学归纳法. 当 $n = 1$ 时, 显然. 假设命题对 $n-1$ 阶方阵成立, 考虑 n 阶方阵 A, 条件已知其所有顺序主子阵可逆.

考虑分块矩阵 $A = \begin{bmatrix} A_1 & \boldsymbol{x} \\ \boldsymbol{y}^{\mathrm{T}} & a \end{bmatrix}$, 其中 A_1 是 $n-1$ 阶方阵, 且显然满足归纳假设, 因此有 LU 分解 $A_1 = L_1U_1$, 而 L_1, U_1 可逆. 对分块矩阵消元, 有

$$
\begin{aligned}
A &= \begin{bmatrix} I_{n-1} & \mathbf{0} \\ \boldsymbol{y}^{\mathrm{T}}A_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} A_1 & \boldsymbol{x} \\ \mathbf{0}^{\mathrm{T}} & a - \boldsymbol{y}^{\mathrm{T}}A_1^{-1}\boldsymbol{x} \end{bmatrix} \\
&= \begin{bmatrix} I_{n-1} & \mathbf{0} \\ \boldsymbol{y}^{\mathrm{T}}A_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} L_1 & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_1 & L_1^{-1}\boldsymbol{x} \\ \mathbf{0}^{\mathrm{T}} & a - \boldsymbol{y}^{\mathrm{T}}A_1^{-1}\boldsymbol{x} \end{bmatrix}
\end{aligned}
$$

$$= \begin{bmatrix} L_1 & \mathbf{0} \\ \boldsymbol{y}^{\mathrm{T}} U_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} U_1 & L_1^{-1}\boldsymbol{x} \\ \mathbf{0}^{\mathrm{T}} & a - \boldsymbol{y}^{\mathrm{T}} U_1^{-1} L_1^{-1} \boldsymbol{x} \end{bmatrix},$$

此即 A 的 LU 分解.

唯一性：设 $A = L_1U_1 = L_2U_2$. 由 A 可逆，L_1, L_2 可逆知，U_1, U_2 可逆. 因此 $L_2^{-1}L_1 = U_2U_1^{-1}$，而前者是单位上三角矩阵，后者是下三角矩阵，因此二者必为单位矩阵，即 $L_2^{-1}L_1 = U_2U_1^{-1} = I$，因此 $L_1 = L_2, U_1 = U_2$. □

例 1.7.4 从 a 个小球中选 b 个，有多少种选法？我们中学学过，这是组合数. 如果 $a \geqslant b \geqslant 0$，我们有公式 $\mathrm{C}_a^b = \dfrac{a!}{b!(a-b)!}$. 如果 $a < b$，我们定义 $\mathrm{C}_a^b = 0$.

考虑对称 **Pascal 矩阵** $A = \left[\mathrm{C}_{i+j-2}^{i-1}\right]_{n\times n}$，对其进行 LU 分解. 为简单起见，假设 $n = 4$，则 $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$. 不难看出，该矩阵的每个元素都是它上边元素和左边元素的和.

计算可知，A 的 LU 分解为

$$A = LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

注意，$U = \left[\mathrm{C}_{j-1}^{i-1}\right]_{n\times n}$ 恰好是上三角 Pascal 矩阵，每个元素都是它左边元素和左上元素的和；而 $L = \left[\mathrm{C}_{i-1}^{j-1}\right]_{n\times n}$ 恰好是下三角 Pascal 矩阵，每个元素都是它上边元素和左上元素的和. 另外，还可以看到 $L = U^{\mathrm{T}}$.

如果直接计算矩阵乘积 LU，可以得到恒等式 $\mathrm{C}_{a+b}^t = \sum\limits_{k=0}^{t} \mathrm{C}_a^k \mathrm{C}_b^{t-k}$. 这个恒等式的意义是，先把 $a+b$ 个小球分成各有 a 个和 b 个的两组，然后从 $a+b$ 个小球中选 t 个的选法，就相当于从 a 个小球的组里面选 k 个，再从 b 个小球的组里面选 $t-k$ 个，在 k 取遍从 0 到 t 的自然数时的所有选法. ☺

LU 分解，就是 Gauss 消元法的简单表达.

可逆矩阵 A 存在 LU 分解，等价于对 A 用 Gauss 消元法把 A 化成阶梯形矩阵，累积的初等矩阵为 L^{-1}，而阶梯形矩阵为 $L^{-1}A$. 之后可以利用 Gauss-Jordan 消元法继续计算：把主元化成 1，这可以通过一系列倍乘行变换得到，记累积的初等矩阵为 D^{-1}，则此时的矩阵为 $D^{-1}L^{-1}A$；再用一系列倍加行变换把矩阵化成单位矩阵 I，记累积的初等矩阵为 U^{-1}，那么 $I = U^{-1}D^{-1}L^{-1}A$. 这给出了如下结论.

命题 1.7.5 (LDU 分解) 如果 n 阶可逆方阵 A 存在 LU 分解，那么存在 n 阶单位下三角矩阵 L，对角元素均不为 0 的 n 阶对角矩阵 D，和 n 阶单位上三角矩阵 U，使得 $A=LDU$，且该分解唯一.

证 只需证明唯一性.假设 $A=L_1D_1U_1=L_2D_2U_2$.六个矩阵都可逆,因此 $L_2^{-1}L_1=D_2U_2U_1^{-1}D_1^{-1}$，而前者是单位下三角矩阵，后者是上三角矩阵，因此二者必为单位矩阵.立得 $L_1=L_2, D_2^{-1}D_1=U_2U_1^{-1}$．类似地，$D_1=D_2, U_1=U_2$. □

分解 $A=LDU$ 称为矩阵 A 的 **LDU 分解**. 注意，事实上，我们只需要计算出 L 和 D，就可以直接得到 U，而不是真需要通过一系列倍加行变换来得到 U.

推论 1.7.6 可逆对称矩阵 A，如果有 LDU 分解 $A=LDU$，则 $L=U^{\mathrm{T}}$.

证 由 A 对称，有 $LDU=A=A^{\mathrm{T}}=U^{\mathrm{T}}D^{\mathrm{T}}L^{\mathrm{T}}$．注意到 $A=LDU=U^{\mathrm{T}}DL^{\mathrm{T}}$ 都是 A 的 LDU 分解，根据 LDU 分解的唯一性，$L=U^{\mathrm{T}}$. □

分解 $A=LDL^{\mathrm{T}}$ 称为对称矩阵 A 的 **LDL$^{\mathrm{T}}$ 分解**.

注意不是任意对称矩阵都有 LDL$^{\mathrm{T}}$ 分解.

我们已经看到，LU 分解可以用来把求解线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的过程简单地形式化.但是并非任意矩阵，也并非任意可逆矩阵都可以进行 LU 分解．另一方面，我们又知道，任意可逆矩阵 A 通过初等行变换一定可以求出线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的解. 其中的差别就在于对换行变换的使用．在解线性方程组时，我们知道使用对换行变换只与其出现的先后次序有关系，而与矩阵元素的值无关．因此我们可以把解线性方程组的过程中使用的这一系列对换行变换率先施加在矩阵上，这样得到的矩阵就只需要倍加行变换就可以化成阶梯形矩阵了．这给出了如下结论.

定理 1.7.7 (PLU 分解) 给定 n 阶可逆矩阵 A，则存在 n 阶置换矩阵 P，n 阶单位下三角矩阵 L，和对角元素不为 0 的上三角矩阵 U，使得 $A=PLU$．特别地，下三角矩阵 L 可以选为所有元素的绝对值都不大于 1 的.

证 采用数学归纳法．当 $n=1$ 时，显然．假设命题对 $n-1$ 阶可逆矩阵成立.

由 A 可逆，A 的第一列元素不能全为零，因此存在置换矩阵 P_1，使得 P_1A 的 $(1,1)$ 元非零，于是

$$P_1A=\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \boldsymbol{y} & A_2 \end{bmatrix}=\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}\boldsymbol{y} & I_{n-1} \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & A_2-a^{-1}\boldsymbol{y}\boldsymbol{x}^{\mathrm{T}} \end{bmatrix}.$$

由 A 可逆且 $a\neq 0$，有 $A_2-a^{-1}\boldsymbol{y}\boldsymbol{x}^{\mathrm{T}}$ 可逆．利用归纳假设，$A_2-a^{-1}\boldsymbol{y}\boldsymbol{x}^{\mathrm{T}}=P_2L_2U_2$．因此

$$A=P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}\boldsymbol{y} & I_{n-1} \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & P_2L_2U_2 \end{bmatrix}$$

$$
\begin{aligned}
&= P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}\boldsymbol{y} & I_{n-1} \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & P_2 \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & L_2 \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & U_2 \end{bmatrix} \\
&= P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}\boldsymbol{y} & P_2 \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & L_2 \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & U_2 \end{bmatrix} \\
&= P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & P_2 \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}P_2^{-1}\boldsymbol{y} & I_{n-1} \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & L_2 \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & U_2 \end{bmatrix} \\
&= P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & P_2 \end{bmatrix}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}P_2^{-1}\boldsymbol{y} & L_2 \end{bmatrix}\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & U_2 \end{bmatrix},
\end{aligned}
$$

其中 $P_1^{-1}\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & P_2 \end{bmatrix}$ 是置换矩阵，$\begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ a^{-1}P_2^{-1}\boldsymbol{y} & L_2 \end{bmatrix}$ 是单位下三角矩阵，$\begin{bmatrix} a & \boldsymbol{x}^{\mathrm{T}} \\ \mathbf{0} & U_2 \end{bmatrix}$ 是上三角矩阵.

特别地，如果每次选择 P_1 时，使得 P_1A 的 $(1,1)$ 元在第一列中最大，就可保证 L 的所有元素的绝对值都不大于 1，请读者自行验证. □

分解 $A=PLU$ 称为矩阵 A 的 **PLU 分解**.

我们把矩阵消元法写成了矩阵的 LU 分解或 PLU 分解，从而求解了线性方程组. 具体来说，给定条件合适的矩阵 A，线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的求解过程为：

1. 首先求出 A 的 LU 分解 $A=LU$；
2. 然后利用前代法求解下三角矩阵对应的线性方程组 $L\boldsymbol{y}=\boldsymbol{b}$；
3. 最后利用回代法求解上三角矩阵对应的线性方程组 $U\boldsymbol{x}=\boldsymbol{y}$.

对 n 阶方阵，上述算法所需要的四则运算的数目可以计算出来：

1. LU 分解需要做 $\dfrac{1}{3}(2n^3-n)$ 次四则运算；
2. 前代法和回代法各需要 n^2 次四则运算.

因此，求解线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 的计算量主要用来计算系数矩阵 A 的 LU 分解. 注意 LU 分解计算量近似与 n 的三次方成正比，因此，矩阵的阶数每次翻倍，对应 LU 分解的计算量将变为原来的 8 倍.

可以看到，LU 分解把矩阵化成两个简单矩阵的乘积. 求解 $A\boldsymbol{x}=\boldsymbol{b}$ 在得到 $A=LU$ 后，就变为了求解一个以上三角矩阵为系数矩阵的方程组和一个以下三角矩阵为系数矩阵的方程组，而这两个方程组的求解要简单得多.（当然得到 LU 分解并不简单.）因此，把矩阵化为一系列矩阵的乘积，可能会给解决问题带来很大便利，也利于分析矩阵的性质，例如通过 LU 分解，我们得到了求解系数矩阵是方阵的线性方程组所需计算量与矩阵阶数的关系.

考虑到矩阵和线性变换的关系，这一过程，就是把一个复杂的线性映射分解成一系

列简单的线性映射的复合的过程；又考虑到线性变换和线性系统的关系，这一过程，也是把一个复杂的线性系统分解成一系列简单的线性系统的“串联”的过程.

一般来说，把矩阵化成一系列简单矩阵的运算结果的过程，称为矩阵分解．我们也可以把矩阵化成若干矩阵的和．例如，对任意方阵 A，都有 $A=\dfrac{A+A^{\mathrm{T}}}{2}+\dfrac{A-A^{\mathrm{T}}}{2}$，而 $\dfrac{A+A^{\mathrm{T}}}{2}$ 是对称矩阵，$\dfrac{A-A^{\mathrm{T}}}{2}$ 是反对称矩阵，这对我们把握 A 的某些性质有一定帮助.

例 1.7.8 回顾例 1.1.10 中的反射变换．关于直线 $x_2\cos\theta-x_1\sin\theta=0$ 的反射变换 $\boldsymbol{H}_\theta$，其表示矩阵为 $\begin{bmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{bmatrix}$．为写出表示矩阵，我们曾考虑 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 的像，这并不容易．然而注意到反射变换的几何含义，$\boldsymbol{H}_\theta(\boldsymbol{x})$ 和 $\boldsymbol{x}$ 的关系是二者关于该直线对称，因此 $\boldsymbol{x}-\boldsymbol{H}_\theta(\boldsymbol{x})=2\boldsymbol{y}$，其中 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 在直线的法方向（就是与直线垂直的方向）的投影．根据中学学习的平面向量的知识，我们知道 $\boldsymbol{y}=(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{v}$，其中 $\boldsymbol{v}$ 是直线的单位法向量 $\begin{bmatrix}\sin\theta\\ -\cos\theta\end{bmatrix}$，因此 $\boldsymbol{H}_\theta(\boldsymbol{x})=\boldsymbol{x}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}=(I_2-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})\boldsymbol{x}$，故反射变换的表示矩阵是 $I_2-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$．可以验证，这和前面含有三角函数的表示矩阵相等．这个矩阵的优点在于我们不再需要三角函数的参与，因为法向量和单位法向量容易写出，具体说来，直线 $a_1x_1+a_2x_2=0$ 的法向量是 $\begin{bmatrix}a_1\\ a_2\end{bmatrix}$. ☺

习题

练习 1.7.1 求下列矩阵的 LU 分解，其中 L 为单位下三角矩阵：

1. $\begin{bmatrix}1&-1&0&0\\-1&2&-1&0\\0&-1&2&-1\\0&0&-1&2\end{bmatrix}$.
2. $\begin{bmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{bmatrix}$.
3. $\begin{bmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&5\end{bmatrix}$.
4. $\begin{bmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&6\end{bmatrix}$.
5. $\begin{bmatrix}1&1&1&1\\1&2&2&2\\1&2&3&4\\1&2&3&4\end{bmatrix}$.
6. $\begin{bmatrix}1&5&5&5\\1&2&6&6\\1&2&3&7\\1&2&3&4\end{bmatrix}$.

注意观察，哪些 0 保留在了 L 中或者 U 中.

练习 1.7.2 利用行变换解下列线性方程组：

1. $\begin{bmatrix}1&5&5&5\\1&2&6&6\\1&2&3&7\\1&2&3&4\end{bmatrix}\boldsymbol{x}=\begin{bmatrix}34\\28\\23\\20\end{bmatrix}$.
2. $\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}\boldsymbol{y}=\begin{bmatrix}34\\28\\23\\20\end{bmatrix}$.

3. $\begin{bmatrix} 1 & 5 & 5 & 5 \\ 0 & -3 & 1 & 1 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & -3 \end{bmatrix} \boldsymbol{x} = \boldsymbol{y}$，这里的 $\boldsymbol{y}$ 是上一问的解.

练习 1.7.3 设 n 阶方阵 $A = \begin{bmatrix} a_{11} & \boldsymbol{w}^{\mathrm{T}} \\ \boldsymbol{v} & B \end{bmatrix}$ 存在 LU 分解 $A = LU$，其中 $\boldsymbol{v} \neq \mathbf{0}$.

1. A 的左上角元素 a_{11} 是否非零?
2. 从上往下对 A 做行变换，将 $\boldsymbol{v}$ 变为 $\mathbf{0}$ 时，需要做多少次加法、乘法?
3. 再设 A 为对称矩阵. 此时 $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 有什么关系? B 有什么性质?
4. 仍令 A 对称. 利用行变换将 $\boldsymbol{v}$ 变为 $\mathbf{0}$ 时，再利用对应的列变换将 $\boldsymbol{w}^{\mathrm{T}}$ 变为 $\mathbf{0}$，求证此时 B 变为对称矩阵. 此时需要做几次加法、乘法?

由此看到，对称矩阵 LU 分解的计算量可以节省一半.

练习 1.7.4 设 n 阶方阵

$$T = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

利用初等变换证明，存在分解式 $T = LU$，其中 L 是下三角矩阵，U 是上三角矩阵. 据此求出 T^{-1}.