*Linear Algebra Done Right*, fourth edition

# 线性代数应该这样学

## （第四版）

作者　[美] Sheldon Axler

译者　吴俊达 何阳

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

**发布日期：2024 年 8 月 4 日**

**修订日期：2025 年 9 月 13 日**

**封面上的等式：** 第 $n$ 个斐波那契数的公式.
5D 节习题的第 21 题用线性代数导出了这个公式.

**原书发布网址：**
`https://linear.axler.net`

**翻译版发布网址：**
`https://oliverwu.top/ladr4e`
`https://linear.he-yang.top`

**原书反馈邮箱：**
`linear@axler.net`

**翻译版反馈邮箱：**
`ladr2024@163.com`

# 译者序

本书是对旧金山州立大学 Sheldon Axler 教授的著作 *Linear Algebra Done Right* 第四版的翻译. 全书共分为 9 章，具体内容见后文“致教师”一节. “致学生”也建议各位读者阅读，有助于各位调整好阅读本书的心态和目标.

本书主要讨论向量空间和线性映射. 对于数学专业的学生而言，这大致对应高等代数课程第二学期的内容. 本书与国内通行教材对于知识的处理方法和编排顺序大不相同（比如“特立独行”地将行列式放在最后引入），可为师生提供全新的视角，因此可作为课程参考书. 对于非数学类的理工科学生而言，现代控制理论、优化方法等许多广阔的研究领域都需要线性映射的理论基础，而单学期的线性代数课程中往往不涉及线性映射的内容. 本书对预备知识的要求低，语言通俗易懂，论述详细，因此可作为非数学类学生的进阶自学用书，教师也可借此教材加开第二学期线性代数进阶课程.

译者在大一的《控制理论中的代数基础》课程中接触本书的第 3 版，阅读过程中深感其既生动又严谨，并从中收获颇丰. 2023 年上半年，译者偶然想起作者在书中留下了个人网址，怀着好奇心点开网页，意外发现作者在准备第四版的消息，当时便萌生了将新版本教材译成中文并借此机会重读全书的愿望. 同年 11 月，译者再次访问上述网址，发现新版本已发布. 正巧，接下来的寒假比较长，有较充裕的时间，译者便踏上了翻译之旅.

翻译的分工情况如下：各序言中，“致谢”一节由何阳翻译，其余由吴俊达翻译；前 6 章习题翻译由何阳完成，其余由吴俊达完成；第 7 章由何阳翻译；第 8 章由吴俊达翻译；第 9 章由两人合译. 马奕斌、彭宣滨两位同学参与校对了部分章节并提出了不少宝贵建议，谨此表示衷心感谢. 还要特别感谢原书作者 Sheldon Axler 教授及其同事的大力支持与热情鼓励. 此外，本翻译使用的模板是在 Ethan Deng 等编写的 ElegantBook 基础上修改而成的，感谢 ElegantLaTeX 项目组精心打造这一美观的模板.

英文的表达习惯与中文有诸多不同，且原书语言生动，若直译难免生硬别扭、失其神韵. 因此，译者在充分尊重原文句意的前提下，对其中部分字眼不求字字对应，而根据自己的理解作了补充，必要时在脚注中作出解释. 翻译本质上是一种再创作，无法代替原著. 希望读者不仅阅读翻译，更要能独立阅读原著，或者至少对照阅读一遍，想必会更有收获.

由于译者水平有限，本翻译一定有不妥之处甚至错误，欢迎读者朋友们多提意见和建议. 请将意见和建议发送至 `ladr2024@163.com`.

感谢 Sheldon Axler 教授将本书开放免费下载！衷心希望本书能帮助各位读者跨越“语言关”，更深入地掌握书中的基本内容！

译者

2024 年 8 月

# 作者简介

Sheldon Axler 在普林斯顿大学（Princeton University）取得学士学位，随后在加利福尼亚大学伯克利分校（University of California at Berkeley）取得数学博士学位.

担任麻省理工学院（MIT）的博士后摩尔讲师（Moore Instructor）期间，Axler 获得了校级教学奖. 之后，他成为密歇根州立大学（Michigan State University）的助理教授、副教授和教授，其间他获得了第一届 J. Sutherland Frame 教学奖和杰出教职工奖.

1996 年，Axler 获得了美国数学协会（Mathematical Association of America）颁发的 Lester R. Ford 解释性写作奖，获奖论文最终被扩充成为本书. 除了发表许多研究论文外，他还是六本数学教材的作者，从大一到研究生学段的内容，这些教材均有涉及. 本书的前几个版本已被超过 375 所院校采用，并已被翻译成三种语言.

Axler 是《数学信使》（*Mathematical Intelligencer*）的主编以及《美国数学月刊》（*American Mathematical Monthly*）的副编辑. 他是美国数学学会（American Mathematical Society）理事会的成员，也是美国国家数学科学研究所（Mathematical Sciences Research Institute）的董事会成员. 他还供职于 Springer 出版社的"本科生数学教材"（Undergraduate Texts in Mathematics）丛书、"研究生数学教材"（Graduate Texts in Mathematics）丛书、Universitext 丛书以及"施普林格数学专著"（Springer Monographs in Mathematics）丛书的编辑委员会.

Axler 是美国数学学会会士，并享受美国国家科学基金会（National Science Foundation）提供的多项津贴.

Axler 于 1997 年入职旧金山州立大学（San Francisco State University），并担任数学系的系主任. 他于 2002 年至 2015 年间担任理学与工程学学院的院长，之后他重新受聘为常规教职，担任数学系教授.

Carrie Heeter, Bishnu Sarangi

作者和他的猫 Moon.

# 目录

# 致学生

这可能是你第二次接触线性代数. 初次接触这个学科时，你可能偏重欧几里得空间（Euclidean space）和矩阵；这次相逢则不同，将聚焦于抽象的向量空间和线性映射. 这些术语将在后面定义，所以如果你不知道它们的意思也不用担心. 这本书从这门学科的初步知识讲起，不需要对线性代数的预备知识. 学习的关键是将全身心沉浸在严谨的数学中，并将重点放在获得对定义、定理和证明深刻的理解上.

你不能像读小说那样读数学. 你如果在不到一个小时里就匆匆浏览过了一页，可能就读得太快了. 当遇到“你应自行验证”这种字眼时，你确实应去验证一下，这常需要你动笔写点东西. 当有些步骤被省略时，你需要补上缺失的部分. 你应该深入思考并用心体会每个定义. 对于每个定理，你都应举例说明为什么每个假设都是必要的. 和其他同学讨论也有助于你理解.

为辅助阅读[1]，在书的彩印版本中，定义置于黄框中，定理则置于蓝框中[2]. 每个定理都有一个非正式的描述性名称.

请查看下面的网站以了解更多关于这本书的信息，包括免费提供的配套视频链接.

非常欢迎你提出建议、评论和更正.

祝你在学习线性代数的过程中获得成功和快乐！

Sheldon Axler

旧金山州立大学

网站: https://linear.axler.net

e-mail: linear@axler.net

[1]原文: As a visual aid. visual aid 意为“视觉教具”，就是让部分知识点更加醒目，并让学生在阅读时可以快速定义关键知识点的一种渠道.

[2]在本翻译版本中，定义和记号在绿框中，定理在橙色框中，以灰色线将例子与其他内容隔开.

# 致教师

您即将讲授的课程可能会让学生第二次接触线性代数. 在您的学生初次接触这门学科时，他们可能主要和欧几里得空间及矩阵打交道. 本课程则不同，将强调抽象的向量空间和线性映射.

这本书的书名值得解释一下. 大多数线性代数教材都是用行列式来证明有限维复向量空间上的每个线性算子都有特征值. 但是，行列式既难懂又不直观；并且，在定义行列式时，往往缺乏动机[3]. 为了证明复向量空间上特征值的存在性定理，大部分教材必须先定义行列式，再证明线性算子不可逆当且仅当其行列式等于 0，然后定义特征多项式. 这条曲折的（或许也很折磨人的）研究途径几乎无法让学生领悟到特征值为何存在.

与之相反，本书给出的简单且不含行列式的证明（例如 5.19）能提供更多的深刻见解. 本书将行列式的讨论移至最后一章，从而开辟了一条新路径，直指线性代数的核心目标——理解线性算子的结构.

本书从线性代数的初步知识开始讲起，除了要求学生具有适当的数学素养以外，不要求预备知识. 一些例子和习题涉及微积分概念，如连续性、微分和积分. 如果您的学生没有学过微积分，那么您可以直接跳过这些例子和习题；如果您的学生学过微积分，那么通过这些例子和习题，他们可以领略数学的不同分支之间的联系，由此丰富经验.

本书中大多数习题需要理解书中的证明才能完成. 即使您的学生已经看过前几章里的一些内容，他们也可能不习惯本书给出的这种习题.

本书要点按章总结如下：

- 第 1 章：本章定义了向量空间，并逐步得出它们的基本性质.
- 第 2 章：本章定义了线性无关、张成空间、基和维数，介绍了有限维向量空间的基本理论.
- 第 3 章：本章引入线性映射. 本章的关键结果是线性映射基本定理：如果 $T$ 是 $V$ 上的线性映射，则 $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$. 本章中的商空间和对偶这两个主题要比书中大部分内容抽象程度更高，您可以跳过这些主题（除了在 9D 节讨论张量积的时候需要用到对偶以外）.
- 第 4 章：本章介绍了理解线性算子所需的那部分多项式理论. 本章不包含线性代数内容. 您可以很快讲完它，特别是如果您的学生已经熟悉这些结果的话.
- 第 5 章：要研究线性算子，可将其限制于更小的子空间中，由这一想法即引出了章首的特征向量. 本章最激动人心的结果，就是“复向量空间上特征值总是存在”的一个简单证明. 随后，这个结果被用于证明复向量空间上的每个线性算子总关于某个基具有上三角矩阵. 在本章乃至本书后续部分，最小多项式起了重要的作用. 例如，本章从最小多项式角度刻画了可对角化的算子. 如果您想节约时间，可以跳过 5E 节.
- 第 6 章：本章定义了内积空间，并利用规范正交基和格拉姆-施密特过程等工具来逐步得出它们的基本性质. 本章还展示了如何使用正交投影来解决某些最小化问题. 接着引入伪逆，它是逆不存在时的有用工具. 如果您想节省一些时间，可以跳过有关伪逆的内容.

[3]原文: often defined without motivation. 大概意思是没有动机（“从天而降”）地引入了行列式的定义.

- 第 7 章：谱定理是本书的重点之一，它刻画了特征向量能组成规范正交基的线性算子. 有了前面章节的铺垫，我们得以给出特别简短的证明. 本章还讨论了正算子、等距映射、幺正算子和矩阵分解，特别是奇异值分解，它进一步引出了极分解和线性映射的范数.
- 第 8 章：本章表明，对于复向量空间上的每个算子，该向量空间中总存在由该算子的广义特征向量构成的基. 接下来，广义特征空间分解描述了复向量空间上的线性算子. 特征值的重数定义为相应广义特征空间的维数. 这些工具被用于证明复向量空间上的每个可逆线性算子都有平方根. 接着，本章证明了复向量空间上的每个线性算子都可以化为若当型. 本章最后对算子的迹进行了研究.
- 第 9 章：本章首先关注双线性型，并证明双线性型构成的向量空间是对称双线性型和交错双线性型各自构成的子空间的直和. 然后对二次型进行对角化. 接着转向对多重线性型的研究，本章表明，$n$ 维向量空间上的交错 $n$ 重线性型构成的子空间维数为 1. 这个结果引出了算子的行列式的定义，它很简洁且不依赖于基. 我们还将得出，对于复向量空间而言，行列式等于特征值的积，该乘积中每个特征值的出现次数就是它的重数. 本章最后介绍了张量积.

本书使用字母 $\mathbf{F}$ 代表实数域或复数域，这样便可同时在实向量空间和复向量空间上探究线性代数理论. 如果您和您的学生更喜欢将 $\mathbf{F}$ 视为任意的域，可参看 1A 节末尾的注记. 我更倾向于避免在这个阶段涉及任意的域，因为这样会引入更多的抽象概念而并不引出更新颖的线性代数内容. 并且，学生更容易接受将多项式看成函数的观点，而系数在有限域中的多项式需要看成更为形式化的对象. 最后，即使理论的起始部分是在任意的域上构建起来的，讨论内积空间时我们还是要回到实向量空间和复向量空间上来.

您可能没法在一学期内涵盖本书的所有内容. 在一学期的授课时长中，仅仅是上完前七章或前八章就已经需要很快的节奏. 如果一定要讲到第 9 章，就得考虑跳过 3E 节中有关商空间的内容和 3F 节的对偶（除非您想讲 9D 节中关于张量积的内容），在半小时内讲完第 4 章多项式，再跳过 5E 节的可交换算子，以及 6C 节中有关伪逆的小节.

培养学生理解和熟练运用线性代数知识的能力，是比讲授任何特定的定理都更为重要的教学目标. 数学只能通过实践来学习. 好在线性代数有许多很好的习题. 在教这门课程时，我通常会在每节课上布置一些习题作为作业，并要求在下节课时上交. 在一节课中，讲解作业有可能占去相当多的时间.

有些习题是为了引导好奇的学生探索一些重要的主题，这些主题通常不被纳入线性代数方面的第二门基础课程中.

**作者心目中的十佳结果**

下面是书中作者最喜欢的十个结果，按照它们在书中的出现顺序排列. 结课时，学生若充分理解了这些关键结果，就将在线性代数方面打下很好的基础.

- 向量空间的任意两个基都有相同的长度（2.34）
- 线性映射基本定理（3.21）
- 如果 $\mathbf{F}=\mathbf{C}$，那么特征值必存在（5.19）
- 如果 $\mathbf{F}=\mathbf{C}$，那么上三角型必存在（5.47）
- 柯西-施瓦兹不等式（6.14）
- 格拉姆-施密特过程（6.32）
- 谱定理（7.29 和 7.31）

- 奇异值分解（7.70）
- $\mathbf{F} = \mathbf{C}$ 情形下的广义特征空间分解定理（8.22）
- 如果 $\dim V = n$，那么 $V$ 上交错 $n$ 重线性型的维数等于 1（9.37）

**第四版中的主要改进和补充**

- 加入二百五十余道新的习题和七十余个新的例子.
- 增加最小多项式的使用，以提供多个结果的更简洁的证明，包括算子具有关于某个基的上三角矩阵的充分必要条件（见 5C 节），可对角化的充分必要条件（见 5D 节）和实谱定理（见 7B 节）.
- 新增关于可交换算子的章节（见 5E 节）.
- 新增关于伪逆的小节（见 6C 节）.
- 新增关于 QR 分解和科列斯基分解的小节（见 7D 节）.
- 现在，奇异值分解适用于从一个内积空间到另一个（可能不同的）内积空间的线性映射，而不仅仅适用于从一个内积空间到自身的线性算子（见 7E 节）.
- 极分解现由奇异值分解证明，这与原来的顺序相反，并使奇异值分解（见 7E 节）和极分解（见 7F 节）的证明都更简洁.
- 新增关于有限维内积空间上线性映射的范数的小节. 利用奇异值分解，在线性映射的范数定义中甚至不用提及上确界（见 7F 节）.
- 新增关于利用具有低维值域的线性映射进行逼近的小节（见 7F 节）.
- 使用初等方法获得了该命题的全新证明：如果 $T$ 是有限维复向量空间 $V$ 上的算子，则存在由 $T$ 的广义特征向量组成的 $V$ 的基（见 8.9）.
- 新增第 9 章多重线性代数，内容包括双线性型、二次型、多重线性型和张量积. 行列式现在通过交错多重线性型来定义，此种定义方式不依赖于基.
- 全新设计版式，以让本书更适于学生阅读. 例如，定义框和结果框现在是圆角而不是直角，这比原来看起来更优雅. 正文字体从 11 号减小到 10.5 号.

要获取更多关于这本书的链接和信息，请访问下面的网站. 欢迎您提出建议、评论和更正.



祝您顺利讲授线性代数课程！

Sheldon Axler
旧金山州立大学

网站: https://linear.axler.net
e-mail: linear@axler.net

# 致谢

承蒙过去两个世纪里创建线性代数的数学家们贡献的智慧财富，感激不尽．本书的所有结果都属于数学的公共遗产．某个定理的某个特例，也许是在很久以前得到了证明，之后又在不同的时期由许多数学家加强改进．为其中所有贡献者冠以恰如其分的赞誉，将是一项艰巨的任务，我也并未这样做．而无论如何，读者都不应该将本书展示的任何结果视为我个人的原创．

本书在许多人的帮助下才得以完善．本书的前三版被全球超过 375 所院校用作教材．从使用本书的师生当中，我收到了数以千计的建议和评论，其中的许多建议促成了这一版的改进．第四版的手稿在 30 所大学中进行了课堂测试．非常感谢在课堂测试期间来自师生的有益反馈．

提出建议的人有很多，为此而作的感谢名单可以写满许多页．而这名单读来乏味，因此我只提及一人，以代表该版本的所有贡献者．他就是 Noel Hinton，澳大利亚国立大学的一名研究生．他为第四版提供的建议和更正比其他任何人都多．而对于贡献建议的每一位，请允许我向你们表达衷心的感谢．万分感谢！

感谢 Springer 出版社在我需要时给予我帮助，并允许我自由地对本书的内容和外观做出最终决定．特别感谢 Springer 出版社的两位极好的数学编辑，她们与我一起完成了这个项目——Loretta Bartolini 参与了第四版工作的前半部分，而 Elizabeth Loew 参与了第四版工作的后半部分．深切感谢 David Kramer，他出色地完成了编辑工作，并使我避免了许多错误．

我还要特别感谢我的好伴侣，Carrie Heeter．她理解我，鼓励我，使我得以专注于为这一新的版本而工作．在整个写作过程中，我们的好小猫 Moon（“关于作者”页面有他的照片）为我们带来了一次次甜蜜的休憩．而在本书完成之际，Moon 突然死于血栓．我们很感激有他陪伴的五年宝贵时光．

Sheldon Axler

# 第 1 章 向量空间

线性代数是研究有限维向量空间上的线性映射的学问. 我们最终会理解这些术语的具体含义. 在本章中，我们将定义向量空间并讨论它们的基本性质.

在线性代数中，如果将复数与实数放在一起研究，就会得到更好的定理和更深刻的见解. 因此，我们将从介绍复数及其基本性质开始.

我们将把平面和三维空间这些例子推广到 $\mathbf{R}^n$ 和 $\mathbf{C}^n$，再进一步推广得到向量空间的概念. 我们将会明白，向量空间是具有满足自然的代数性质的加法和标量乘法运算的集合.

接着，我们将讨论子空间. 子空间之于向量空间，就类似子集之于集合. 最后，我们将关注子空间的和（类似于子集的并集）与子空间的直和（类似于不相交集合的并集）.

Pierre Louis Dumesnil, Nils Forsberg

图为勒内·笛卡尔（René Descartes）向瑞典女王克里斯蒂娜（Queen Christina of Sweden）讲解他的工作. 笛卡尔于 1637 年发表了用两个坐标来描述平面的观点，向量空间就是这一观点的推广.

# 1A $\mathbf{R}^n$ 和 $\mathbf{C}^n$

## 复数

你应该已经熟悉了实数集合 $\mathbf{R}$ 的基本性质．发明复数，是为了让我们可以取负数的平方根．我们的想法是，假设 $-1$ 有平方根，将其用 $\mathrm{i}$ 表示，并且它遵守通常的算术规则．正式的定义如下．

**1.1 定义：复数（complex numbers）、C**

- 一个**复数**是一个有序对 $(a,b)$，其中 $a,b\in\mathbf{R}$，不过我们会把这写成 $a+b\mathrm{i}$.
- 全体复数所构成的集合用 $\mathbf{C}$ 表示：

$$\mathbf{C}=\{a+b\mathrm{i}:a,b\in\mathbf{R}\}.$$

- $\mathbf{C}$ 上的**加法（addition）**和**乘法（multiplication）**定义为

$$(a+b\mathrm{i})+(c+d\mathrm{i})=(a+c)+(b+d)\mathrm{i},$$

$$(a+b\mathrm{i})(c+d\mathrm{i})=(ac-bd)+(ad+bc)\mathrm{i};$$

其中 $a,b,c,d\in\mathbf{R}$.

如果 $a\in\mathbf{R}$，那么我们将 $a+0\mathrm{i}$ 等同于实数 $a$．由此，我们将 $\mathbf{R}$ 视为 $\mathbf{C}$ 的子集．我们通常将 $0+b\mathrm{i}$ 简写作 $b\mathrm{i}$，将 $0+1\mathrm{i}$ 简写作 $\mathrm{i}$.

莱昂哈德·欧拉（Leonhard Euler）于 1777 年首先使用符号 i 来代表 $\sqrt{-1}$.

上述复数乘法定义式的来由可以这样说明：先假设已知 $\mathrm{i}^2=-1$，并用一般的算术规则来导出两复数乘积的公式，再用它反过来验证定义式的确满足

$$\mathrm{i}^2=-1.$$

不要去背两个复数乘积的公式：只要回忆起 $\mathrm{i}^2=-1$，再运用一般的算术规则（在 1.3 中给出），你总是可以把它重新推导出来．接下来的示例说明了此过程.

**1.2 例：复数的算术运算**

运用 1.3 中给出的分配性质和可交换性，可以算出 $(2+3\mathrm{i})(4+5\mathrm{i})$ 的值：

$$\begin{aligned}(2+3\mathrm{i})(4+5\mathrm{i})&=2\cdot(4+5\mathrm{i})+(3\mathrm{i})(4+5\mathrm{i})\\&=2\cdot4+2\cdot5\mathrm{i}+3\mathrm{i}\cdot4+(3\mathrm{i})(5\mathrm{i})\\&=8+10\mathrm{i}+12\mathrm{i}-15\\&=-7+22\mathrm{i}.\end{aligned}$$

下面是全书的第一个结果．它指出，复数加法和复数乘法具有我们熟悉的性质，正如我们所预期的那样.

**1.3 复数的算术性质**

**可交换性（commutativity）**

对于所有 $\alpha, \beta \in \mathbf{C}$，都有 $\alpha + \beta = \beta + \alpha$ 以及 $\alpha\beta = \beta\alpha$.

**可结合性（associativity）**

对于所有 $\alpha, \beta, \lambda \in \mathbf{C}$，都有 $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ 以及 $(\alpha\beta)\lambda = \alpha(\beta\lambda)$.

**恒等元（identities）**

对于所有 $\lambda \in \mathbf{C}$，都有 $\lambda + 0 = \lambda$ 以及 $\lambda 1 = \lambda$.

**加法逆元（additive inverse）**

对于每个 $\alpha \in \mathbf{C}$，都存在唯一的 $\beta \in \mathbf{C}$ 使得 $\alpha + \beta = 0$.

**乘法逆元（multiplicative inverse）**

对于每个 $\alpha \in \mathbf{C}$ 且 $\alpha \neq 0$，都存在唯一的 $\beta \in \mathbf{C}$ 使得 $\alpha\beta = 1$.

**分配性质（distributive property）**

对于所有 $\lambda, \alpha, \beta \in \mathbf{C}$，都有 $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$.

上述性质可用我们熟悉的实数性质和复数加法、复数乘法的定义证明. 接下来的例子展示了如何证明复数乘法的可交换性，而其他性质的证明则留作习题.

**1.4 例：复数乘法的可交换性**

为了说明对于所有 $\alpha, \beta \in \mathbf{C}$ 均有 $\alpha\beta = \beta\alpha$，假设

$$\alpha = a + b\mathrm{i} \quad 且 \quad \beta = c + d\mathrm{i},$$

其中 $a, b, c, d \in \mathbf{R}$. 接着由复数乘法的定义得

$$\begin{aligned} \alpha\beta &= (a + b\mathrm{i})(c + d\mathrm{i}) \\ &= (ac - bd) + (ad + bc)\mathrm{i} \end{aligned}$$

以及

$$\begin{aligned} \beta\alpha &= (c + d\mathrm{i})(a + b\mathrm{i}) \\ &= (ca - db) + (cb + da)\mathrm{i}. \end{aligned}$$

将上面两式结合实数乘法与实数加法的可交换性可得 $\alpha\beta = \beta\alpha$.

接下来，我们定义复数的加法逆元和乘法逆元，然后用这些逆元定义复数的减法和除法运算.

**1.5 定义：$-\alpha$、减法（subtraction），$1/\alpha$、除法（division）**

假设 $\alpha, \beta \in \mathbf{C}$.

- 令 $-\alpha$ 表示 $\alpha$ 的加法逆元. 于是 $-\alpha$ 是唯一使得

$$\alpha + (-\alpha) = 0$$

成立的复数.

转下页

- **C** 上的**减法**的定义为

$$\beta-\alpha=\beta+(-\alpha).$$

- 对于 $\alpha\neq 0$，令 $1/\alpha$ 和 $\frac{1}{\alpha}$ 表示 $\alpha$ 的乘法逆元．于是 $1/\alpha$ 是唯一使得

$$\alpha(1/\alpha)=1$$

  成立的复数．
- 对于 $\alpha\neq 0$，**除以** $\alpha$ 的定义为

$$\beta/\alpha=\beta(1/\alpha).$$

♣

为便于下定义，也便于证明对于实数和复数都适用的定理，我们采用以下记号．

**1.6 记号：F**

在全书中，**F** 代表 **R** 或 **C**. ♣

因此，如果我们证明了一个涉及 **F** 的定理，我们就会知道当把 **F** 替换为 **R** 或 **C** 时，这个定理也是成立的．

使用字母 **F** 是因为 **R** 和 **C** 都是所谓**域**（**field**）的实例．

称 **F** 中的元素为**标量**（**scalar**）．通常，当我们想要强调一个对象是数，而不是向量（稍后将给出定义）时，就使用“标量”这个词（它只是“数”的一个花哨的表达法）．

对于 $\alpha\in\mathbf{F}$ 以及正整数 $m$，我们定义 $\alpha^m$ 表示 $\alpha$ 自乘 $m$ 次：

$$\alpha^m=\underbrace{\alpha\cdots\alpha}_{m\text{ 个 }\alpha}.$$

这个定义蕴涵着，对于所有 $\alpha,\beta\in\mathbf{F}$ 和所有正整数 $m,n$，有

$$(\alpha^m)^n=\alpha^{mn}\quad\text{及}\quad(\alpha\beta)^m=\alpha^m\beta^m.$$

## 组

在定义 $\mathbf{R}^n$ 和 $\mathbf{C}^n$ 之前，我们先看两个重要的例子．

**1.7 例：$\mathbf{R}^2$ 和 $\mathbf{R}^3$**

- 集合 $\mathbf{R}^2$（你可以将其视作一个平面）是全体有序实数对所构成的集合：

$$\mathbf{R}^2=\{(x,y):x,y\in\mathbf{R}\}.$$

- 集合 $\mathbf{R}^3$（你可以将其视作通常的三维空间）是全体有序实数三元组所构成的集合：

$$\mathbf{R}^3=\{(x,y,z):x,y,z\in\mathbf{R}\}.$$

为了将 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 推广至更高维数，我们首先需要讨论组的概念．

**1.8 定义：组（list）、长度（length）**

- 假设 $n$ 是非负整数．一个**长度**为 $n$ 的**组**是 $n$ 个有顺序的元素，这些元素可能是数、其他组或是更抽象的对象．
- 两个组是相等的，当且仅当它们具有相同的长度和按相同顺序排列的相同元素．

许多数学家将长度为 $n$ 的组称为 $n$ **元组（$n$-tuple）**．

组的通常写法，是将其中元素以逗号分隔并用圆括号括起来．于是，长度为 2 的组就是有序对，可以写成 $(a, b)$．长度为 3 的组就是有序三元组，可以写成 $(x, y, z)$．长度为 $n$ 的组可能看起来是这样的：

$$(z_1, \ldots, z_n).$$

有时我们会单用**组**这个词而不明说其长度．但请记住，根据定义，每个组都具有有限长度，且这长度是非负整数．从而，对于形如 $(x_1, x_2, \ldots)$ 的对象，我们可以说它“具有无限的长度”，所以它不是组．

长度是 0 的组看起来是这样的：$(\,)$．我们将这样的对象看成组，是为了使一些定理不出现平凡的例外情形[1]．

组与有限集有两方面差异：在组中，顺序很重要，并且重复是有含义的；而在集合里，顺序和重复都无关紧要．

**1.9 例：组 VS 集合**

- 组 $(3, 5)$ 和 $(5, 3)$ 是不相等的，但集合 $\{3, 5\}$ 和 $\{5, 3\}$ 是相等的．
- 组 $(4, 4)$ 和 $(4, 4, 4)$ 是不相等的（长度不等），但集合 $\{4, 4\}$ 和 $\{4, 4, 4\}$ 都等于集合 $\{4\}$．

## $\mathbf{F}^n$

要定义与 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 类似的高维对象，我们将 $\mathbf{R}$ 换成 $\mathbf{F}$（它等于 $\mathbf{R}$ 或 $\mathbf{C}$）并且将 2 或 3 换成任意的正整数即可．

**1.10 记号：$n$**

在本章剩余内容中，将 $n$ 取为某一固定的正整数．

**1.11 定义：$\mathbf{F}^n$、坐标（coordinate）**

$\mathbf{F}^n$ 是全体具有 $n$ 个 $\mathbf{F}$ 中元素的组所构成的集合：

$$\mathbf{F}^n = \{(x_1, \ldots, x_n) : \text{对于 } k = 1, \ldots, n \text{ 有 } x_k \in \mathbf{F}\}.$$

对于 $(x_1, \ldots, x_n) \in \mathbf{F}^n$ 和 $k \in \{1, \ldots, n\}$，我们称 $x_k$ 是 $(x_1, \ldots, x_n)$ 的第 $k$ 个**坐标**．

如果 $\mathbf{F} = \mathbf{R}$，且 $n$ 等于 2 或 3，那么 $\mathbf{F}^n$ 的上述定义就与前面 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 的定义相吻合．

[1] “平凡”直译自“trivial”，形容数学中最显然、易证的情形．此类情形固然简单，但不能不考虑．比如此处“长度为零的组”，就会在 2.19 中 $m = 1$ 的特殊情况中遇到，见作者在 2.19 证明后所作的说明．

**1.12 例：$\mathbf{C}^4$**

$\mathbf{C}^4$ 就是全体由四个复数组成的组所构成的集合：

$$\mathbf{C}^4 = \{(z_1, z_2, z_3, z_4) : z_1, z_2, z_3, z_4 \in \mathbf{C}\}.$$

如果 $n \geq 4$，我们就无法将 $\mathbf{R}^n$ 可视化为物理实体；类似地，$\mathbf{C}^1$ 可以被视作一个平面，但是对于 $n \geq 2$ 情形，人脑就不能想象出 $\mathbf{C}^n$ 的全貌了．然而，即便 $n$ 很大，我们也可以如在 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中那样简便地在 $\mathbf{F}^n$ 中进行代数运算．例如，$\mathbf{F}^n$ 中的加法运算定义如下．

可阅读埃德温・A・艾勃特（Edwin A. Abbott）的《平面国》（*Flatland: A Romance of Many Dimensions*），其中有趣地描述了生活在 $\mathbf{R}^2$ 的生物是如何认知 $\mathbf{R}^3$ 的．这本发表于 1884 年的小说也许能帮助你想象四维或更高维的物理空间．

**1.13 定义：$\mathbf{F}^n$ 中的加法（addition in $\mathbf{F}^n$）**

$\mathbf{F}^n$ 中的**加法**定义为将对应坐标分别相加：

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

♣

如果我们使用单个字母来表示 $n$ 个数组成的组，而不是显式地写出坐标的话，往往可以更简洁地表达有关 $\mathbf{F}^n$ 的数学内容．例如，在陈述接下来的结果时，我们用的是 $\mathbf{F}^n$ 中的 $x$ 和 $y$，即便其证明仍需 $(x_1, \ldots, x_n)$ 和 $(y_1, \ldots, y_n)$ 这些更繁琐的记号．

**1.14 $\mathbf{F}^n$ 中加法的可交换性**

如果 $x, y \in \mathbf{F}^n$，那么 $x + y = y + x$.

♡

**证明** 假设 $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$ 且 $y = (y_1, \ldots, y_n) \in \mathbf{F}^n$．那么

$$\begin{aligned}
x + y &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) \\
&= (x_1 + y_1, \ldots, x_n + y_n) \\
&= (y_1 + x_1, \ldots, y_n + x_n) \\
&= (y_1, \ldots, y_n) + (x_1, \ldots, x_n) \\
&= y + x,
\end{aligned}$$

其中第二个和第四个等号成立是由于 $\mathbf{F}^n$ 中的加法定义，第三个等号成立是因为 $\mathbf{F}$ 中加法的通常的可交换性． ■

如果用单个字母来表示 $\mathbf{F}^n$ 中的元素，那么在必须列出坐标时，就用同一个字母加上合适的下标来表示．例如，如果 $x \in \mathbf{F}^n$，那么令 $x$ 等于 $(x_1, \ldots, x_n)$ 是个好的记法，如上面的证明所示．如果可行的话，只使用 $x$ 并避免显式使用坐标则更好．

符号 ■ 表示"证明完毕".

**1.15 记号：0**

令 0 表示长度为 $n$ 且所有坐标都是 0 的组：

$$0 = (0, \ldots, 0).$$

♣

这里我们以两种不同的方式使用符号 0：在上面等式的左边，符号 0 表示长度为 $n$ 的组，它是 $\mathbf{F}^n$ 中的元素；而在右边，每个 0 都表示一个数．这种做法看似可能令人困惑，实际上不会造成任何问题，因为根据上下文就能明确使用的是哪种 0.

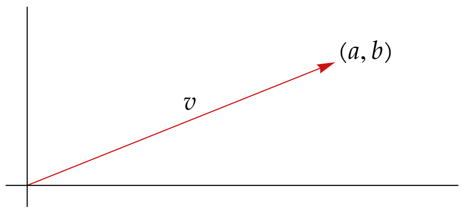
**1.16　例：根据上下文确定使用的是哪种 0**

考虑“0 是 $\mathbf{F}^n$ 中的加法恒等元”这一表达式：

$$\text{对于所有的 } x \in \mathbf{F}^n\text{，有 } x + 0 = x.$$

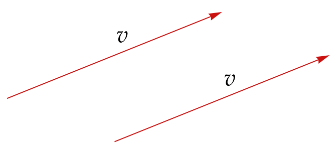
此处式中的 0 是 1.15 中定义的组，而不是数 0，因为我们并未定义 $\mathbf{F}^n$ 中元素（此处即 $x$）和数 0 的加法.

图形有助于我们直观理解．我们将在 $\mathbf{R}^2$ 中绘制图形，因为我们可以在二维的表面（如纸和计算机屏幕上）勾画这个空间．$\mathbf{R}^2$ 中的元素常表示为点 $v = (a, b)$．有时我们不将 $v$ 看成点，而看成从原点开始、到 $(a, b)$ 结束的箭头，如图所示．当我们把 $\mathbf{R}^2$ 中的元素看作箭头时，我们就称它为**向量（vector）**.



可将 $\mathbf{R}^2$ 的元素视为点或向量.

当我们把 $\mathbf{R}^2$ 中的向量看作箭头时，我们可以把箭头平行移动（不改变它的长度或方向），并且仍然认为它是同一个向量．带着这个观点，摒除坐标轴和显式的坐标而只考虑向量（如图所示），你常常可以获得更好的理解．这里展示的两个箭头有相同的长度和方向，所以我们认为它们是相同的向量.



一个向量.

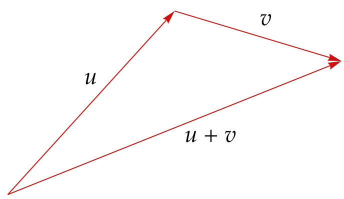
当我们利用 $\mathbf{R}^2$ 中的图形或使用点和向量这种有些模糊的说法时，请记住，这些只是辅助我们理解，而不能代替我们将要研究的真正数学内容．虽然我们画不好高维空间中的图形，但是这些空间的元素和 $\mathbf{R}^2$ 的元素一样有着严格的定义.

例如，$(2, -3, 17, \pi, \sqrt{2})$ 是 $\mathbf{R}^5$ 的一个元素，我们偶尔也将其称为 $\mathbf{R}^5$ 中的一个点或 $\mathbf{R}^5$ 中的一个向量，而无需担心 $\mathbf{R}^5$ 的几何是否有任何物理意义.

经济学中的数学模型可能有数千个变量，例如 $x_1, \ldots, x_{5000}$，也就是说，我们必须在 $\mathbf{R}^{5000}$ 中研究．这样的空间无法用几何方法来处理．然而，代数方法效果很好．因此，我们的学科叫做**线性代数**.

回忆一下，我们将 $\mathbf{F}^n$ 中两个元素之和定义为，由对应的坐标相加得到的 $\mathbf{F}^n$ 中的元素（见 1.13）．现在我们将看到，在 $\mathbf{R}^2$ 这个特殊情形中，加法具有简洁的几何解释.

假设我们想把 $\mathbf{R}^2$ 中的两个向量 $u$ 和 $v$ 相加．把 $v$ 平行移动，使它的起点与向量 $u$ 的终点重合，如右图所示．那么，$u$ 与 $v$ 之和 $u + v$ 就等于始于 $u$ 的起点，止于 $v$ 的终点的向量.



两向量之和.

在接下来的定义中，单行列出的等式右侧的 0 是组 $0 \in \mathbf{F}^n$.

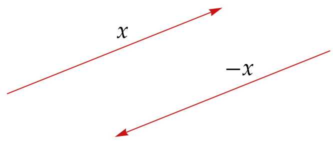
**1.17 定义：$\mathbf{F}^n$ 中的加法逆元（additive inverse in $\mathbf{F}^n$）、$-x$**

对于 $x \in \mathbf{F}^n$，$x$ 的**加法逆元**，记作 $-x$，是满足下式的向量 $-x \in \mathbf{F}^n$：

$$x + (-x) = 0.$$

由此，如果 $x = (x_1, \ldots, x_n)$，那么 $-x = (-x_1, \ldots, -x_n)$.

$\mathbf{R}^2$ 中一向量的加法逆元是与之长度相同但指向相反方向的向量．这里的图说明了这种思考 $\mathbf{R}^2$ 中加法逆元的方式．正如你所见，标记为 $-x$ 的向量与标记为 $x$ 的向量具有相同的长度，但是指向相反的方向.




一向量及其加法逆元.

讨论完 $\mathbf{F}^n$ 中的加法后，我们现在转而研究乘法．我们本可以用与加法类似的方式定义 $\mathbf{F}^n$ 中的乘法：取出 $\mathbf{F}^n$ 的两个元素，将它们对应的坐标相乘，得出 $\mathbf{F}^n$ 中的另一个元素．经验表明，这种定义无助于实现我们的目的．另一类乘法，称为标量乘法，将成为我们讨论的核心．具体地说，我们需要定义将 $\mathbf{F}^n$ 中的一个元素乘以 $\mathbf{F}$ 中的一个元素是什么含义.

**1.18 定义：$\mathbf{F}^n$ 中的标量乘法（scalar multiplication in $\mathbf{F}^n$）**

数 $\lambda$ 与 $\mathbf{F}^n$ 中的向量之**乘积**（**product**）是通过将这向量的每一个坐标都乘以 $\lambda$ 计算得到的：

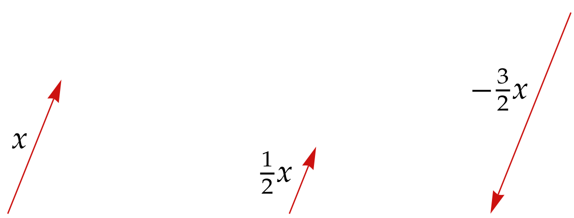
$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

此处 $\lambda \in \mathbf{F}$ 且 $(x_1, \ldots, x_n) \in \mathbf{F}^n$.

$\mathbf{F}^n$ 中的标量乘法将一个标量和一个向量相乘，得到一个向量．相比之下，$\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中的点积将两个向量相乘，得到一个标量．在第 6 章中，点积的推广形式将变得很重要.

标量乘法在 $\mathbf{R}^2$ 中具有很漂亮的几何解释．如果 $\lambda > 0$ 且 $x \in \mathbf{R}^2$，那么 $\lambda x$ 就是与 $x$ 指向相同且长度是 $x$ 的 $\lambda$ 倍的向量．换句话说，为了得到 $\lambda x$，我们把 $x$ 缩短或者延长到原来的 $\lambda$ 倍，至于是缩短还是延长取决于 $\lambda$ 是小于 1 还是大于 1.

如果 $\lambda < 0$ 且 $x \in \mathbf{R}^2$，那么 $\lambda x$ 就是与 $x$ 指向相反且长度是 $x$ 的 $|\lambda|$ 倍的向量，如此处所示.




标量乘法.

## 关于域的题外话

一个**域**是这样一个集合：它包含至少两个不同的元素（称作 0 和 1），且带有满足 1.3 中列出的所有性质的加法和乘法运算．因此，$\mathbf{R}$ 和 $\mathbf{C}$ 都是域，定义了通常的加法和乘法运算的有理数集合也是域．域的另一个例子是集合 $\{0, 1\}$，它具有通常的加法和乘法运算，除了 $1 + 1$ 被定义为等于 0.

在这本书中我们将不会与除 **R** 和 **C** 之外的域打交道. 然而，线性代数中许多适用于域 **R** 和 **C** 的定义、定理和证明无需更改即可在其他任意的域中成立. 在这本书中的多半部分（除了第 6 章和第 7 章，这两章研究内积空间），你都可以将 **F** 看成代表任意的域（而非 **R** 和 **C**），如果你喜欢这样做的话. 有些结果以“**F** 是 **C**”为前提，对于这些结果（除在关于内积的章节外），你大概率可以将这一前提替换成“**F** 是代数闭域”（意即每个不是常值且系数在 **F** 中的多项式都有零点）. 一些结果，例如 1C 节习题 13，需要在 **F** 上假定 $1+1 \neq 0$.

## 习题 1A

**1** 证明：$\alpha+\beta=\beta+\alpha$ 对所有 $\alpha, \beta \in \mathbf{C}$ 成立.

**2** 证明：$(\alpha+\beta)+\lambda=\alpha+(\beta+\lambda)$ 对所有 $\alpha, \beta, \lambda \in \mathbf{C}$ 成立.

**3** 证明：$(\alpha\beta)\lambda=\alpha(\beta\lambda)$ 对所有 $\alpha, \beta, \lambda \in \mathbf{C}$ 成立.

**4** 证明：$\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$ 对所有 $\lambda, \alpha, \beta \in \mathbf{C}$ 成立.

**5** 证明：对任一 $\alpha \in \mathbf{C}$，都存在唯一的 $\beta \in \mathbf{C}$ 使得 $\alpha+\beta=0$.

**6** 证明：对任一 $\alpha \in \mathbf{C}$（$\alpha \neq 0$），都存在唯一的 $\beta \in \mathbf{C}$ 使得 $\alpha\beta=1$.

**7** 证明：
$$\frac{-1+\sqrt{3}\mathrm{i}}{2}$$
是 1 的立方根（意即它的立方等于 1）.

**8** 求 i 的两个相异平方根.

**9** 求 $x \in \mathbf{R}^4$ 使得
$$(4,-3,1,7)+2x=(5,9,-6,8).$$

**10** 解释为什么不存在 $\lambda \in \mathbf{C}$ 使得
$$\lambda(2-3\mathrm{i}, 5+4\mathrm{i}, -6+7\mathrm{i})=(12-5\mathrm{i}, 7+22\mathrm{i}, -32-9\mathrm{i}).$$

**11** 证明：$(x+y)+z=x+(y+z)$ 对所有 $x, y, z \in \mathbf{F}^n$ 成立.

**12** 证明：$(ab)x=a(bx)$ 对所有 $x \in \mathbf{F}^n$ 和 $a, b \in \mathbf{F}$ 成立.

**13** 证明：$1x=x$ 对所有 $x \in \mathbf{F}^n$ 成立.

**14** 证明：$\lambda(x+y)=\lambda x+\lambda y$ 对所有 $\lambda \in \mathbf{F}$ 和 $x, y \in \mathbf{F}^n$ 成立.

**15** 证明：$(a+b)x=ax+bx$ 对所有 $a, b \in \mathbf{F}$ 和 $x \in \mathbf{F}^n$ 成立.

“你会加法吗？”白皇后问道，“一加一加一加一加一加一加一加一加一加一等于多少？”
“我不知道.”爱丽丝回答，“我数不清了.”

——刘易斯·卡罗尔（Lewis Carroll）《爱丽丝镜中奇遇记》（*Through the Looking Glass*）

# 1B 向量空间的定义

定义向量空间的动机源于 $\mathbf{F}^n$ 中加法和标量乘法的性质：加法满足交换律和结合律，并且有恒等元；每个元素都有一个加法逆元；标量乘法满足结合律；与 1 作标量乘法就像我们预期的那样不改变被乘的元素；加法和标量乘法通过分配性质联系起来.

我们将向量空间定义为一个集合 $V$，这个集合 $V$ 具有满足上一段所述性质的加法和标量乘法运算.

**1.19 定义：加法（addition）、标量乘法（scalar multiplication）**

- 集合 $V$ 上的**加法**是一个函数，它将每一对 $u, v \in V$ 对应到一个元素 $u + v \in V$.
- 集合 $V$ 上的**标量乘法**是一个函数，它将每个 $\lambda \in \mathbf{F}$ 和每个 $v \in V$ 对应到一个元素 $\lambda v \in V$.

♣

现在我们做好了充足的准备，可以给出向量空间的正式定义了.

**1.20 定义：向量空间（vector space）**

一个**向量空间**是一个集合 $V$，$V$ 上的加法和标量乘法满足下列性质：

**可交换性（commutativity）**

对于所有 $u, v \in V$，都有 $u + v = v + u$.

**可结合性（associativity）**

对于所有 $u, v, w \in V$ 以及所有 $a, b \in \mathbf{F}$，都有 $(u + v) + w = u + (v + w)$ 以及 $(ab)v = a(bv)$.

**加法恒等元（additive identity）**

对于所有 $v \in V$，都存在 $0 \in V$ 使得 $v + 0 = v$.

**加法逆元（additive inverse）**

对于每个 $v \in V$，都存在 $w \in V$ 使得 $v + w = 0$.

**乘法恒等元（multiplicative identity）**

对于所有 $v \in V$，都有 $1v = v$.

**分配性质（distributive properties）**

对于所有 $u, v \in V$ 以及所有 $a, b \in \mathbf{F}$，都有 $a(u + v) = au + av$ 且 $(a + b)v = av + bv$. ♣

以下的几何语言有时有助于我们直观理解.

**1.21 定义：向量（vector）、点（point）**

向量空间的元素被称作**向量**或**点**.

♣

向量空间上的标量乘法依赖于 $\mathbf{F}$ 的选取. 由此，当我们需要描述得更确切时，我们会说 $V$ 是 **$\mathbf{F}$ 上的向量空间（vector space over F）**，而不是仅仅说 $V$ 是向量空间. 例如，$\mathbf{R}^n$ 是 $\mathbf{R}$ 上的向量空间，而 $\mathbf{C}^n$ 是 $\mathbf{C}$ 上的向量空间.

**1.22 定义：实向量空间（real vector space）、复向量空间（complex vector space）**

- $\mathbf{R}$ 上的向量空间称作实向量空间.
- $\mathbf{C}$ 上的向量空间称作复向量空间.

通常，$\mathbf{F}$ 的选取要么可以从上下文明确得知，要么无关紧要. 因此我们一般假定 $\mathbf{F}$ 暗含于语境中，而无需专门提及它.

最简单的向量空间是 $\{0\}$，它只包含一个点.

带有通常的加法和标量乘法运算的 $\mathbf{F}^n$ 是 $\mathbf{F}$ 上的向量空间（你应自行验证这一点）. $\mathbf{F}^n$ 这个例子为我们定义向量空间提供了动因.

**1.23 例：$\mathbf{F}^\infty$**

$\mathbf{F}^\infty$ 定义为全体由 $\mathbf{F}$ 中元素组成的序列所构成的集合：

$$\mathbf{F}^\infty = \{(x_1, x_2, \ldots) : x_k \in \mathbf{F}, \text{ 其中 } k = 1, 2, \ldots\}.$$

$\mathbf{F}^\infty$ 上的加法和标量乘法的定义也和我们预期的一样：

$$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots),$$
$$\lambda(x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \ldots).$$

有了这些定义，$\mathbf{F}^\infty$ 就成为 $\mathbf{F}$ 上的向量空间（你应自行验证）. 这个向量空间中的加法恒等元是全 0 序列.

接下来这个向量空间的实例涉及函数集合.

**1.24 记号：$\mathbf{F}^S$**

- 如果 $S$ 是集合，那么 $\mathbf{F}^S$ 表示从 $S$ 到 $\mathbf{F}$ 的所有函数构成的集合.
- 对于 $f, g \in \mathbf{F}^S$，**和** $f + g \in \mathbf{F}^S$ 是由下式定义的函数：对于所有 $x \in S$，
$$(f + g)(x) = f(x) + g(x).$$
- 对于 $\lambda \in \mathbf{F}$ 与 $f \in \mathbf{F}^S$，**乘积** $\lambda f \in \mathbf{F}^S$ 是由下式定义的函数：对于所有 $x \in S$，
$$(\lambda f)(x) = \lambda f(x).$$

举个上述记号的具体例子：如果 $S$ 是区间 $[0,1]$ 且 $\mathbf{F} = \mathbf{R}$，那么 $\mathbf{R}^{[0,1]}$ 是全体定义在区间 $[0,1]$ 上的实值函数所构成的集合.

你应当自行验证下面例子中给出的三点结论.

**1.25 例：$\mathbf{F}^S$ 是向量空间**

- 如果 $S$ 是非空集合，那么 $\mathbf{F}^S$（带有定义如上的加法和标量乘法）是 $\mathbf{F}$ 上的向量空间.
- $\mathbf{F}^S$ 的加法恒等元是由下式定义的函数 $0 : S \to \mathbf{F}$：对于所有 $x \in S$，
$$0(x) = 0.$$

转下页

- 对于 $f \in \mathbf{F}^S$，$f$ 的加法逆元是由下式定义的函数 $-f : S \to \mathbf{F}$：对于所有 $x \in S$，

$$(-f)x = -f(x).$$

向量空间 $\mathbf{F}^n$ 是向量空间 $\mathbf{F}^S$ 的一个特例，因为每个 $(x_1, \dots, x_n) \in \mathbf{F}^n$ 都可被视作一个从集合 $\{1, 2, \dots, n\}$ 到 $\mathbf{F}$ 的函数 $x$，只要将 $(x_1, \dots, x_n)$ 的第 $k$ 个坐标写成 $x(k)$ 而不是 $x_k$ 就可看出．换句话说，我们可以将 $\mathbf{F}^n$ 看成 $\mathbf{F}^{\{1,2,\dots,n\}}$．类似地，我们可以将 $\mathbf{F}^\infty$ 看成 $\mathbf{F}^{\{1,2,\dots\}}$．

向量空间 $\mathbf{R}^{[0,1]}$ 的元素是区间 $[0,1]$ 上的实值函数，而不是组．通常，一个向量空间是一个抽象实体，其元素可能是组、函数或奇怪的对象．

稍后我们将看到向量空间的更多例子，但在此之前，我们需要先探究向量空间的一些基本性质．

向量空间的定义要求它存在加法恒等元，而接下来的结果表明这个恒等元是唯一的．

**1.26 加法恒等元唯一**

向量空间有唯一的加法恒等元．

**证明** 设 $0$ 和 $0'$ 是同一个向量空间 $V$ 的加法恒等元．那么

$$0' = 0' + 0 = 0 + 0' = 0,$$

其中第一个等号成立是因为 $0$ 是加法恒等元，第二个等号成立是因为可交换性，第三个等号成立是因为 $0'$ 是加法恒等元．于是 $0' = 0$，证明了 $V$ 仅有一个加法恒等元． ■

向量空间中的每个元素 $v$ 都存在加法逆元 $w$，它是向量空间中使得 $v + w = 0$ 成立的元素．接下来的结果表明，向量空间中的每个元素都只有一个加法逆元．

**1.27 加法逆元唯一**

向量空间里的每个元素都有唯一的加法逆元．

**证明** 设 $V$ 是向量空间．令 $v \in V$．假设 $w$ 和 $w'$ 都是 $v$ 的加法逆元．那么

$$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w',$$

由此 $w = w'$，命题得证． ■

因为加法逆元是唯一的，所以下面的记号现在就有意义了．

**1.28 记号：$-v$、$w - v$**

令 $v, w \in V$，那么

- $-v$ 表示 $v$ 的加法逆元；
- $w - v$ 定义为 $w + (-v)$.

本书中几乎所有的结果都涉及向量空间．为了避免频繁地重申 $V$ 是向量空间，我们现在作出以下声明，便可一劳永逸：

**1.29 记号：$V$**

在本书的剩余部分中，$V$ 表示 $\mathbf{F}$ 上的向量空间.

在接下来的结果中，等式左侧的 0 表示的是标量（数 $0 \in \mathbf{F}$），等式右侧的 0 表示的是向量（$V$ 的加法恒等元）.

**1.30 向量与数 0 相乘**

对于每个 $v \in V$，都有 $0v = 0$.

**证明** 对于 $v \in V$，我们有

$$0v = (0+0)v = 0v + 0v.$$

在此式两边都加上 $0v$ 的加法逆元就得到 $0 = 0v$，命题得证. ■

在接下来的结果中，0 表示 $V$ 的加法恒等元．虽然 1.30 和 1.31 的证明是相似的，但它们并不完全相同．更准确地说，1.30 表明标量 0 和任何向量的乘积等于向量 0，而 1.31 表明任何标量和向量 0 的乘积等于向量 0.

1.30 中的结果涉及 $V$ 的加法恒等元和标量乘法．在向量空间的定义中，只有分配性质将向量加法和标量乘法关联起来．因此，在 1.30 的证明中必须使用分配性质.

**1.31 数与向量 0 相乘**

对于每个 $a \in \mathbf{F}$，都有 $a0 = 0$.

**证明** 对于 $a \in \mathbf{F}$，我们有

$$a0 = a(0+0) = a0 + a0.$$

在此式两边都加上 $a0$ 的加法逆元就得到 $0 = a0$，命题得证. ■

现在我们证明，如果把 $V$ 中的一个元素乘以标量 $-1$，就会得到该元素的加法逆元.

**1.32 向量与数 −1 相乘**

对于每个 $v \in V$，都有 $(-1)v = -v$.

**证明** 对于 $v \in V$，我们有

$$v + (-1)v = 1v + (-1)v = \big(1 + (-1)\big)v = 0v = 0.$$

这个式子说明 $(-1)v$ 加上 $v$ 会得到 0．那么 $(-1)v$ 就是 $v$ 的加法逆元，命题得证. ■

## 习题 1B

**1** 证明：$-(-v) = v$ 对任一 $v \in V$ 都成立.

**2** 设 $a \in \mathbf{F}$，$v \in V$ 且 $av = 0$，证明：$a = 0$ 或 $v = 0$.

**3** 设 $v, w \in V$，解释为什么存在唯一的 $x \in V$ 使得 $v + 3x = w$.

**4** 空集不是向量空间．对于列写在向量空间定义（1.20）中的要求，空集仅不满足其中的一条．是哪一条？

**5** 证明：在向量空间的定义（1.20）中，加法逆元条件可以替换成这一条件——

$$0v = 0 \text{ 对所有 } v \in V \text{ 成立}.$$

这里，左侧的 0 是数 0，而右侧的 0 是 $V$ 的加法恒等元.

注 在定义中“条件可以替换”，指原来的条件换成新条件后，满足定义的对象还是原来那些.

**6** 令 $\infty$ 和 $-\infty$ 表示不在 $\mathbf{R}$ 中的两个不同对象. 根据记号就能猜到 $\mathbf{R}\cup\{\infty,-\infty\}$ 上的加法和标量乘法可以如何定义. 具体而言，两个实数的和与积照常定义，而对于 $t\in\mathbf{R}$，我们定义：

$$t\infty=\begin{cases}-\infty, & \text{若 } t<0,\\ 0, & \text{若 } t=0,\\ \infty, & \text{若 } t>0;\end{cases}\qquad t(-\infty)=\begin{cases}\infty, & \text{若 } t<0,\\ 0, & \text{若 } t=0,\\ -\infty, & \text{若 } t>0;\end{cases}$$

以及

$$\begin{aligned}t+\infty&=\infty+t=\infty+\infty=\infty,\\ t+(-\infty)&=(-\infty)+t=(-\infty)+(-\infty)=-\infty,\\ \infty+(-\infty)&=(-\infty)+\infty=0.\end{aligned}$$

具有这样的加法和标量乘法运算的 $\mathbf{R}\cup\{\infty,-\infty\}$ 是 $\mathbf{R}$ 上的向量空间吗？解释一下.

**7** 设 $S$ 是非空集合，令 $V^S$ 表示从 $S$ 到 $V$ 的函数构成的集合. 在 $V^S$ 上定义自然的加法和标量乘法，并证明具有这些定义的 $V^S$ 是向量空间.

**8** 设 $V$ 是实向量空间.

- $V$ 的**复化**（**complexification**）记为 $V_{\mathbf{C}}$，等于 $V\times V$. $V_{\mathbf{C}}$ 中的元素为有序对 $(u,v)$，其中 $u,v\in V$，不过我们把它写作 $u+\mathrm{i}v$.
- $V_{\mathbf{C}}$ 上的加法定义为

$$(u_1+\mathrm{i}v_1)+(u_2+\mathrm{i}v_2)=(u_1+u_2)+\mathrm{i}(v_1+v_2)$$

  对所有 $u_1,v_1,u_2,v_2\in V$ 成立.
- $V_{\mathbf{C}}$ 上的复标量乘法定义为

$$(a+b\mathrm{i})(u+\mathrm{i}v)=(au-bv)+\mathrm{i}(av+bu)$$

  对所有 $a,b\in\mathbf{R}$ 和所有 $u,v\in V$ 成立.

证明具有如上加法和标量乘法定义的 $V_{\mathbf{C}}$ 是复向量空间.

注 将 $u\in V$ 等同于 $u+\mathrm{i}0$ 从而将 $V$ 视为 $V_{\mathbf{C}}$ 的子集. 这样一来，由 $V$ 构造 $V_{\mathbf{C}}$ 就可以视作由 $\mathbf{R}^n$ 构造 $\mathbf{C}^n$ 的推广.

# 1C 子空间

通过考虑子空间，我们可以大大扩充向量空间的例子.

**1.33 定义：子空间（subspace）**

如果 $V$ 的子集 $U$ 是与 $V$ 具有相同的加法恒等元、加法和标量乘法运算的向量空间，那么 $U$ 就称为 $V$ 的**子空间**.

接下来的结果给出了检验向量空间的子集是否为子空间的最简单方法.

有人使用术语"线性子空间"（**linear subspace**），这和子空间意思一样.

**1.34 子空间的条件**

当且仅当 $V$ 的子集 $U$ 满足以下三个条件时，$U$ 是 $V$ 的子空间.

**加法恒等元（additive identity）**

$0 \in U$.

**对于加法封闭（closed under addition）**

$u, w \in U$ 意味着 $u + w \in U$.

**对于标量乘法封闭（closed under scalar multiplication）**

$a \in \mathbf{F}$ 且 $u \in U$ 意味着 $au \in U$.

**证明** 如果 $U$ 是 $V$ 的子空间，那么根据向量空间的定义，$U$ 必满足上述三个条件.

反之，假设 $U$ 满足上述三个条件. 第一个条件保证了 $V$ 的加法恒等元必在 $U$ 中. 第二个条件保证了 $V$ 上的加法在 $U$ 上是有意义的. 第三个条件保证了 $V$ 上的标量乘法在 $U$ 上是有意义的.

上面的加法恒等元条件，可以用 $U$ 非空来代替（因为取 $u \in U$ 并将其乘 0 意味着 $0 \in U$）. 然而，如果 $V$ 的子集 $U$ 确实是子空间，那么通常证明 $U$ 非空的最快方法就是证明 $0 \in U$.

如果 $u \in U$，那么由上述第三个条件，$-u$【由 1.32 可得其等于 $(-1)u$】也在 $U$ 中. 因此 $U$ 中的每个元素均有在 $U$ 中的加法逆元.

向量空间定义的其余部分，如可结合性和可交换性，对于 $U$ 自动满足，因为它们在更大的空间 $V$ 中成立. 因此，$U$ 是向量空间，进而是 $V$ 的子空间. ■

给定 $V$ 的一个子集，用上述结果中的三个条件即可快速确定它是否为 $V$ 的子空间. 你应该自行验证接下来这个示例中的所有结论.

**1.35 例：子空间**

(a) 如果 $b \in \mathbf{F}$，那么当且仅当 $b = 0$ 时，

$$\{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = 5x_4 + b\}$$

是 $\mathbf{F}^4$ 的子空间.

(b) 定义在区间 $[0,1]$ 上的全体连续实值函数构成的集合是 $\mathbf{R}^{[0,1]}$ 的子空间.

(c) 定义在 $\mathbf{R}$ 上的全体可微实值函数构成的集合是 $\mathbf{R}^{\mathbf{R}}$ 的子空间.

转下页

(d) 当且仅当 $b=0$ 时，定义在区间 $(0,3)$ 上且满足 $f'(2)=b$ 的全体可微实值函数 $f$ 构成的集合是 $\mathbf{R}^{(0,3)}$ 的子空间.

(e) 极限为 0 的所有复数序列所构成的集合是 $\mathbf{C}^{\infty}$ 的子空间.

验证上面的某些结论，可让我们一瞥隐含在微积分中的线性结构. 例如，上面的 (b) 需要用到"两个连续函数的和是连续的"这一结论；又例如，上面的 (d) 需要用到"对于常数 $c$，$cf$ 的导数等于 $c$ 乘以 $f$ 的导数"这一结论.

集合 $\{0\}$ 是 $V$ 的最小子空间，而 $V$ 本身是 $V$ 的最大子空间. 空集不是 $V$ 的子空间，因为子空间必须是向量空间，因此必须包含至少一个元素，即加法恒等元.

$\mathbf{R}^2$ 的子空间恰有 $\{0\}$，$\mathbf{R}^2$ 中所有过原点的直线，以及 $\mathbf{R}^2$ 本身. $\mathbf{R}^3$ 的子空间恰有 $\{0\}$，$\mathbf{R}^3$ 中所有过原点的直线，$\mathbf{R}^3$ 中所有过原点的平面，以及 $\mathbf{R}^3$ 本身. 证明这些是子空间很容易，困难的是证明 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 仅有这些子空间. 当我们在下一章中引进更多工具之后，这项工作会变得更容易些.

## 子空间的和

讨论向量空间时，我们通常只对子空间而不是任意子集感兴趣. 子空间之和的概念会很有用.

子空间的并往往不是子空间（见习题 12），这也就是我们通常讨论和而不是并的原因.

**1.36 定义：子空间的和（sum of subspaces）**

假设 $V_1,\ldots,V_m$ 是 $V$ 的子空间. $V_1,\ldots,V_m$ 的**和**是由 $V_1,\ldots,V_m$ 中元素所有可能的和所构成的集合，记作 $V_1+\cdots+V_m$. 更确切地说，

$$V_1+\cdots+V_m=\{v_1+\cdots+v_m : v_1\in V_1,\ldots,v_m\in V_m\}.$$

♣

我们看几个有关子空间之和的例子.

**1.37 例：$\mathbf{F}^3$ 的子空间之和**

假设 $U$ 是 $\mathbf{F}^3$ 中第二个和第三个坐标等于 0 的所有元素构成的集合，$W$ 是 $\mathbf{F}^3$ 中第一个和第三个坐标等于 0 的所有元素构成的集合：

$$U=\{(x,0,0)\in\mathbf{F}^3 : x\in\mathbf{F}\}, W=\{(0,y,0)\in\mathbf{F}^3 : y\in\mathbf{F}\}.$$

那么，

$$U+W=\{(x,y,0)\in\mathbf{F}^3 : x,y\in\mathbf{F}\},$$

你应自行验证.

**1.38 例：$\mathbf{F}^4$ 的子空间之和**

假设

$$U=\{(x,x,y,y)\in\mathbf{F}^4 : x,y\in\mathbf{F}\}, W=\{(x,x,x,y)\in\mathbf{F}^4 : x,y\in\mathbf{F}\}.$$

转下页

如果使用自然语言而非符号来描述，我们可以说，$U$ 是由 $\mathbf{F}^4$ 中前两个坐标相等且后两个坐标也相等的所有元素构成的集合；类似地，$W$ 是由 $\mathbf{F}^4$ 中前三个坐标相等的所有元素构成的集合.

为求出 $U+W$ 的表达式，考虑 $U$ 中的一个元素 $(a,a,b,b)$ 和 $W$ 中的一个元素 $(c,c,c,d)$，其中 $a,b,c,d\in\mathbf{F}$. 我们有

$$(a,a,b,b)+(c,c,c,d)=(a+c,a+c,b+c,b+d),$$

这表明 $U+W$ 中每个元素的前两个坐标都相等. 因此

$$U+W\subseteq\{(x,x,y,z)\in\mathbf{F}^4 : x,y,z\in\mathbf{F}\}. \tag{1.39}$$

为证明 $\{(x,x,y,z)\in\mathbf{F}^4 : x,y,z\in\mathbf{F}\}$ 同样含于 $U+W$，假设 $x,y,z\in\mathbf{F}$，那么

$$(x,x,y,z)=(x,x,y,y)+(0,0,0,z-y),$$

上式右端第一个向量在 $U$ 中，第二个向量在 $W$ 中. 于是 $(x,x,y,z)\in U+W$，说明式 (1.39) 表示的包含关系反过来也成立. 因此

$$U+W=\{(x,x,y,z)\in\mathbf{F}^4 : x,y,z\in\mathbf{F}\},$$

这表明，$U+W$ 是由 $\mathbf{F}^4$ 中前两个坐标相等的所有元素构成的集合.

---

接下来的结果表明，子空间的和仍是子空间，而且实际上还是包含其中所有求和项的最小子空间（意即，任一子空间，如果包含所有求和项，也就包含这些求和项之和）. [2]

**1.40 子空间的和是包含这些子空间的最小子空间**

假设 $V_1,\dots,V_m$ 是 $V$ 的子空间，那么 $V_1+\dots+V_m$ 是最小的包含 $V_1,\dots,V_m$ 的子空间.

向量空间理论中子空间的和类似于集合论中子集的并集. 给定一向量空间的两个子空间，包含它们的最小子空间是它们的和. 类似地，给定一集合的两个子集，包含它们的最小子集是它们的并集.

**证明** 读者可自行验证 $V_1+\dots+V_m$ 包含加法恒等元 0 且对于加法和标量乘法都是封闭的. 进而 1.34 表明，$V_1+\dots+V_m$ 是 $V$ 的子空间.

$V_1,\dots,V_m$ 都包含于 $V_1+\dots+V_m$ 中（为了看出这点，考虑 $v_1+\dots+v_m$，其中除某个 $v_k$ 不是 0 外剩余各项都是 0）. 反之，每个包含 $V_1,\dots,V_m$ 的子空间都包含 $V_1+\dots+V_m$（因为子空间必须包含其中有限个元素的和）. 由此，$V_1+\dots+V_m$ 是最小的包含 $V_1,\dots,V_m$ 的子空间. ■

## 直和

设 $V_1,\dots,V_m$ 是 $V$ 的子空间. $V_1+\dots+V_m$ 的每个元素都可以被写成这种形式

$$v_1+\dots+v_m,$$

其中每个 $v_k\in V_k$. 我们特别感兴趣的是 $V_1+\dots+V_m$ 中的每个向量都能唯一地用上述形式表示的情况. 这种情况十分重要，所以它获得了专属的名称（直和）以及专属的符号（$\oplus$）.

[2]此处“求和项”指的就是用于求和的子空间.

**1.41 定义：直和（direct sum）、$\oplus$**

设 $V_1,\ldots,V_m$ 是 $V$ 的子空间.

- 如果 $V_1+\cdots+V_m$ 中的每个元素都能用 $v_1+\cdots+v_m$（其中各 $v_k\in V_k$）这种形式唯一地表示出来，则称子空间之和 $V_1+\cdots+V_m$ 为**直和**.
- 如果 $V_1+\cdots+V_m$ 是直和，那么用 $V_1\oplus\cdots\oplus V_m$ 来表示 $V_1+\cdots+V_m$，其中记号 $\oplus$ 表示此处的和是直和.

♣

**1.42 例：两个子空间的直和**

假设 $U$ 是 $\mathbf{F}^3$ 中最后一个坐标等于 0 的所有向量构成的子空间，$W$ 是 $\mathbf{F}^3$ 中前两个坐标等于 0 的所有向量构成的子空间：

$$U=\{(x,y,0)\in\mathbf{F}^3 : x,y\in\mathbf{F}\}, W=\{(0,0,z)\in\mathbf{F}^3 : z\in\mathbf{F}\}.$$

那么 $\mathbf{F}^3=U\oplus W$. 你应自行验证.

**1.43 例：多个子空间的直和**

假设 $V_k$ 是 $\mathbf{F}^n$ 中除了第 $k$ 个坐标外，其余坐标均为 0 的所有向量构成的子空间. 例如，$V_2=\{(0,x,0,\ldots,0)\in\mathbf{F}^n : x\in\mathbf{F}\}$. 那么 $\mathbf{F}^n=V_1\oplus\cdots\oplus V_n$. 你应自行验证.

输入 `\oplus` 即可在 TEX 中打出 $\oplus$.

有时反例能同正例一样增进我们的理解.

**1.44 例：一个不是直和的和**

假设

$$\begin{aligned}V_1&=\{(x,y,0)\in\mathbf{F}^3 : x,y\in\mathbf{F}\},\\V_2&=\{(0,0,z)\in\mathbf{F}^3 : z\in\mathbf{F}\},\\V_3&=\{(0,y,y)\in\mathbf{F}^3 : y\in\mathbf{F}\}.\end{aligned}$$

那么 $\mathbf{F}^3=V_1+V_2+V_3$，因为每个向量 $(x,y,z)\in\mathbf{F}^3$ 都可以写成

$$(x,y,z)=(x,y,0)+(0,0,z)+(0,0,0),$$

上述等式右端的第一个向量在 $V_1$ 中，第二个向量在 $V_2$ 中，第三个向量在 $V_3$ 中.

然而，$\mathbf{F}^3$ 不等于 $V_1,V_2,V_3$ 的直和，因为将向量 $(0,0,0)$ 写成和式 $v_1+v_2+v_3$（其中各 $v_k\in V_k$）的方式不止一种. 具体地说，我们有

$$(0,0,0)=(0,1,0)+(0,0,1)+(0,-1,-1),$$

并且显然也有

$$(0,0,0)=(0,0,0)+(0,0,0)+(0,0,0),$$

上面两个等式右端的第一个向量在 $V_1$ 中，第二个向量在 $V_2$ 中，第三个向量在 $V_3$ 中. 进而 $V_1+V_2+V_3$ 不是直和.

直和的定义要求每个向量都能唯一地表示成一个适当的和. 接下来的结果表明, 在判断子空间的和是否为直和时，我们只需考虑是否能将 0 唯一地表示成一个适当的和.

符号 $\oplus$ 由一个圆圈和其中的加号组成，它提醒我们正在讨论一种特殊类型的子空间之和——直和中的每个元素只能唯一地由给定子空间的元素之和表示.

**1.45　直和的条件**

假定 $V_1, \ldots, V_m$ 是 $V$ 的子空间. 那么 $V_1 + \cdots + V_m$ 是直和，当且仅当用 $v_1 + \cdots + v_m$（其中各 $v_k \in V_k$）表示 0 的唯一方式是将每个 $v_k$ 都取 0.

**证明**　首先假定 $V_1 + \cdots + V_m$ 是直和. 那么直和的定义就表明用 $v_1 + \cdots + v_m$（其中各 $v_k \in V_k$）表示 0 的唯一方式是将每个 $v_k$ 都取 0.

现在假设用 $v_1 + \cdots + v_m$（其中各 $v_k \in V_k$）表示 0 的唯一方式是将每个 $v_k$ 都取 0. 为了说明 $V_1 + \cdots + V_m$ 是直和，令 $v \in V_1 + \cdots + V_m$，则对某 $v_1 \in V_1, \ldots, v_m \in V_m$，有

$$v = v_1 + \cdots + v_m.$$

为了说明这个表示法是唯一的，假设对 $u_1 \in V_1, \ldots, u_m \in V_m$，我们也有

$$v = u_1 + \cdots + u_m,$$

将这两个式子相减，我们就得到

$$0 = (v_1 - u_1) + \cdots + (v_m - u_m).$$

因为 $v_1 - u_1 \in V_1, \ldots, v_m - u_m \in V_m$，所以上式表明，每个 $v_k - u_k$ 都等于 0. 进而 $v_1 = u_1, \ldots, v_m = u_m$，唯一性得证. ■

接下来的结果给出了检验两个子空间的和是否为直和的一个简单条件.

下面使用的符号 $\Longleftrightarrow$ 意思是“当且仅当”；也可以认为这个符号是“等价于”的意思.

**1.46　两个子空间的直和**

假定 $U$ 和 $W$ 是 $V$ 的子空间. 那么

$$U + W \text{ 是直和} \Longleftrightarrow U \cap W = \{0\}.$$

**证明**　首先假定 $U + W$ 是直和. 如果 $v \in U \cap W$，那么 $0 = v + (-v)$（其中 $v \in U$ 且 $-v \in W$）. 因为 0 要被 $U$ 中向量与 $W$ 中向量的和唯一表示，所以我们得到 $v = 0$. 由此 $U \cap W = \{0\}$，我们完成了一个方向的证明.

为了证明另一方向，现在假设 $U \cap W = \{0\}$. 为了证明 $U + W$ 是直和，假设 $u \in U, w \in W$ 且有

$$0 = u + w.$$

为了完成证明，我们只需说明 $u = w = 0$（根据 1.45）. 上式表明 $u = -w \in W$. 由此 $u \in U \cap W$. 因此 $u = 0$，结合上式又可得 $w = 0$，这样就完成了证明. ■

上面的结果只适用于两个子空间的情况. 当考虑有两个以上子空间的直和的问题时, 仅仅检验每一对子空间只交于 0 处是不够的. 想想例 1.44 就能明白: 在这个反例中, 我们有 $V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = \{0\}$.

子空间的和类似于子集的并集. 类似地, 子空间的直和类似于子集的不相交并. 向量空间的两个子空间不可能不相交, 因为它们都包含 0. 所以至少在两个子空间这种情形下, "不相交"被替换成"交集为 $\{0\}$".

## 习题 1C

**1** 对于 $\mathbf{F}^3$ 的下列各子集, 判断其是否为 $\mathbf{F}^3$ 的子空间:

(a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$

(b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$

(c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$

(d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$

**2** 验证例 1.35 中关于子空间的所有结论.

**3** 证明在区间 $(-4, 4)$ 上的满足 $f'(-1) = 3f(2)$ 的可微实值函数 $f$ 的集合是 $\mathbf{R}^{(-4,4)}$ 的子空间.

**4** 设 $b \in \mathbf{R}$, 证明: 在区间 $[0, 1]$ 上的满足 $\int_0^1 f = b$ 的连续实值函数 $f$ 的集合是 $\mathbf{R}^{[0,1]}$ 的子空间, 当且仅当 $b = 0$.

**5** $\mathbf{R}^2$ 是不是复向量空间 $\mathbf{C}^2$ 的子空间?

**6** (a) $\{(a, b, c) \in \mathbf{R}^3 : a^3 = b^3\}$ 是不是 $\mathbf{R}^3$ 的子空间?

(b) $\{(a, b, c) \in \mathbf{C}^3 : a^3 = b^3\}$ 是不是 $\mathbf{C}^3$ 的子空间?

**7** 证明或给出一反例: 如果 $U$ 是 $\mathbf{R}^2$ 的非空子集, 满足对加法封闭和对"取加法逆元"封闭 (意即只要 $u \in U$ 就有 $-u \in U$), 那么 $U$ 是 $\mathbf{R}^2$ 的子空间.

**8** 给出一例: $\mathbf{R}^2$ 的非空子集 $U$, 满足对标量乘法封闭, 但不是 $\mathbf{R}^2$ 的子空间.

**9** 函数 $f : \mathbf{R} \to \mathbf{R}$ 称为**周期的** (**periodic**), 如果存在一正数 $p$ 使得 $f(x) = f(x + p)$ 对所有 $x \in \mathbf{R}$ 成立. 从 $\mathbf{R}$ 到 $\mathbf{R}$ 的周期函数构成的集合是不是 $\mathbf{R}^{\mathbf{R}}$ 的子空间? 请作解释.

**10** 设 $V_1$ 和 $V_2$ 是 $V$ 的子空间, 证明: 交集 $V_1 \cap V_2$ 是 $V$ 的子空间.

**11** 证明: $V$ 的任意一族子空间的交集都是 $V$ 的子空间.

**12** 证明: $V$ 的两个子空间的并集是 $V$ 的子空间, 当且仅当其中一个包含于另一个.

**13** 证明: $V$ 的三个子空间的并集是 $V$ 的子空间, 当且仅当其中一个包含另两个.

注 这道习题比习题 12 难多了, 也许是因为如果我们把 $\mathbf{F}$ 换成仅含两个元素的域, 这道习题的结论就不对了.

**14** 设

$$U = \{(x, -x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad 与 \quad W = \{(x, x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\}.$$

用符号描述 $U + W$, 并给出不使用符号的描述[3].

**15** 设 $U$ 是 $V$ 的子空间, 那么 $U + U$ 是什么?

**16** $V$ 的"子空间加法"运算[4]可交换吗? 换句话说, 如果 $U$ 和 $W$ 是 $V$ 的子空间, 是否有 $U + W = W + U$?

---

[3]不使用符号, 说的是用自然语言来描述, 就像在例 1.38 里那样.

[4]即子空间之和. 可参看索引项"addition of subspaces".

**17** $V$ 的“子空间加法”运算可结合吗？换句话说，如果 $V_1, V_2, V_3$ 是 $V$ 的子空间，是否有

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)?$$

**18** $V$ 的“子空间加法”运算有没有加法恒等元？哪些子空间有加法逆元？

**19** 证明或给出一反例：如果是 $V_1, V_2, U$ 是 $V$ 的子空间，并满足

$$V_1 + U = V_2 + U,$$

则有 $V_1 = V_2$.

**20** 设

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\},$$

求 $\mathbf{F}^4$ 的一个子空间 $W$，使得 $\mathbf{F}^4 = U \oplus W$.

**21** 设

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\},$$

求 $\mathbf{F}^5$ 的一个子空间 $W$，使得 $\mathbf{F}^5 = U \oplus W$.

**22** 设

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\},$$

求 $\mathbf{F}^5$ 的三个都不为 $\{0\}$ 的子空间 $W_1, W_2, W_3$，使得 $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

**23** 证明或给出一反例：如果 $V_1, V_2, U$ 是 $V$ 的子空间，并满足

$$V = V_1 \oplus U \quad 且 \quad V = V_2 \oplus U,$$

则有 $V_1 = V_2$.

提示　在尝试发现线性代数中的一个猜测是对是错时，先在 $\mathbf{F}^2$ 里试试，经常是很有帮助的.

**24** 函数 $f : \mathbf{R} \to \mathbf{R}$ 被称为**偶的**（**even**），若

$$f(-x) = f(x)$$

对所有 $x \in \mathbf{R}$ 成立．函数 $f : \mathbf{R} \to \mathbf{R}$ 被称为**奇的**（**odd**），若

$$f(-x) = -f(x)$$

对所有 $x \in \mathbf{R}$ 成立．令 $V_e$ 代表 $\mathbf{R}$ 上的实值偶函数构成的集合，令 $V_o$ 代表 $\mathbf{R}$ 上的实值奇函数构成的集合，证明：$\mathbf{R}^{\mathbf{R}} = V_e \oplus V_o$.

# 第 2 章 有限维向量空间

在上一章中我们学习了向量空间. 线性代数关注的不是任意的向量空间，而是在本章中介绍的有限维向量空间.

我们从考虑向量组[1]的线性组合开启本章，这引导我们得到线性无关性这个关键概念. 线性相关性引理将成为我们最有力的工具之一.

向量空间的基是该向量空间中这样一个向量组：它既小到可以满足线性无关性，又大到使其线性组合能填满这整个空间. 我们将看到，向量空间的每个基都有相等的长度，这使我们能定义向量空间的维数.

本章以两个子空间之和的维数公式收尾.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 代表 $\mathbf{F}$ 上的向量空间.

图为坐落于新泽西州普林斯顿（Princeton, New Jersey）的普林斯顿高等研究院（Institute for Advanced Study）主楼. 保罗・哈尔莫斯（Paul Halmos，1916-2006）在这座大楼里写出了第一本现代的线性代数书. 哈尔莫斯所著的线性代数书于 1942 年出版（第二版于 1958 年出版），此书的书名与本章的标题相同.

[1] 本翻译中，a list of vectors 视情况翻译为"一组向量"或"一个向量组"（由于其本身的多义性）， lists of vectors 统一翻译为向量组.

# 2A 张成空间和线性无关性

我们一直把由数构成的组写成用圆括号括起来的数. 之后，对于 $\mathbf{F}^n$ 中的元素，我们将沿用这种写法，例如 $(2,-7,8) \in \mathbf{F}^3$. 然而，现在我们需要考虑向量组（这里的向量可能是 $\mathbf{F}^n$ 或其他向量空间的元素）. 为了避免混淆，我们书写向量组时通常不用圆括号括起来. 例如，$(4,1,6),(9,5,7)$ 是 $\mathbf{R}^3$ 中长度为 2 的向量组.

**2.1 记号：向量组**

书写向量组时，我们通常不用圆括号括起来.

## 线性组合和张成空间

将一个向量组中的向量乘以标量再求和的结果，就称为该向量组的线性组合. 下面是正式的定义.

**2.2 定义：线性组合（linear combination）**

$V$ 中向量组 $v_1,\ldots,v_m$ 的**线性组合**[2]是形如

$$a_1v_1+\cdots+a_mv_m$$

的向量，其中 $a_1,\ldots,a_m \in \mathbf{F}$.

---

**2.3 例：$\mathbf{R}^3$ 中的线性组合**

- $(17,-4,2)$ 是 $\mathbf{R}^3$ 中长度为 2 的向量组 $(2,1,-3),(1,-2,4)$ 的线性组合，因为

$$(17,-4,2)=6(2,1,-3)+5(1,-2,4).$$

- $(17,-4,5)$ 不是 $\mathbf{R}^3$ 中长度为 2 的向量组 $(2,1,-3),(1,-2,4)$ 的线性组合，因为不存在数 $a_1,a_2 \in \mathbf{F}$ 使得下式成立

$$(17,-4,5)=a_1(2,1,-3)+a_2(1,-2,4).$$

换言之，线性方程组

$$\begin{aligned}17&=2a_1+a_2\\-4&=a_1-2a_2\\5&=-3a_1+4a_2\end{aligned}$$

无解（你应自行验证）.

---

[2]作者有时说 “linear combination of **the list**”（见英文原文定义 2.2 的上一行），有时说 “linear combination of the previous ones”（ones 指代 **vectors**，见 2.22 的英文原文），即“向量的线性组合”和“向量组的线性组合”的说法都可能出现. 只要明白都指的是将一个向量空间中的有限个向量分别乘以标量再求和即可.

**2.4 定义：张成空间（span）**

$V$ 中向量组 $v_1, \ldots, v_m$ 的所有线性组合所构成的集合称为 $v_1, \ldots, v_m$ 的**张成空间**[3]，记作 $\operatorname{span}(v_1, \ldots, v_m)$. 换言之，

$$\operatorname{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \ldots, a_m \in \mathbf{F}\}.$$

定义空向量组 ( ) 的张成空间为 $\{0\}$.

**2.5 例：张成空间**

上一个例子说明，在 $\mathbf{F}^3$ 中，

- $(17, -4, 2) \in \operatorname{span}\big((2, 1, -3), (1, -2, 4)\big)$；
- $(17, -4, 5) \notin \operatorname{span}\big((2, 1, -3), (1, -2, 4)\big)$.

**2.6 向量组的张成空间是最小的包含组中所有向量的子空间**

$V$ 中向量组的张成空间是最小的包含这向量组中所有向量的 $V$ 的子空间.

**证明** 假设 $v_1, \ldots, v_m$ 是 $V$ 中的向量组.

首先我们说明 $\operatorname{span}(v_1, \ldots, v_m)$ 是 $V$ 的子空间. 加法恒等元属于 $\operatorname{span}(v_1, \ldots, v_m)$，因为

> 有些数学家使用术语“**线性张成空间**”（**linear span**），它和张成空间的意思一样.

$$0 = 0v_1 + \cdots + 0v_m.$$

而且，$\operatorname{span}(v_1, \ldots, v_m)$ 对加法封闭，因为

$$(a_1 v_1 + \cdots + a_m v_m) + (c_1 v_1 + \cdots + c_m v_m) = (a_1 + c_1)v_1 + \cdots + (a_m + c_m)v_m.$$

此外，$\operatorname{span}(v_1, \ldots, v_m)$ 对标量乘法封闭，因为

$$\lambda(a_1 v_1 + \cdots + a_m v_m) = \lambda a_1 v_1 + \cdots + \lambda a_m v_m.$$

于是，$\operatorname{span}(v_1, \ldots, v_m)$ 就是 $V$ 的子空间（由 1.34 可得）.

每个 $v_k$ 都是 $v_1, \ldots, v_m$ 的线性组合（为了说明这点，在 2.2 的公式中，令 $a_k = 1$ 且其他各 $a$ 都等于 0）. 于是 $\operatorname{span}(v_1, \ldots, v_m)$ 包含每个 $v_k$. 反之，因为子空间对于标量乘法和加法都是封闭的，所以 $V$ 的每个包含所有 $v_k$ 的子空间都包含 $\operatorname{span}(v_1, \ldots, v_m)$. 由此，$\operatorname{span}(v_1, \ldots, v_m)$ 是最小的包含所有向量 $v_1, \ldots, v_m$ 的 $V$ 的子空间. ■

**2.7 定义：张成（spans）**

如果 $\operatorname{span}(v_1, \ldots, v_m)$ 等于 $V$，我们就说 $v_1, \ldots, v_m$ **张成**[4] $V$.

[3] 原文 span，是个名词. 在后文中，和“线性组合”一样，“（一组）向量的张成空间”和“向量组的张成空间”的说法都可能出现，它们所指相同，都是一个向量空间中的有限个向量分别乘以标量再求和的所有结果构成的集合.

[4] 原文 spans，动词的第三人称单数形式.

**2.8 例：一个张成 $\mathbf{F}^n$ 的组**

假设 $n$ 为正整数. 我们想要证明

$$(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$$

张成 $\mathbf{F}^n$. 上述向量组中的第 $k$ 个向量的第 $k$ 个坐标是 1，而其他坐标都是 0.

假设 $(x_1,\ldots,x_n)\in\mathbf{F}^n$，那么

$$(x_1,\ldots,x_n)=x_1(1,0,\ldots,0)+x_2(0,1,0,\ldots,0)+\cdots+x_n(0,\ldots,0,1).$$

于是，$(x_1,\ldots,x_n)\in\operatorname{span}\big((1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\big)$.

现在我们给出线性代数中一个关键的定义.

**2.9 定义：有限维向量空间（finite-dimensional vector space）**

如果一个向量空间可由其中某个向量组张成，则称该向量空间是**有限维的**.

上面的例 2.8 表明，对于每个正整数 $n$，$\mathbf{F}^n$ 都是有限维的向量空间.

多项式的定义对你来说无疑是很熟悉的.

回想一下：根据定义，每个组都具有有限长度.

**2.10 定义：多项式（polynomial）、$\mathcal{P}(\mathbf{F})$**

- 对于函数 $p:\mathbf{F}\to\mathbf{F}$，如果存在 $a_0,\ldots,a_m\in\mathbf{F}$ 使得对所有 $z\in\mathbf{F}$ 都有

$$p(z)=a_0+a_1z+a_2z^2+\cdots+a_mz^m,$$

  则称 $p$ 为系数在 $\mathbf{F}$ 中的**多项式**.
- $\mathcal{P}(\mathbf{F})$ 是系数在 $\mathbf{F}$ 中的全体多项式所构成的集合.

带有通常的加法和标量乘法运算的 $\mathcal{P}(\mathbf{F})$ 是 $\mathbf{F}$ 上的向量空间（你应自行验证）. 因此，$\mathcal{P}(\mathbf{F})$ 是 $\mathbf{F}^{\mathbf{F}}$（全体由 $\mathbf{F}$ 到 $\mathbf{F}$ 的函数所构成的向量空间）的子空间.

如果一个多项式（视为一个由 $\mathbf{F}$ 到 $\mathbf{F}$ 的函数）可由两组系数表示，那么将其中一种表示法减去另一种，就能得到这样一个多项式——$\mathbf{F}$ 上的恒等于 0 的函数，因此它的所有系数都为 0（如果你对这个事实不熟悉，暂且相信它是对的，之后我们会证明它——参见 4.8）. **结论：**一个多项式的系数由该多项式唯一决定. 于是，下述定义唯一地规定了多项式的次数.

**2.11 定义：多项式的次数（degree of a polynomial）、$\deg p$**

- 对于多项式 $p\in\mathcal{P}(\mathbf{F})$，如果存在 $a_0,a_1,\ldots,a_m\in\mathbf{F}$ 且 $a_m\neq 0$ 使得对每个 $z\in\mathbf{F}$，都有

$$p(z)=a_0+a_1z+\cdots+a_mz^m,$$

  那么就说 $p$ 的**次数**是 $m$.
- 规定恒等于 0 的多项式的次数为 $-\infty$.
- 多项式 $p$ 的次数记为 $\deg p$.

在下面的定义中，我们约定 $-\infty<m$，这意味着多项式 0 属于 $\mathcal{P}_m(\mathbf{F})$.

**2.12 记号：$\mathcal{P}_m(\mathbf{F})$**

对于非负整数 $m$，$\mathcal{P}_m(\mathbf{F})$ 表示系数在 $\mathbf{F}$ 中且次数不高于 $m$ 的所有多项式所构成的集合. ♣

如果 $m$ 是非负整数，那么 $\mathcal{P}_m(\mathbf{F}) = \operatorname{span}(1, z, \ldots, z^m)$（此处我们令 $z^k$ 表示一个函数，这有点滥用记号[5]）. 于是，对于每个非负整数 $m$，$\mathcal{P}_m(\mathbf{F})$ 都是有限维向量空间.

**2.13 定义：无限维向量空间（infinite-dimensional vector space）**

如果一个向量空间不是有限维的，就称它是**无限维的**. ♣

**2.14 例：$\mathcal{P}(\mathbf{F})$ 是无限维的**

考虑 $\mathcal{P}(\mathbf{F})$ 中的任意一组元素. 令 $m$ 表示这个组中多项式的最高次数. 那么这个组的张成空间中每个多项式的次数都不超过 $m$. 从而 $z^{m+1}$ 不在这个组的张成空间里. 因此没有组能张成 $\mathcal{P}(\mathbf{F})$，于是 $\mathcal{P}(\mathbf{F})$ 是无限维的.

## 线性无关性

假设 $v_1, \ldots, v_m \in V$ 且 $v \in \operatorname{span}(v_1, \ldots, v_m)$. 根据张成空间的定义，存在 $a_1, \ldots, a_m \in \mathbf{F}$，使得

$$v = a_1 v_1 + \cdots + a_m v_m.$$

考虑这个问题：上式中标量的选取是否唯一？假设 $c_1, \ldots, c_m$ 是另一组标量且满足

$$v = c_1 v_1 + \cdots + c_m v_m.$$

将上述两式相减，我们有

$$0 = (a_1 - c_1)v_1 + \cdots + (a_m - c_m)v_m.$$

于是，我们将 0 写成了 $(v_1, \ldots, v_m)$ 的线性组合. 如果这样表达 0 的唯一方式，是将线性组合式中的所有标量都取为 0，那么每个 $a_k - c_k$ 都等于 0，这也就意味着每个 $a_k$ 等于相应的 $c_k$（由此可见标量的选取确实是唯一的）. 这种情形很重要，所以我们赋予其特别的名称——线性无关——并在下面定义.

**2.15 定义：线性无关（linearly independent）**

- 对于 $V$ 中的向量组 $v_1, \ldots, v_m$，如果使得

$$a_1 v_1 + \cdots + a_m v_m = 0$$

  成立的 $a_1, \ldots, a_m \in \mathbf{F}$ 的唯一选取方式是 $a_1 = \cdots = a_m = 0$，那么称该向量组为**线性无关的**.
- 规定空向量组 ( ) 也是线性无关的. ♣

上面的推导说明，$v_1, \ldots, v_m$ 是线性无关的，当且仅当 $\operatorname{span}(v_1, \ldots, v_m)$ 中的每个向量都只能唯一地表示成 $v_1, \ldots, v_m$ 的线性组合.

[5] 为什么说“这有点滥用记号”呢？因为 $z^k$ 这个记号本来是用于表示数 $z \in \mathbf{F}$ 的 $k$ 次幂，这里却用来指代函数 $f: z \mapsto z^k, z \in \mathbf{F}$. 这样其实有损于严谨性和准确性，但也方便了表达和理解，还能避免冗长的描述.

**2.16 例：线性无关组**

(a) 为了说明组 $(1,0,0,0),(0,1,0,0),(0,0,1,0)$ 在 $\mathbf{F}^4$ 中是线性无关的，假设 $a_1,a_2,a_3\in\mathbf{F}$ 且

$$a_1(1,0,0,0)+a_2(0,1,0,0)+a_3(0,0,1,0)=(0,0,0,0).$$

由此

$$(a_1,a_2,a_3,0)=(0,0,0,0).$$

所以 $a_1=a_2=a_3=0$. 于是组 $(1,0,0,0),(0,1,0,0),(0,0,1,0)$ 在 $\mathbf{F}^4$ 中是线性无关的.

(b) 设 $m$ 是非负整数. 为了说明组 $1,z,\dots,z^m$ 在 $\mathcal{P}(\mathbf{F})$ 中是线性无关的，假设 $a_0,a_1,\dots,a_m\in\mathbf{F}$ 且

$$a_0+a_1z+\cdots+a_mz^m=0,$$

这里我们将式子两侧都看成 $\mathcal{P}(\mathbf{F})$ 中的元素. 那么对于所有 $z\in\mathbf{F}$，都有

$$a_0+a_1z+\cdots+a_mz^m=0.$$

正如前面讨论过的那样（由 4.8 亦可得），这意味着 $a_0=a_1=\cdots=a_m=0$. 于是 $1,z,\dots,z^m$ 是 $\mathcal{P}(\mathbf{F})$ 中的线性无关组.

(c) 向量空间中长度为 1 的向量组是线性无关的，当且仅当组中的向量不是 0.

(d) 向量空间中长度为 2 的向量组是线性无关的，当且仅当组中任一向量都不是另一个向量的标量倍[6].

如果从一个线性无关组中移除某些向量，余下的向量构成的向量组仍线性无关. 你应自行验证这一点.

**2.17 定义：线性相关（linearly dependent）**

- 如果 $V$ 中的一个向量组不是线性无关的，就称它是**线性相关的**.
- 换言之，对于 $V$ 中的向量组 $v_1,\dots,v_m$，如果存在不全为 0 的 $a_1,\dots,a_m\in\mathbf{F}$ 使得 $a_1v_1+\cdots+a_mv_m=0$，那么该向量组是线性相关的.

♣

**2.18 例：线性相关组**

- $(2,3,1),(1,-1,2),(7,3,8)$ 在 $\mathbf{F}^3$ 中是线性相关的，因为

$$2(2,3,1)+3(1,-1,2)+(-1)(7,3,8)=(0,0,0).$$

- 组 $(2,3,1),(1,-1,2),(7,3,c)$ 在 $\mathbf{F}^3$ 中是线性相关的，当且仅当 $c=8$（你应自行验证）.
- 如果 $V$ 中一个向量组中的某个向量是其他向量的线性组合，那么这个向量组是线性相关的.【证明：先将这个组中的一个向量写成等于其他向量的线性组合的形式，然后将那个向量移到等式的另一侧（这样它前面会乘以 $-1$）即可.】
- $V$ 中每个包含向量 0 的向量组都是线性相关的.（这是上一点的特例）

[6]一向量的标量倍就是该向量和某个标量相乘的结果.

接下来这条引理是一个绝佳的工具．它是说，给定一个线性相关的向量组，其中就有某个向量处于排在其之前的向量的张成空间里．进而，我们可从该组中去掉那个向量，而不改变该组的张成空间．

**2.19 线性相关性引理（linear dependence lemma）**

设 $v_1, \ldots, v_m$ 是 $V$ 中的线性相关组．那么存在 $k \in \{1, 2, \ldots, m\}$ 满足

$$v_k \in \operatorname{span}(v_1, \ldots, v_{k-1}).$$

进而，如果 $k$ 满足上述条件且从 $v_1, \ldots, v_m$ 中移除第 $k$ 项，那么剩余向量组成的向量组的张成空间仍等于 $\operatorname{span}(v_1, \ldots, v_m)$.

**证明** 因为向量组 $v_1, \ldots, v_m$ 是线性相关的，所以存在不全为 0 的数 $a_1, \ldots, a_m \in \mathbf{F}$ 使得

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

令 $k$ 是 $\{1, \ldots, m\}$ 中使 $a_k \neq 0$ 的最大者．那么

$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1},$$

这证明了 $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$.

现在假设 $k$ 是 $\{1, \ldots, m\}$ 中任一使得 $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$ 成立的元素.令 $b_1, \ldots, b_{k-1} \in \mathbf{F}$ 满足

$$v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}. \tag{2.20}$$

假设 $u \in \operatorname{span}(v_1, \ldots, v_m)$．那么存在 $c_1, \ldots, c_m \in \mathbf{F}$ 使得

$$u = c_1 v_1 + \cdots + c_m v_m.$$

在上述等式中，我们可以将 $v_k$ 替换成式 (2.20) 的右端，所得等式表明 $u$ 处于将 $v_1, \ldots, v_m$ 的第 $k$ 项去掉所得的向量组的张成空间中．于是，从向量组 $v_1, \ldots, v_m$ 去掉第 $k$ 项不改变该向量组的张成空间． ■

如果线性相关性引理中 $k = 1$，那么 $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$ 意味着 $v_1 = 0$，因为 $\operatorname{span}(\,) = \{0\}$．还请注意，如果 $k = 1$，那么线性相关性引理的证明过程的一部分需要作修改．一般而言，书中其余部分的证明将不会关注一些必须考虑的特殊情况，包括长度为 0 的组、子空间 $\{0\}$，以及其余平凡的情形——结论对于这些情形也正确，只是证明过程稍有不同．请务必自行检验一下这些特殊情况.

**2.21 例：线性相关性引理中 $k$ 的最小值**

考虑 $\mathbf{R}^3$ 中的组

$$(1, 2, 3), (6, 5, 4), (15, 16, 17), (8, 9, 7).$$

我们马上就会看到，这个长度为 4 的组是线性相关的．于是线性相关性引理告诉我们，存在 $k \in \{1, 2, 3, 4\}$ 使得组中第 $k$ 个向量是该组中排在它前面的向量的线性组合．我们来看怎么得出 $k$ 的最小值.

转下页

在线性相关性引理中，当且仅当组中第一个向量等于 0 时才可取 $k=1$. 因为 $(1,2,3)$ 不是向量 0，所以对于该组我们不能取 $k=1$.

在线性相关性引理中，当且仅当组中第二个向量是第一个向量的标量倍时才可取 $k=2$. 然而，并不存在 $c\in\mathbf{R}$ 使得 $(6,5,4)=c(1,2,3)$. 于是对于该组我们不能取 $k=2$.

在线性相关性引理中，当且仅当组中第三个向量是前两个向量的线性组合时才可取 $k=3$. 于是对于本例中的组，我们想要知道是否存在 $a,b\in\mathbf{R}$ 使得

$$(15,16,17)=a(1,2,3)+b(6,5,4)$$

成立. 上述方程等价于一个由三个方程组成、含有两个未知数 $a,b$ 的线性方程组. 使用高斯消元法或者合适的软件，我们求出 $a=3,b=2$ 是上述方程的解（你可自行验证）. 于是对于本例中的组， $k=3$ 是满足线性相关性引理的最小 $k$ 值.

---

现在我们得到一个关键结果. 它是说，$V$ 中的线性无关组不会长于 $V$ 中的张成组.

**2.22 线性无关组的长度 $\leq$ 张成组的长度**

在有限维向量空间中，每个线性无关向量组的长度小于或等于每个张成向量组的长度.

**证明** 假设 $u_1,\ldots,u_m$ 是 $V$ 中的线性无关组，并假设 $w_1,\ldots,w_n$ 张成 $V$. 我们需要证明 $m\leq n$. 我们通过下述过程来完成这个证明，其中共有 $m$ 步. 注意，每一步中我们都加入某个 $u$ 且去掉某个 $w$.

**步骤 1**

令 $B$ 为组 $w_1,\ldots,w_n$，其张成空间是 $V$. 在该组开头加入 $u_1$ 可得一个线性相关组（因为 $u_1$ 能被写成 $w_1,\ldots,w_n$ 的线性组合）. 换言之，组

$$u_1,w_1,\ldots,w_n$$

是线性相关的.

于是，由线性相关性引理（2.19），上述组中的某一个向量是该组中排在它前面的向量的线性组合. 向量组 $u_1,\ldots,u_m$ 线性无关，因此 $u_1\neq 0$. 于是， $u_1$ 不属于上述组中排在它前面的向量的张成空间（也就是空组的张成空间 $\{0\}$）. 所以，线性相关性引理表明，我们可以移除某个 $w$ 以使由 $u_1$ 和剩余各 $w$ 构成的新组 $B$（长度为 $n$）仍张成 $V$.

**步骤 $k$（$k=2,\ldots,m$）**

由第 $k-1$ 步得到的组 $B$（长度为 $n$）张成 $V$. 特别地[7]，$u_k$ 在 $B$ 的张成空间里. 于是，将 $u_k$ 紧接在 $u_1,\ldots,u_{k-1}$ 后面插入 $B$ 中，可得一个线性相关组，其长度为 $n+1$. 根据线性相关性引理（2.19），该组中的某一向量处于排在它之前的向量的张成空间中. 又由于 $u_1,\ldots,u_k$ 是线性无关的，这个向量不可能是某个 $u$.

因此，在这一步肯定仍剩下至少一个 $w$. 我们可以从新组（通过将 $u_k$ 插入合适的位置所得）中去除某个 $w$（它是排在它前面若干个向量的线性组合），由此得到新的组 $B$（长度为 $n$），它由 $u_1,\ldots,u_k$ 和剩余各 $w$ 构成，且仍张成 $V$.

[7] "特别地"对应英文"in particular"，其所接的结论隐含于其之前的结论，是其之前的结论的"特殊情况"或"弱化版本". 例如，"所有的猫都是动物. 特别地，所有的黑猫都是动物". 对于此处，$B$ 张成 $V$，而 $u_k$ 属于 $V$，则 $u_k$ 自然在 $B$ 的张成空间里.

在步骤 $m$ 后，我们将所有的 $u$ 都加入了，上述过程结束．在每一步中，当我们往 $B$ 中加入一个 $u$ 时，线性相关性引理就表明肯定有某个 $w$ 将被移除．于是各 $w$ 至少有各 $u$ 那么多． ■

接下来的两个例子展示的是，如何运用上述结果而无需任何计算，来说明某些组不是线性无关的，或说明某些组不能张成一个给定的向量空间．

**2.23 例：$\mathbf{R}^3$ 中长度为 4 的向量组都不是线性无关的**

向量组 $(1,0,0),(0,1,0),(0,0,1)$ 长度为 3 且张成 $\mathbf{R}^3$．于是，$\mathbf{R}^3$ 中长度超过 3 的组都不是线性无关的．

例如，我们现在知道长度为 4 的向量组 $(1,2,3),(4,5,8),(9,6,7),(-3,2,8)$ 在 $\mathbf{R}^3$ 中不是线性无关的．

**2.24 例：长度为 3 的向量组都不能张成 $\mathbf{R}^4$**

向量组 $(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)$ 长度为 4 且在 $\mathbf{R}^4$ 中是线性无关的．于是，长度少于 4 的向量组都不能张成 $\mathbf{R}^4$．

例如，我们现在知道长度为 3 的向量组 $(1,2,3,-5),(4,5,8,3),(9,6,7,-1)$ 不能张成 $\mathbf{R}^4$．

直觉告诉我们，有限维向量空间的各个子空间也应是有限维的．现在我们来证明这个直觉是正确的．

**2.25 有限维的子空间**

有限维向量空间的每个子空间都是有限维的．

**证明** 假设 $V$ 是有限维的，且 $U$ 是 $V$ 的子空间．我们需要证明 $U$ 是有限维的．我们通过多步的构造来完成证明，如下所示．

**步骤 1**

如果 $U=\{0\}$，那么 $U$ 是有限维的，我们的证明完成了．如果 $U\neq\{0\}$，那么选取一非零向量 $u_1\in U$．

**步骤 $k$**

如果 $U=\operatorname{span}(u_1,\dots,u_{k-1})$，那么 $U$ 是有限维的，我们的证明完成了．如果 $U\neq\operatorname{span}(u_1,\dots,u_{k-1})$，那么选取一向量 $u_k\in U$ 使得

$$u_k\notin\operatorname{span}(u_1,\dots,u_{k-1}).$$

经过每一步，只要这个过程还在继续，我们都构造出一组向量，并使得其中没有哪个向量处于排在它前面的向量的张成空间里．于是，根据线性相关性引理（2.19），经过每一步我们都构造出一个线性无关组．这个线性无关组不能长于 $V$ 的任一张成组（由 2.22）．于是，这个过程终会停止，意味着 $U$ 是有限维的． ■

## 习题 2A

**1** 求 $\mathbf{F}^3$ 中的四个不同向量，其张成空间等于

$$\{(x,y,z)\in\mathbf{F}^3 : x+y+z=0\}.$$

**2** 证明或给出一反例：如果 $v_1, v_2, v_3, v_4$ 张成 $V$，那么组

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

也张成 $V$.

**3** 设 $v_1, \ldots, v_m$ 是 $V$ 中的一组向量．对于 $k \in \{1, \ldots, m\}$，令

$$w_k = v_1 + \cdots + v_k.$$

证明 $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(w_1, \ldots, w_m)$.

**4** (a) 证明：向量空间中长度为一的组线性无关，当且仅当组中的该向量不是 0.

(b) 证明：向量空间中长度为二的组线性无关，当且仅当组中两个向量的任一个都不是另一个的标量倍.

**5** 求一数 $t$ 使得

$$(3, 1, 4), (2, -3, 5), (5, 9, t)$$

在 $\mathbf{R}^3$ 中不是线性无关的.

**6** 证明：组 $(2, 3, 1), (1, -1, 2), (7, 3, c)$ 在 $\mathbf{F}^3$ 中线性相关，当且仅当 $c = 8$.

**7** (a) 证明：如果我们将 $\mathbf{C}$ 视为 $\mathbf{R}$ 上的向量空间，那么组 $1 + \mathrm{i}, 1 - \mathrm{i}$ 是线性无关的.

(b) 证明：如果我们将 $\mathbf{C}$ 视为 $\mathbf{C}$ 上的向量空间，那么组 $1 + \mathrm{i}, 1 - \mathrm{i}$ 是线性相关的.

**8** 设 $v_1, v_2, v_3, v_4$ 在 $V$ 中线性无关．证明组

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

也线性无关.

**9** 证明或给出一反例：如果 $v_1, v_2, \ldots, v_m$ 是 $V$ 中的线性无关向量组，那么

$$5v_1 - 4v_2, v_2, v_3, \ldots, v_m$$

线性无关.

**10** 证明或给出一反例：如果 $v_1, v_2, \ldots, v_m$ 是 $V$ 中的线性无关向量组，且 $\lambda \in \mathbf{F}$（$\lambda \neq 0$），那么 $\lambda v_1, \lambda v_2, \ldots, \lambda v_m$ 线性无关.

**11** 证明或给出一反例：如果 $v_1, \ldots, v_m$ 和 $w_1, \ldots, w_m$ 是 $V$ 中线性无关的向量组，那么组 $v_1 + w_1, \ldots, v_m + w_m$ 线性无关.

**12** 设 $v_1, \ldots, v_m$ 在 $V$ 中线性无关，且 $w \in V$．证明：如果 $v_1 + w, \ldots, v_m + w$ 线性相关，那么 $w \in \operatorname{span}(v_1, \ldots, v_m)$.

**13** 设 $v_1, \ldots, v_m$ 在 $V$ 中线性无关，且 $w \in V$．证明：

$$v_1, \ldots, v_m, w \text{ 线性无关} \iff w \notin \operatorname{span}(v_1, \ldots, v_m).$$

**14** 设 $v_1, \ldots, v_m$ 是 $V$ 中一组向量．对于 $k \in \{1, \ldots, m\}$，令

$$w_k = v_1 + \cdots + v_k.$$

证明：组 $v_1, \ldots, v_m$ 线性无关，当且仅当 $w_1, \ldots, w_m$ 线性无关.

**15** 解释为什么在 $\mathcal{P}_4(\mathbf{F})$ 中不存在由六个多项式组成的线性无关组.

**16** 解释为什么由四个多项式构成的组不能张成 $\mathcal{P}_4(\mathbf{F})$.

**17** 证明：$V$ 是无限维的，当且仅当 $V$ 中有一序列 $v_1, v_2, \ldots$ 使得对任一正整数 $m$ 均有 $v_1, \ldots, v_m$ 线性无关.

**18** 证明 $\mathbf{F}^\infty$ 是无限维的.

**19** 证明：由区间 $[0,1]$ 上的所有连续实值函数构成的实向量空间是无限维的.

**20** 设 $p_0, p_1, \ldots, p_m$ 是 $\mathcal{P}_m(\mathbf{F})$ 中的多项式，其满足对任一 $k \in \{0, \ldots, m\}$ 都有 $p_k(2) = 0$. 证明：$p_0, p_1, \ldots, p_m$ 在 $\mathcal{P}_m(\mathbf{F})$ 中不是线性无关的.

## 2B 基

在上一节中，我们讨论了线性无关组和张成组. 现在我们考虑同时具备这两个性质的组，从而将这两个概念融到一起.

**2.26 定义：基（basis）**

$V$ 中线性无关且张成 $V$ 的向量组称为 $V$ 的**基**.

**2.27 例：基**

(a) 向量组 $(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ 是 $\mathbf{F}^n$ 的基，它被称为 $\mathbf{F}^n$ 的**标准基**（**standard basis**）.

(b) 向量组 $(1,2),(3,5)$ 是 $\mathbf{F}^2$ 的基. 注意，这个组的长度是 2，与 $\mathbf{F}^2$ 的标准基的长度一样. 在下一节中我们将看到，这并非巧合.

(c) 向量组 $(1,2,-4),(7,-5,6)$ 在 $\mathbf{F}^3$ 中是线性无关的，但不是 $\mathbf{F}^3$ 的基，因为它不张成 $\mathbf{F}^3$.

(d) 向量组 $(1,2),(3,5),(4,13)$ 张成 $\mathbf{F}^2$ 但不是 $\mathbf{F}^2$ 的基，因为它不是线性无关的.

(e) 向量组 $(1,1,0),(0,0,1)$ 是 $\{(x,x,y)\in\mathbf{F}^3 : x,y\in\mathbf{F}\}$ 的基.

(f) 向量组 $(1,-1,0),(1,0,-1)$ 是 $\{(x,y,z)\in\mathbf{F}^3 : x+y+z=0\}$ 的基.

(g) 向量组 $1,z,\ldots,z^m$ 是 $\mathcal{P}_m(\mathbf{F})$ 的基，它被称为 $\mathcal{P}_m(\mathbf{F})$ 的**标准基**.

除了标准基，$\mathbf{F}^n$ 也有许多别的基. 例如

$$(7,5),(-4,9) \quad 与 \quad (1,2),(3,5)$$

都是 $\mathbf{F}^2$ 的基.

接下来的结果有助于说明为什么基是很有用的.

**2.28 基的判定准则**

$V$ 中向量组 $v_1,\ldots,v_n$ 是 $V$ 的基，当且仅当每个 $v\in V$ 都可以被唯一地写成这样的形式：

$$v=a_1v_1+\cdots+a_nv_n\,, \tag{2.29}$$

其中 $a_1,\ldots,a_n\in\mathbf{F}$.

> 这个证明本质上是再现了我们定义线性无关性时的思想.

**证明** 首先假设 $v_1,\ldots,v_n$ 是 $V$ 的基. 令 $v\in V$. 因为 $v_1,\ldots,v_n$ 张成 $V$，所以存在 $a_1,\ldots,a_n\in\mathbf{F}$ 使得式 (2.29) 成立. 为了说明式 (2.29) 的表示法是唯一的，假设还有一组标量 $c_1,\ldots,c_n$ 使得下式成立：

$$v=c_1v_1+\cdots+c_nv_n.$$

用式 (2.29) 减去上式，我们得到

$$0=(a_1-c_1)v_1+\cdots+(a_n-c_n)v_n.$$

这表明，各 $a_k-c_k$ 都等于 0（因为 $v_1,\ldots,v_n$ 线性无关）. 所以 $a_1=c_1,\ldots,a_n=c_n$. 我们得到了我们欲证的唯一性，完成了一个方向的证明.

对于另一方向，假设每个 $v \in V$ 都能被唯一地由式 (2.29) 给出的形式表示，这表明组 $v_1, \ldots, v_n$ 张成 $V$. 为了说明 $v_1, \ldots, v_n$ 线性无关，假设 $a_1, \ldots, a_n \in \mathbf{F}$ 使得下式成立

$$0 = a_1 v_1 + \cdots + a_n v_n.$$

式 (2.29) 所示表达法（取 $v = 0$）的唯一性表明，$a_1 = \cdots = a_n = 0$. 于是 $v_1, \ldots, v_n$ 是线性无关的，因而是 $V$ 的基. ■

向量空间中的张成组可能不是基，因为它可能不是线性无关的. 接下来这条结果表明，给定任一张成组，其中某些（也可能没有）向量可被抛弃，而使得剩下的向量所构成的向量组线性无关，且仍能张成同一向量空间.

举个向量空间 $\mathbf{F}^2$ 中的例子，如果将下述证明中各步骤应用于组 $(1,2),(3,6),(4,7),(5,9)$ 上，那么第二个和第四个向量会被移除，剩下的 $(1,2),(4,7)$ 是 $\mathbf{F}^2$ 的基.

**2.30 每个张成组都包含基**

向量空间中的每个张成组都能被削减成该向量空间的基.

**证明** 假设 $v_1, \ldots, v_n$ 张成 $V$. 我们想从 $v_1, \ldots, v_n$ 中去除某些向量以使剩下的向量形成 $V$ 的基. 为此，我们采用下述含有多步的过程.

首先，令 $B$ 等于向量组 $v_1, \ldots, v_n$.

**步骤 1**

如果 $v_1 = 0$，那么从 $B$ 中删去 $v_1$. 如果 $v_1 \neq 0$，那么保持 $B$ 不变.

**步骤 $k$**

如果 $v_k$ 在 $\operatorname{span}(v_1, \ldots, v_{k-1})$ 中，那么从 $B$ 中删去 $v_k$. 如果 $v_k$ 不在 $\operatorname{span}(v_1, \ldots, v_{k-1})$ 中，那么保持 $B$ 不变.

在步骤 $n$ 后停止该过程，得到了经过处理的向量组 $B$. 这个组张成 $V$，因为原本的组就张成 $V$，而我们只是将那些处于排在自身之前的向量的张成空间中的向量给去掉. 这个过程还保证了 $B$ 中没有一个向量处于排在自身之前的向量的张成空间中，于是根据线性相关性引理（2.19），$B$ 是线性无关的. 所以 $B$ 是 $V$ 的基. ■

我们现在得到上述结果的一个重要推论.

**2.31 有限维向量空间的基**

每个有限维向量空间都有基.

**证明** 由定义，有限维向量空间含有张成组. 上一个结果告诉我们，每个张成组都能被削减为一个基. ■

接下来这个结果在某种意义上说是 2.30 的对偶. 2.30 是说每个张成组都能被削减成基. 现在，我们证明，给定某个空间中任一线性无关组，我们可以添加一些向量（也可能一个向量也不用添加），使得经扩充后的组仍是线性无关的但能张成这个空间.

**2.32 每个线性无关组都可被扩充成基**

有限维向量空间中每个线性无关向量组都可被扩充成该向量空间的基.

**证明** 假设 $u_1,\ldots,u_m$ 在有限维向量空间 $V$ 中线性无关. 令 $w_1,\ldots,w_n$ 是 $V$ 中张成 $V$ 的向量组. 于是向量组

$$u_1,\ldots,u_m,w_1,\ldots,w_n$$

张成 $V$.将 2.30 的证明中各步骤应用于该向量组,将其削减成 $V$ 的一个基,可得一个由 $u_1,\ldots,u_m$ 和某些 $w$ 构成的基(此过程中没有哪个 $u$ 被删掉,因为 $u_1,\ldots,u_m$ 是线性无关的). ■

举个 $\mathbf{F}^3$ 中的例子:取线性无关组 $(2,3,4),(9,6,8)$,如果我们将 $w_1,w_2,w_3$ 取为 $\mathbf{F}^3$ 的标准基,那么运用上述证明中各步骤,可得向量组

$$(2,3,4),(9,6,8),(0,1,0),$$

它是 $\mathbf{F}^3$ 的基.

运用相同的思想但是更加高级的工具,无需假设 $V$ 是有限维的,就能证明接下来的结果.

应用上述结果,我们现在证明,对于有限维向量空间的每个子空间,都能找到另一个子空间,使得整个空间等于这两个子空间的直和.

**2.33 $V$ 的每个子空间都是等于 $V$ 的直和的组成部分**

假设 $V$ 是有限维的,$U$ 是 $V$ 的子空间. 那么存在 $V$ 的子空间 $W$,使得 $V=U\oplus W$.

**证明** 因为 $V$ 是有限维的,所以 $U$ 也如此(见 2.25). 于是,$U$ 中存在一个基 $u_1,\ldots,u_m$(由 2.31). 显然 $u_1,\ldots,u_m$ 是 $V$ 中的线性无关向量组. 所以,这个组可以被扩充为 $V$ 的一个基 $u_1,\ldots,u_m,w_1,\ldots,w_n$(由 2.32). 令 $W=\operatorname{span}(w_1,\ldots,w_n)$.

为证明 $V=U\oplus W$,根据 1.46,我们只需证明

$$V=U+W \quad 且 \quad U\cap W=\{0\}.$$

为了证明上述第一个式子,假设 $v\in V$. 那么,由于组 $u_1,\ldots,u_m,w_1,\ldots,w_n$ 张成 $V$,所以存在 $a_1,\ldots,a_m,b_1,\ldots,b_n\in\mathbf{F}$ 使得

$$v=\underbrace{a_1u_1+\cdots+a_mu_m}_{u}+\underbrace{b_1w_1+\cdots+b_nw_n}_{w}.$$

则有 $v=u+w$,其中 $u\in U$ 且 $w\in W$,定义如上. 于是 $v\in U+W$,完成了 $V=U+W$ 的证明.

为了说明 $U\cap W=\{0\}$,假设 $v\in U\cap W$. 那么存在标量 $a_1,\ldots,a_m,b_1,\ldots,b_n\in\mathbf{F}$ 使得

$$v=a_1u_1+\cdots+a_mu_m=b_1w_1+\cdots+b_nw_n.$$

于是

$$a_1u_1+\cdots+a_mu_m-b_1w_1-\cdots-b_nw_n=0.$$

由于 $u_1,\ldots,u_m,w_1,\ldots,w_n$ 是线性无关的,这表明

$$a_1=\cdots=a_m=b_1=\cdots=b_n=0.$$

于是 $v=0$,完成了 $U\cap W=\{0\}$ 的证明. ■

## 习题 2B

**1** 求出所有恰好有一个基的向量空间.

**2** 验证例 2.27 中的所有结论.

**3** (a) 设 $U$ 为 $\mathbf{R}^5$ 的子空间，定义为

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ 且 } x_3 = 7x_4\}.$$

求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathbf{R}^5$ 的一个基.

(c) 求 $\mathbf{R}^5$ 的一个子空间 $W$ 使得 $\mathbf{R}^5 = U \oplus W$.

**4** (a) 设 $U$ 为 $\mathbf{C}^5$ 的子空间，定义为

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbf{C}^5 : 6z_1 = z_2 \text{ 且 } z_3 + 2z_4 + 3z_5 = 0\}.$$

求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathbf{C}^5$ 的一个基.

(c) 求 $\mathbf{C}^5$ 的一个子空间 $W$ 使得 $\mathbf{C}^5 = U \oplus W$.

**5** 设 $V$ 是有限维的，$U, W$ 是 $V$ 的子空间，且 $V = U + W$. 证明 $V$ 有一个由 $U \cup W$ 中向量组成的基.

**6** 证明或给出一反例：如果 $p_0, p_1, p_2, p_3$ 是 $\mathcal{P}_3(\mathbf{F})$ 中的组，该组中每个多项式次数都不为 2，那么 $p_0, p_1, p_2, p_3$ 不是 $\mathcal{P}_3(\mathbf{F})$ 的基.

**7** 设 $v_1, v_2, v_3, v_4$ 是 $V$ 的基，证明

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

也是 $V$ 的基.

**8** 证明或给出一反例：如果 $v_1, v_2, v_3, v_4$ 是 $V$ 的基，且 $U$ 是 $V$ 的一个满足 $v_1, v_2 \in U$ 而 $v_3 \notin U$ 且 $v_4 \notin U$ 的子空间，那么 $v_1, v_2$ 是 $U$ 的基.

**9** 设 $v_1, \ldots, v_m$ 是 $V$ 中的一组向量. 对于 $k \in \{1, \ldots, m\}$，令

$$w_k = v_1 + \cdots + v_k.$$

证明：$v_1, \ldots, v_m$ 是 $V$ 的基，当且仅当 $w_1, \ldots, w_m$ 是 $V$ 的基.

**10** 设 $U$ 和 $W$ 是 $V$ 的子空间，$V = U \oplus W$. 又设 $u_1, \ldots, u_m$ 是 $U$ 的基，$w_1, \ldots, w_n$ 是 $W$ 的基. 证明：

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

是 $V$ 的基.

**11** 设 $V$ 是实向量空间，证明：如果 $v_1, \ldots, v_n$ 是 $V$（视为实向量空间）的基，那么 $v_1, \ldots, v_n$ 也是其复化 $V_{\mathbf{C}}$（视为复向量空间）的基.

注 复化 $V_{\mathbf{C}}$ 的定义见 1B 节习题 8.

## 2C 维数

尽管我们一直在讨论有限维数的向量空间，我们仍未定义这种对象的“维数”究竟是什么. 应该怎么定义维数呢？一个合理的定义，应该能使 $\mathbf{F}^n$ 的维数等于 $n$. 注意到，$\mathbf{F}^n$ 的标准基

$$(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,\ldots,0,1)$$

的长度是 $n$. 于是，我们想把维数定义成基的长度. 然而，一个有限维向量空间往往有很多不同的基，只有当给定向量空间中所有的基都有相同长度时，我们期望的定义才是合理的. 幸好事实就是如此——现在我们就给出证明.

**2.34 基的长度不依赖于基的选取**

有限维向量空间的任意两个基都有相同的长度.

**证明** 假设 $V$ 是有限维的. 令 $B_1$ 和 $B_2$ 是 $V$ 的两个基. 那么 $B_1$ 在 $V$ 中是线性无关的，$B_2$ 张成 $V$. 所以 $B_1$ 的长度不会超过 $B_2$ 的长度（由 2.22）. 互换 $B_1$ 和 $B_2$ 的角色，我们同样能看到，$B_2$ 的长度不会超过 $B_1$ 的长度. 于是，$B_1$ 的长度等于 $B_2$ 的长度，命题得证. ■

既然我们知道一个有限维向量空间的任意两个基有相同的长度，我们可以正式地定义这类空间的维数了.

**2.35 定义：维数（dimension）、$\dim V$**

- 有限维向量空间的**维数**是这个向量空间中任意一个基的长度.
- 有限维向量空间 $V$ 的维数记作 $\dim V$.

**2.36 例：维数**

- $\dim \mathbf{F}^n = n$，因为 $\mathbf{F}^n$ 的标准基的长度是 $n$.
- $\dim \mathcal{P}_m(\mathbf{F}) = m+1$，因为 $\mathcal{P}_m(\mathbf{F})$ 的标准基 $1, z, \ldots, z^m$ 的长度是 $m+1$.
- 如果 $U = \{(x,x,y) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}$，那么 $\dim U = 2$，这是因为 $(1,1,0),(0,0,1)$ 是 $U$ 的基.
- 如果 $U = \{(x,y,z) \in \mathbf{F}^3 : x+y+z=0\}$，那么 $\dim U = 2$，这是因为 $(1,-1,0),(1,0,-1)$ 是 $U$ 的基.

有限维向量空间的每个子空间都是有限维的（由 2.25），则其也有维数. 接下来的结果给出了我们所期望的有关子空间维数的不等式.

**2.37 子空间的维数**

如果 $V$ 是有限维的且 $U$ 是 $V$ 的子空间，那么 $\dim U \le \dim V$.

**证明** 假设 $V$ 是有限维的且 $U$ 是 $V$ 的子空间. 将 $U$ 中的基看成 $V$ 中的线性无关组，并将 $V$ 的基看成 $V$ 中的张成组. 这样，利用 2.22 即可得出结论 $\dim U \le \dim V$. ■

为了确定 $V$ 中一个向量组是 $V$ 的基，根据定义，我们必须证明这个组满足两个性质：它必须是线性无关的，还必须张成 $V$. 下面两个结论表明，如果所讨论的组具有恰当的长度，那么我们只需要证明它满足所要求的两个性质之一. 首先我们证明，每个具有恰当长度的线性无关组都是基.

实向量空间 $\mathbf{R}^2$ 的维数是 2；复向量空间 $\mathbf{C}$ 的维数是 1. 作为集合，$\mathbf{R}^2$ 可以被认为与 $\mathbf{C}$ 等同（并且，在两个空间上，加法是相同的，如果标量取自实数域，那么标量乘法也相同）. 于是，当我们讨论向量空间的维数时，不可忽视 $\mathbf{F}$ 的选取带来的影响.

**2.38 长度恰当的线性无关组是基**

假设 $V$ 是有限维的. 那么 $V$ 中每个长度为 $\dim V$ 的线性无关向量组都是 $V$ 的基.

**证明** 假设 $\dim V = n$ 且 $v_1, \dots, v_n$ 在 $V$ 中是线性无关的. 组 $v_1, \dots, v_n$ 可以被扩充为 $V$ 的基（由 2.32）. 然而，$V$ 的每个基长度都是 $n$，所以这里的扩充是个平凡的情况，也就是说没有元素被加进 $v_1, \dots, v_n$ 中. 于是 $v_1, \dots, v_n$ 是 $V$ 的基，命题得证. ■

接下来这条结果是上述结论的一个有用推论.

**2.39 某空间中与之维数相同的子空间等于这整个空间**

假设 $V$ 是有限维的，$U$ 是 $V$ 的子空间且满足 $\dim U = \dim V$. 那么 $U = V$.

**证明** 令 $u_1, \dots, u_n$ 是 $U$ 的基. 那么 $n = \dim U$，根据前提条件又知 $n = \dim V$. 于是 $u_1, \dots, u_n$ 是 $V$ 中的线性无关向量组（因为它是 $U$ 的基）且长度为 $\dim V$. 由 2.38，我们知道 $u_1, \dots, u_n$ 是 $V$ 的基. 特别地，$V$ 中每个向量都是 $u_1, \dots, u_n$ 的线性组合. 于是 $U = V$. ■

---

**2.40 例：$\mathbf{F}^2$ 的一个基**

考虑 $\mathbf{F}^2$ 中向量组 $(5,7),(4,3)$. 这个长度为 2 的组在 $\mathbf{F}^2$ 中是线性无关的（因为其中任一向量都不是另一向量的标量倍）. 注意到 $\mathbf{F}^2$ 的维数是 2. 于是 2.38 表明长度为 2 的线性无关向量组 $(5,7),(4,3)$ 是 $\mathbf{F}^2$ 的基（我们无需再费心思去检验其是否张成 $\mathbf{F}^2$）.

---

**2.41 例：$\mathcal{P}_3(\mathbf{R})$ 的一个子空间的一个基**

令 $U$ 是 $\mathcal{P}_3(\mathbf{R})$ 的子空间，由下式定义：

$$U = \{p \in \mathcal{P}_3(\mathbf{R}) : p'(5) = 0\}.$$

为了求出 $U$ 的一个基，首先注意到多项式 $1,(x-5)^2$ 和 $(x-5)^3$ 在 $U$ 中.

假设 $a, b, c \in \mathbf{R}$ 且对所有 $x \in \mathbf{R}$ 有

$$a + b(x-5)^2 + c(x-5)^3 = 0.$$

无需将上式左侧完全展开，我们就能发现上式左侧含有 $cx^3$ 项. 因为等式右侧没有 $x^3$ 项，所以这意味着 $c = 0$. 因为 $c = 0$，我们可见等式左侧含 $bx^2$ 项，这意味着 $b = 0$. 同理，由于 $b = c = 0$，我们可以推断出 $a = 0$. 于是，上述等式表明 $a = b = c = 0$. 因此，$1,(x-5)^2,(x-5)^3$ 这个组在 $U$ 中是线性无关的. 于是 $3 \le \dim U$. 所以

$$3 \le \dim U \le \dim \mathcal{P}_3(\mathbf{R}) = 4,$$

转下页

其中用到了 2.37.

多项式 $x$ 不在 $U$ 中，因为它的导数等于常值函数 1. 于是 $U \neq \mathcal{P}_3(\mathbf{R})$. 所以 $\dim U \neq 4$（由 2.39）. 这样，上述的不等式就表明 $\dim U = 3$. 于是，$U$ 中的线性无关组 $1,(x-5)^2,(x-5)^3$ 的长度为 $\dim U$，因此是 $U$ 的基（由 2.38）.

---

现在我们证明，具有恰当长度的张成组就是基.

**2.42 长度恰当的张成组是基**

假设 $V$ 是有限维的. 那么 $V$ 中每个长度为 $\dim V$ 的张成组都是 $V$ 的基.

**证明** 假设 $\dim V = n$ 且 $v_1, \ldots, v_n$ 张成 $V$. 组 $v_1, \ldots, v_n$ 可以被削减成 $V$ 的基（由 2.30）. 然而，$V$ 中每个基的长度都是 $n$，所以此处的削减是个平凡的情况，也就是说没有一个元素被从 $v_1, \ldots, v_n$ 里删除掉. 于是 $v_1, \ldots, v_n$ 就是 $V$ 的基，命题得证. ■

接下来的结果，给出了有限维向量空间的两子空间之和的维数公式. 这个公式类似于我们熟悉的计数公式：两个有限集之并集的元素个数，等于第一个集合的元素个数，加上第二个集合的元素个数，再减去这两个集合之交集的元素个数.

**2.43 子空间之和的维数**

如果 $V_1$ 和 $V_2$ 是一个有限维向量空间的子空间，那么

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

**证明** 令 $v_1, \ldots, v_m$ 是 $V_1 \cap V_2$ 的基. 于是 $\dim(V_1 \cap V_2) = m$. 因为 $v_1, \ldots, v_m$ 是 $V_1 \cap V_2$ 的基，所以在 $V_1$ 中它是线性无关的. 因此这个组可以被扩充为 $V_1$ 的基 $v_1, \ldots, v_m, u_1, \ldots, u_j$（由 2.32）. 于是 $\dim V_1 = m + j$. 同样，将 $v_1, \ldots, v_m$ 扩充为 $V_2$ 的基 $v_1, \ldots, v_m, w_1, \ldots, w_k$. 于是 $\dim V_2 = m + k$.

我们将证明，

$$v_1, \ldots, v_m, u_1, \ldots, u_j, w_1, \ldots, w_k \tag{2.44}$$

是 $V_1 + V_2$ 的基，这是因为由此我们就能得到

$$\begin{aligned} \dim(V_1 + V_2) &= m + j + k \\ &= (m + j) + (m + k) - m \\ &= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2), \end{aligned}$$

进而完成证明.

(2.44) 这个组含于 $V_1 \cup V_2$，则也含于 $V_1 + V_2$. 又知该组的张成空间包含 $V_1$ 和 $V_2$，因此该组的张成空间就等于 $V_1 + V_2$. 于是，要证明 (2.44) 是 $V_1 + V_2$ 的基，我们只需证明它是线性无关的.

为了证明 (2.44) 是线性无关的，假设

$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_j u_j + c_1 w_1 + \cdots + c_k w_k = 0,$$

其中各 $a, b, c$ 都是标量. 我们需要证明各 $a, b, c$ 都等于 0. 可将上式重写成

$$c_1 w_1 + \cdots + c_k w_k = -a_1 v_1 - \cdots - a_m v_m - b_1 u_1 - \cdots - b_j u_j, \tag{2.45}$$

这证明了 $c_1w_1+\cdots+c_kw_k \in V_1$. 又因为各 $w$ 都在 $V_2$ 中，所以上式进一步表明 $c_1w_1+\cdots+c_kw_k \in V_1 \cap V_2$. 因为 $v_1,\ldots,v_m$ 是 $V_1 \cap V_2$ 的基，所以存在标量 $d_1,\ldots,d_m$ 使得

$$c_1w_1+\cdots+c_kw_k = d_1v_1+\cdots+d_mv_m,$$

但是 $v_1,\ldots,v_m,w_1,\ldots,w_k$ 是线性无关的，所以由上式推出各 $c$（以及各 $d$）都等于 0. 于是，式 (2.45) 成为下式

$$a_1v_1+\cdots+a_mv_m+b_1u_1+\cdots+b_ju_j = 0.$$

因为组 $v_1,\ldots,v_m,u_1,\ldots,u_j$ 是线性无关的，所以由此式推出各 $a$ 和各 $b$ 都等于 0，证毕. ■

对于有限集合 $S$，令 $\#S$ 代表 $S$ 的元素个数. 下表将有限集与有限维向量空间做了对比，显示了 $\#S$（对于集合）和 $\dim V$（对于向量空间）的类似性，以及子集的并（就集合而言）和子空间的和（就子空间而言）的类似性.

| **集合（sets）** | **向量空间（vector spaces）** |
|---|---|
| $S$ 是有限集 | $V$ 是有限维向量空间 |
| $\#S$ | $\dim V$ |
| 对于 $S$ 的子集 $S_1,S_2$，它们的并集 $S_1\cup S_2$ 是 $S$ 中最小的包含 $S_1$ 和 $S_2$ 的子集. | 对于 $V$ 的子空间 $V_1,V_2$，它们的和 $V_1+V_2$ 是 $V$ 中最小的包含 $V_1$ 和 $V_2$ 的子空间. |
| $\#(S_1\cup S_2)$ $= \#S_1+\#S_2-\#(S_1\cap S_2)$ | $\dim(V_1+V_2)$ $=\dim V_1+\dim V_2-\dim(V_1\cap V_2)$ |
| $\#(S_1\cup S_2)=\#S_1+\#S_2$ $\iff S_1\cap S_2=\varnothing$ | $\dim(V_1+V_2)=\dim V_1+\dim V_2$ $\iff V_1\cap V_2=\{0\}$ |
| $S_1\cup\cdots\cup S_m$ 是不相交集合的并 $\iff \#(S_1\cup\cdots\cup S_m)=\#S_1+\cdots+\#S_m$ | $V_1+\cdots+V_m$ 是直和 $\iff \dim(V_1+\cdots+V_m)=\dim V_1+\cdots+\dim V_m$ |

上表中最后一行关注的是不相交集合的并（对于集合）和直和（对于向量空间）之间的类似性. 上表中最后一格中的结果[8] 将在 3.94 中证明.

通过类比，你应该能得出习题 12 题至 18 题中，有关向量空间的那些结论所对应到的有关集合的结论.

## 习题 2C

**1** 证明：$\mathbf{R}^2$ 的子空间恰有 $\{0\}$，$\mathbf{R}^2$ 中所有过原点的直线，以及 $\mathbf{R}^2$ 本身.

**2** 证明：$\mathbf{R}^3$ 的子空间恰有 $\{0\}$，$\mathbf{R}^3$ 中所有过原点的直线，$\mathbf{R}^3$ 中所有过原点的平面，以及 $\mathbf{R}^3$ 本身.

**3** (a) 令 $U=\{p\in\mathcal{P}_4(\mathbf{F}) : p(6)=0\}$，求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathcal{P}_4(\mathbf{F})$ 的基.

(c) 求 $\mathcal{P}_4(\mathbf{F})$ 的一个子空间 $W$ 使得 $\mathcal{P}_4(\mathbf{F})=U\oplus W$.

[8] 即最后一行右边那列中的结果.

**4** (a) 令 $U=\{p\in\mathcal{P}_4(\mathbf{R}):p''(6)=0\}$，求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathcal{P}_4(\mathbf{R})$ 的基.

(c) 求 $\mathcal{P}_4(\mathbf{R})$ 的一个子空间 $W$ 使得 $\mathcal{P}_4(\mathbf{R})=U\oplus W$.

**5** (a) 令 $U=\{p\in\mathcal{P}_4(\mathbf{F}):p(2)=p(5)\}$，求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathcal{P}_4(\mathbf{F})$ 的基.

(c) 求 $\mathcal{P}_4(\mathbf{F})$ 的一个子空间 $W$ 使得 $\mathcal{P}_4(\mathbf{F})=U\oplus W$.

**6** (a) 令 $U=\{p\in\mathcal{P}_4(\mathbf{F}):p(2)=p(5)=p(6)\}$，求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathcal{P}_4(\mathbf{F})$ 的基.

(c) 求 $\mathcal{P}_4(\mathbf{F})$ 的一个子空间 $W$ 使得 $\mathcal{P}_4(\mathbf{F})=U\oplus W$.

**7** (a) 令 $U=\left\{p\in\mathcal{P}_4(\mathbf{R}):\int_{-1}^{1}p=0\right\}$，求 $U$ 的一个基.

(b) 将 (a) 中的基扩充为 $\mathcal{P}_4(\mathbf{R})$ 的基.

(c) 求 $\mathcal{P}_4(\mathbf{R})$ 的一个子空间 $W$ 使得 $\mathcal{P}_4(\mathbf{R})=U\oplus W$.

**8** 设 $v_1,\dots,v_m$ 在 $V$ 中线性无关，$w\in V$，证明

$$\dim\operatorname{span}(v_1+w,\dots,v_m+w)\geq m-1.$$

**9** 设 $m$ 是一正整数，$p_0,p_1,\dots,p_m\in\mathcal{P}(\mathbf{F})$，其中 $p_k$ 次数为 $k$. 证明 $p_0,p_1,\dots,p_m$ 是 $\mathcal{P}_m(\mathbf{F})$ 的基.

**10** 设 $m$ 是一正整数，对于 $0\leq k\leq m$，令

$$p_k(x)=x^k(1-x)^{m-k}.$$

证明 $p_0,\dots,p_m$ 是 $\mathcal{P}_m(\mathbf{F})$ 的基.

注　这道习题中的基能够引出所谓**伯恩斯坦多项式（Bernstein polynomials）**. 你可以上网搜索，了解伯恩斯坦多项式如何用于逼近 $[0,1]$ 上的连续函数.

**11** 设 $U$ 和 $W$ 都是 $\mathbf{C}^6$ 的四维子空间，证明：在 $U\cap W$ 中存在两个向量，其中任一个都不是另一个的标量倍.

**12** 设 $U$ 和 $W$ 是 $\mathbf{R}^8$ 的子空间，$\dim U=3$，$\dim W=5$，且 $U+W=\mathbf{R}^8$，证明 $\mathbf{R}^8=U\oplus W$.

**13** 设 $U$ 和 $W$ 都是 $\mathbf{R}^9$ 的五维子空间，证明 $U\cap W\neq\{0\}$.

**14** 设 $V$ 是十维的向量空间，$V_1,V_2,V_3$ 是 $V$ 的子空间，$\dim V_1=\dim V_2=\dim V_3=7$，证明 $V_1\cap V_2\cap V_3\neq\{0\}$.

**15** 设 $V$ 是有限维的，$V_1,V_2,V_3$ 是 $V$ 的子空间，$\dim V_1+\dim V_2+\dim V_3>2\dim V$，证明 $V_1\cap V_2\cap V_3\neq\{0\}$.

**16** 设 $V$ 是有限维的，$U$ 是 $V$ 的子空间，$U\neq V$. 令 $n=\dim V$，$m=\dim U$，证明：$V$ 存在这样 $n-m$ 个子空间，其中每个子空间维数都为 $n-1$ 且它们的交集等于 $U$.

**17** 设 $V_1,\dots,V_m$ 是 $V$ 的有限维子空间，证明：$V_1+\cdots+V_m$ 是有限维的，而且

$$\dim(V_1+\cdots+V_m)\leq\dim V_1+\cdots+\dim V_m.$$

注　以上不等式当且仅当 $V_1+\cdots+V_m$ 是直和时取等，这将在 3.94 中得到证明.

**18** 设 $V$ 是有限维的，$\dim V=n\geq 1$. 证明存在 $V$ 的一维子空间 $V_1,\dots,V_n$ 使得

$$V=V_1\oplus\cdots\oplus V_n.$$

**19** 受三个有限集之并集的元素数量公式的启发，你可能会这样猜测：如果 $V_1, V_2, V_3$ 是一有限维向量空间的子空间，那么有

$$\begin{aligned}&\dim(V_1+V_2+V_3)\\&\quad=\dim V_1+\dim V_2+\dim V_3\\&\qquad-\dim(V_1\cap V_2)-\dim(V_1\cap V_3)-\dim(V_2\cap V_3)\\&\qquad+\dim(V_1\cap V_2\cap V_3).\end{aligned}$$

解释一下为什么这样猜测，然后证明以上公式或给出一反例.

**20** 证明：如果 $V_1, V_2$ 和 $V_3$ 是一有限维向量空间的子空间，那么

$$\begin{aligned}&\dim(V_1+V_2+V_3)\\&\quad=\dim V_1+\dim V_2+\dim V_3\\&\qquad-\frac{\dim(V_1\cap V_2)+\dim(V_1\cap V_3)+\dim(V_2\cap V_3)}{3}\\&\qquad-\frac{\dim\big((V_1+V_2)\cap V_3\big)+\dim\big((V_1+V_3)\cap V_2\big)+\dim\big((V_2+V_3)\cap V_1\big)}{3}.\end{aligned}$$

注 以上公式可能显得有点奇怪，因为右边看着不像一个整数.

我立即放弃了以前的研究，将自然史和从它所派生的一切学问看作畸形的、发育不全的怪胎，对那种甚至不配踏入真正知识大门的所谓的科学嗤之以鼻. 正是怀着这样的心情，我开始研究数学及其相关学科，认为它们有着坚实的根基，因而值得我去认真考察.

——玛丽·沃斯通克拉夫特·雪莱（Mary Wollstonecraft Shelley）《弗兰肯斯坦》（*Frankenstein*）

# 第 3 章 线性映射

到目前为止，我们的注意力都集中在向量空间上. 不会有人为向量空间而兴奋. 线性代数中真正有趣的部分，是我们现在要转而研究的主题——线性映射.

我们将频繁地运用线性映射基本定理这一强大工具. 它指出，线性映射的定义空间的维数，等于被映射到 0 的子空间的维数再加上值域的维数. 由此可得一个非同寻常的结论：对于从有限维向量空间到其本身的线性映射，当且仅当它的值域为它所作用的整个空间时，它才是一对一的映射[1].

我们将在本章引入一个重要概念——矩阵，它与线性映射、定义空间的基及目标空间的基这三者相关联[2]. 线性映射和矩阵的对应关系有助于我们深入理解线性代数的关键内容.

本章最后引入向量空间的积、商空间以及对偶空间.

在本章中，除了 $V$，我们还需用到另外几个向量空间，并称它们为 $U$ 和 $W$. 于是，现在开始我们总采用以下假定.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $U, V$ 和 $W$ 代表 $\mathbf{F}$ 上的向量空间.

♣



图为建于 12 世纪的丹克沃德罗德城堡（Dankwarderode Castle），坐落于不伦瑞克（Brunswick 或 Braunschweig），卡尔·弗里德里希·高斯（Carl Friedrich Gauss，1777-1855）在此出生并长大. 1809 年，高斯发表了一种求解线性方程组的方法. 这种方法【现被称为高斯消元法（Gaussian elimination）】在一部中国古籍[3]里已有运用，比高斯早 1600 多年.

[1]用后面将要学习的术语来表达就是：该映射的单射性与满射性必同时成立.

[2]在论及线性映射时，原文 domain space 或 domain 翻译为“定义空间”，指的是 $\mathcal{L}(V, W)$ 中的前者（$V$），即线性映射之自变量所取自的空间；target space 翻译为“目标空间”，指的是 $\mathcal{L}(V, W)$ 中的后者（$W$），即线性映射之输出所属的空间，注意它和值域的区别（值域是目标空间的子空间）.

[3]此书即《九章算术》.

# 3A 线性映射所成的向量空间

## 线性映射的定义和实例

现在我们给出线性代数中的一个关键定义.

**3.1 定义：线性映射（linear map）**

从 $V$ 到 $W$ 的**线性映射**是满足下列性质的函数 $T: V \to W$.

**可加性（additivity）**

对于所有 $u, v \in V$，均有 $T(u+v) = Tu + Tv$.

**齐次性（homogeneity）**

对于所有 $\lambda \in \mathbf{F}$ 和所有 $v \in V$，均有 $T(\lambda v) = \lambda(Tv)$.

♣

注意，对于线性映射，我们常常用记号 $Tv$，也会用一般的函数记号 $T(v)$.

有些数学家使用术语**“线性变换”（linear transformation）**，这和线性映射同义.

**3.2 记号：$\mathcal{L}(V,W)$、$\mathcal{L}(V)$**

- 从 $V$ 到 $W$ 的全体线性映射构成的集合记作 $\mathcal{L}(V,W)$.
- 从 $V$ 到 $V$ 的全体线性映射构成的集合记作 $\mathcal{L}(V)$. 换言之，$\mathcal{L}(V) = \mathcal{L}(V,V)$.

♣

我们看些线性映射的实例. 请务必自行验证下面的例子中定义的每个函数都确实是线性映射.

---

**3.3 例：线性映射**

**零映射（zero）**

除其他用途外，我们还令符号 0 表示一个这样的线性映射：它将某向量空间中每个元素都对应到另一个（也可能是同一个）向量空间的加法恒等元. 具体而言，$0 \in \mathcal{L}(V,W)$ 定义为

$$0v = 0.$$

上式左侧的 0 是从 $V$ 到 $W$ 的函数，而右侧的 0 是 $W$ 中的加法恒等元. 如往常一样，你根据上下文就可区分符号 0 的多种用法.

**恒等算子（identity operator）**

**恒等算子**记作 $I$，是某个向量空间上的一个这样的线性映射：它将该空间中每个元素都对应到该元素本身. 具体而言，$I \in \mathcal{L}(V)$ 定义为

$$Iv = v.$$

**微分映射（differentiation）**

定义 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 为

$$Dp = p'.$$

断言该函数为线性映射，是用另一种方式表达了有关微分的这个基本结论：对于任意可微函数 $f, g$ 和常数 $\lambda$，$(f+g)' = f' + g'$ 和 $(\lambda f)' = \lambda f'$ 总成立.

转下页

**积分映射（integration）**

定义 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathbf{R})$ 为

$$Tp = \int_0^1 p.$$

断言该函数是线性的，是用另一种方式表达了有关积分的这两个基本结论：两个函数之和的积分等于这两个函数的积分之和，以及一个函数乘以常数再作积分等于该函数作积分后乘以同一常数.

**与 $x^2$ 相乘映射（multiplication by $x^2$）**

定义线性映射 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 为：对于所有 $x \in \mathbf{R}$，

$$(Tp)(x) = x^2 p(x).$$

**后向移位映射（backward shift）**

回忆一下，$\mathbf{F}^\infty$ 表示元素属于 $\mathbf{F}$ 的全体序列构成的向量空间. 定义线性映射 $T \in \mathcal{L}(\mathbf{F}^\infty)$ 为：

$$T(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots).$$

**从 $\mathbf{R}^3$ 到 $\mathbf{R}^2$ 的映射**

定义线性映射 $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$ 为

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

**从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的映射**

为推广上个例子，令 $m$ 和 $n$ 为正整数，并令 $A_{j,k} \in \mathbf{F}$（$j = 1, \ldots, m$ 且 $k = 1, \ldots, n$），定义线性映射 $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ 为

$$T(x_1, \ldots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \ldots, A_{m,1}x_1 + \cdots + A_{m,n}x_n).$$

实际上，每个从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的线性映射都是这种形式.

**多项式复合映射（composition）**

取定一个多项式 $q \in \mathcal{P}(\mathbf{R})$. 定义线性映射 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 为

$$(Tp)(x) = p\big(q(x)\big).$$

下个结论的存在性部分意味着我们总能构造一个线性映射，以将一个基中各向量对应到任意我们想要的值上，唯一性部分则表明一个线性映射完全决定于它将一个基映射到的值.

**3.4 线性映射引理（linear map lemma）**

假定 $v_1, \ldots, v_n$ 是 $V$ 的基且 $w_1, \ldots, w_n \in W$. 那么存在唯一的线性映射 $T : V \to W$ 使得对每个 $k = 1, \ldots, n$ 都有

$$Tv_k = w_k.$$

**证明** 首先我们证明具有我们所期望性质的线性映射 $T$ 存在. 定义 $T : V \to W$ 为

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n,$$

其中 $c_1, \ldots, c_n$ 是 $\mathbf{F}$ 中任意元素. 组 $v_1, \ldots, v_n$ 是 $V$ 的基，由此，上述方程的确定义了一个从 $V$

到 $W$ 的函数 $T$（因为 $V$ 中每个元素都能被唯一地写成 $c_1v_1+\cdots+c_nv_n$ 的形式）.

对于每个 $k$，取上式中 $c_k=1$ 且其他各 $c$ 都等于 0 就证明了 $Tv_k=w_k$.

如果 $u,v\in V$，且有 $u=a_1v_1+\cdots+a_nv_n$ 和 $v=c_1v_1+\cdots+c_nv_n$，则

$$\begin{aligned}T(u+v)&=T\big((a_1+c_1)v_1+\cdots+(a_n+c_n)v_n\big)\\&=(a_1+c_1)w_1+\cdots+(a_n+c_n)w_n\\&=(a_1w_1+\cdots+a_nw_n)+(c_1w_1+\cdots+c_nw_n)\\&=Tu+Tv.\end{aligned}$$

类似地，如果 $\lambda\in\mathbf{F}$ 且 $v=c_1v_1+\cdots+c_nv_n$，那么

$$\begin{aligned}T(\lambda v)&=T(\lambda c_1v_1+\cdots+\lambda c_nv_n)\\&=\lambda c_1w_1+\cdots+\lambda c_nw_n\\&=\lambda(c_1w_1+\cdots+c_nw_n)\\&=\lambda Tv.\end{aligned}$$

于是，$T$ 就是从 $V$ 到 $W$ 的线性映射.

为了证明唯一性，现在假设 $T\in\mathcal{L}(V,W)$ 且对于每个 $k=1,\ldots,n$ 都有 $Tv_k=w_k$. 令 $c_1,\ldots,c_n\in\mathbf{F}$. 那么 $T$ 的齐次性表明对于每个 $k=1,\ldots,n$，都有 $T(c_kv_k)=c_kw_k$. 现结合 $T$ 的可加性可得

$$T(c_1v_1+\cdots+c_nv_n)=c_1w_1+\cdots+c_nw_n.$$

于是 $T$ 在 $\operatorname{span}(v_1,\ldots,v_n)$ 上由上式唯一确定. 因为 $v_1,\ldots,v_n$ 是 $V$ 的基，所以这表明 $T$ 在 $V$ 上唯一确定，即 $T$ 的唯一性得证. ■

## $\mathcal{L}(V,W)$ 上的代数运算

我们从定义 $\mathcal{L}(V,W)$ 上的加法和标量乘法开始.

**3.5 定义：$\mathcal{L}(V,W)$ 上的加法和标量乘法【addition and scalar multiplication on $\mathcal{L}(V,W)$】**

假设 $S,T\in\mathcal{L}(V,W)$ 且 $\lambda\in\mathbf{F}$. 和 $S+T$ 与积 $\lambda T$ 都是从 $V$ 到 $W$ 的线性映射，分别定义如下：对于所有 $v\in V$，

$$(S+T)(v)=Sv+Tv,\quad(\lambda T)(v)=\lambda(Tv).$$

♣

你应该验证定义如上的 $S+T$ 和 $\lambda T$ 确实是线性映射. 换言之，如果 $S,T\in\mathcal{L}(V,W)$ 且 $\lambda\in\mathbf{F}$，那么 $S+T\in\mathcal{L}(V,W)$ 且 $\lambda T\in\mathcal{L}(V,W)$.

因为我们特地在 $\mathcal{L}(V,W)$ 上定义了加法和标量乘法，所以下面结果就不足为奇了.

线性映射在数学中普遍存在. 然而，它们并不如有些人想象的那般无处不在. 当这些人错误地写出 $\cos(x+y)$ 等于 $\cos x+\cos y$ 和 $\cos 2x$ 等于 $2\cos x$ 时，似乎连 cos 都能看成从 $\mathbf{R}$ 到 $\mathbf{R}$ 的线性映射.

**3.6 $\mathcal{L}(V,W)$ 是向量空间**

有了上面定义的加法和标量乘法，$\mathcal{L}(V,W)$ 就是向量空间.

♡

上述结论的证明很常规，留给读者完成．注意 $\mathcal{L}(V,W)$ 的加法恒等元是例 3.3 中定义的零线性映射．

通常，将一个向量空间中的两个元素相乘没有意义，但在某些线性映射之间却能构造出有用的乘积，接下来给出定义．

**3.7 定义：线性映射的乘积（product of linear maps）**

如果 $T \in \mathcal{L}(U,V)$ 且 $S \in \mathcal{L}(V,W)$，那么**乘积** $ST \in \mathcal{L}(U,W)$ 就定义为：对于所有 $u \in U$，

$$(ST)(u) = S(Tu).$$

由此可见，$ST$ 就是一般的两函数复合 $S \circ T$，不过当两个函数都是线性函数时，我们通常使用 $ST$ 这个记号而不用 $S \circ T$．用 $ST$ 这样的乘积表示法，有助于使下个结果中的分配性质看起来更自然些．

注意只有当 $T$ 映射到 $S$ 的定义空间中时，$ST$ 才有定义．你应该验证对于任意 $T \in \mathcal{L}(U,V)$ 和 $S \in \mathcal{L}(V,W)$，$ST$ 的确是从 $U$ 到 $W$ 的线性映射．

**3.8 线性映射乘积的代数性质**

**可结合性（associativity）**

对于任意使乘积有意义的线性映射 $T_1, T_2, T_3$（意即 $T_3$ 映射到 $T_2$ 的定义空间中，$T_2$ 映射到 $T_1$ 的定义空间中），有 $(T_1T_2)T_3 = T_1(T_2T_3)$．

**恒等元（identity）**

对于任意 $T \in \mathcal{L}(V,W)$，有 $TI = IT = T$．这里第一个 $I$ 是 $V$ 上的恒等算子，而第二个 $I$ 是 $W$ 上的恒等算子．

**分配性质（distributive properties）**

对于任意 $T, T_1, T_2 \in \mathcal{L}(U,V)$ 和 $S, S_1, S_2 \in \mathcal{L}(V,W)$，有 $(S_1 + S_2)T = S_1T + S_2T$ 且 $S(T_1 + T_2) = ST_1 + ST_2$．

上述结论的证明很常规，留给读者完成．

线性映射的乘法不满足交换律．换言之，$ST = TS$ 并不一定正确，即便式子两侧的乘积都是有意义的．

**3.9 例：不可交换的两个从 $\mathcal{P}(\mathbf{R})$ 到 $\mathcal{P}(\mathbf{R})$ 的线性映射**

假设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是例 3.3 中定义的微分映射，$T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是例 3.3 中定义的“与 $x^2$ 相乘”映射．那么

$$\big((TD)p\big)(x) = x^2p'(x) \quad \text{但} \quad \big((DT)p\big)(x) = x^2p'(x) + 2xp(x).$$

于是 $TD \neq DT$——先作微分再乘以 $x^2$，和先乘以 $x^2$ 再作微分是不一样的．

**3.10 线性映射将 0 映射为 0**

假设 $T$ 是由 $V$ 到 $W$ 的线性映射．那么 $T(0) = 0$．

**证明** 根据可加性，我们有

$$T(0)=T(0+0)=T(0)+T(0).$$

在等式两侧同时加上 $T(0)$ 的加法逆元，即可得出结论 $T(0)=0$. ■

假设 $m,b\in\mathbf{R}$. 当且仅当 $b=0$ 时，由

$$f(x)=mx+b$$

定义的函数 $f:\mathbf{R}\to\mathbf{R}$ 是线性映射（利用 3.10）. 由此可见，高中代数中的线性函数和线性代数所考虑的线性映射是不一样的.

## 习题 3A

**1** 设 $b,c\in\mathbf{R}$，定义 $T:\mathbf{R}^3\to\mathbf{R}^2$ 为：

$$T(x,y,z)=(2x-4y+3z+b,6x+cxyz).$$

证明：$T$ 是线性的，当且仅当 $b=c=0$.

**2** 设 $b,c\in\mathbf{R}$，定义 $T:\mathcal{P}(\mathbf{R})\to\mathbf{R}^2$ 为：

$$Tp=\left(3p(4)+5p'(6)+bp(1)p(2),\int_{-1}^{2}x^3p(x)\,\mathrm{d}x+c\sin p(0)\right).$$

证明：$T$ 是线性的，当且仅当 $b=c=0$.

**3** 设 $T\in\mathcal{L}(\mathbf{F}^n,\mathbf{F}^m)$，证明：存在标量 $A_{j,k}\in\mathbf{F}$（$j=1,\ldots,m$，$k=1,\ldots,n$），使得

$$T(x_1,\ldots,x_n)=(A_{1,1}x_1+\cdots+A_{1,n}x_n,\ldots,A_{m,1}x_1+\cdots+A_{m,n}x_n)$$

对任一 $(x_1,\ldots,x_n)\in\mathbf{F}^n$ 都成立.

**注** 这道习题表明线性映射 $T$ 具有在例 3.3 的倒数第二例中预示的形式.

**4** 设 $T\in\mathcal{L}(V,W)$ 且 $v_1,\ldots,v_m$ 是 $V$ 中一组向量，使得 $Tv_1,\ldots,Tv_m$ 是 $W$ 中一线性无关组. 证明 $v_1,\ldots,v_m$ 线性无关.

**5** 证明 $\mathcal{L}(V,W)$ 是向量空间，即 3.6 的结论.

**6** 证明线性映射的乘法具有可结合性、恒等元以及分配性质，即 3.8 的结论.

**7** 证明：任何从一维向量空间到其自身的线性映射，就是乘以某一标量. 更准确地说，证明：如果 $\dim V=1$ 且 $T\in\mathcal{L}(V)$，那么存在 $\lambda\in\mathbf{F}$ 使得对所有 $v\in V$ 有 $Tv=\lambda v$.

**8** 给出一例：函数 $\varphi:\mathbf{R}^2\to\mathbf{R}$，使得

$$\varphi(av)=a\varphi(v)$$

对所有 $a\in\mathbf{R}$ 和所有 $v\in\mathbf{R}^2$ 成立，但 $\varphi$ 不是线性的.

**注** 本题和下一题表明，仅有齐次性或仅有可加性，都不足以推导出一个函数是线性映射.

**9** 给出一例：函数 $\varphi:\mathbf{C}\to\mathbf{C}$，使得

$$\varphi(w+z)=\varphi(w)+\varphi(z)$$

对所有 $w,z\in\mathbf{C}$ 成立，但 $\varphi$ 不是线性的.（此处将 $\mathbf{C}$ 视为复向量空间.）

**注** 满足上述可加性条件而不是线性的函数 $\varphi:\mathbf{R}\to\mathbf{R}$ 也是存在的. 然而，证明这样一个函数的存在性要用到更高等（高等得多）的工具.

**10** 证明或给出一反例：如果 $q \in \mathcal{P}(\mathbf{R})$，$T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ 定义为 $Tp = q \circ p$，那么 $T$ 是线性映射.

注 这里定义的函数 $T$，不同于 3.3 中最后一例定义的函数 $T$，区别在于函数复合的次序.

**11** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 证明：$T$ 是恒等算子的标量倍，当且仅当 $ST = TS$ 对任意 $S \in \mathcal{L}(V)$ 都成立.

**12** 设 $U$ 是 $V$ 的子空间（$U \neq V$），设 $S \in \mathcal{L}(U, W)$ 且 $S \neq 0$（意即对某个 $u \in U$ 有 $Su \neq 0$）. 定义 $T : V \to W$ 如下：

$$Tv = \begin{cases} Sv & \text{若 } v \in U, \\ 0 & \text{若 } v \in V \text{ 且 } v \notin U. \end{cases}$$

证明 $T$ 不是 $V$ 上的线性映射.

**13** 设 $V$ 是有限维的. 证明：$V$ 的子空间上的任一线性映射都可以扩充为 $V$ 上的线性映射. 换句话说，证明：如果 $U$ 是 $V$ 的子空间，且 $S \in \mathcal{L}(U, W)$，那么存在 $T \in \mathcal{L}(V, W)$ 使得 $Tu = Su$ 对所有 $u \in U$ 成立.

注 3.125 的证明会用到本题的结果.

**14** 设 $V$ 是有限维的（$\dim V > 0$），又设 $W$ 是无限维的. 证明 $\mathcal{L}(V, W)$ 是无限维的.

**15** 设 $v_1, \ldots, v_m$ 是 $V$ 中的线性相关组，又设 $W \neq \{0\}$. 证明：存在 $w_1, \ldots, w_m \in W$ 使得不存在 $T \in \mathcal{L}(V, W)$ 对每个 $k = 1, \ldots, m$ 都满足 $Tv_k = w_k$.

**16** 设 $V$ 是有限维的（$\dim V > 1$）. 证明存在 $S, T \in \mathcal{L}(V)$ 使得 $ST \neq TS$.

**17** 设 $V$ 是有限维的，证明 $\mathcal{L}(V)$ 仅有的双边理想是 $\{0\}$ 和 $\mathcal{L}(V)$.

注 $\mathcal{L}(V)$ 的子空间 $\mathcal{E}$ 被称为 $\mathcal{L}(V)$ 的**双边理想（two-sided ideal）**，如果 $TE \in \mathcal{E}$ 且 $ET \in \mathcal{E}$ 对所有 $E \in \mathcal{E}$ 和所有 $T \in \mathcal{L}(V)$ 都成立.

# 3B 零空间和值域

## 零空间和单射性

在这一节中，我们将学习与每个线性映射都紧密相关的两个子空间. 首先是被映射为 0 的向量所构成的集合.

**3.11 定义：零空间（null space）、null $T$**

对于 $T \in \mathcal{L}(V, W)$，$T$ 的**零空间**记为 null $T$，是 $V$ 的子集，其由被 $T$ 映射到 0 的所有向量构成：

$$\operatorname{null} T = \{v \in V : Tv = 0\}.$$

♣

**3.12 例：零空间**

- 如果 $T$ 是从 $V$ 到 $W$ 的零映射，意即对于每个 $v \in V$ 都有 $Tv = 0$，那么 null $T = V$.
- 假设 $\varphi \in \mathcal{L}(\mathbf{C}^3, \mathbf{C})$，定义为 $\varphi(z_1, z_2, z_3) = z_1 + 2z_2 + 3z_3$. 那么 null $\varphi$ 等于 $\{(z_1, z_2, z_3) \in \mathbf{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$，它是 $\varphi$ 的定义空间的子空间. 我们很快就会看到，每个线性映射的零空间都是其定义空间的子空间.
- 假设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是微分映射，即其定义为 $Dp = p'$. 导函数为 0 的函数仅有常值函数. 于是，$D$ 的零空间等于常值函数构成的集合.
- 假设 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是与 $x^2$ 相乘映射，即其定义为 $Tp(x) = x^2 p(x)$. 唯一使得 $x^2 p(x) = 0$ 对所有 $x \in \mathbf{R}$ 都成立的多项式 $p$ 是多项式 0. 于是 null $T = \{0\}$.
- 假设 $T \in \mathcal{L}(\mathbf{F}^\infty)$ 是后向移位映射，即其定义为 $T(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots)$. 那么 $T(x_1, x_2, x_3, \ldots)$ 等于 0 当且仅当数 $x_2, x_3, \ldots$ 全为 0. 于是 null $T = \{(a, 0, 0, \ldots) : a \in \mathbf{F}\}$.

接下来的结果说明，每个线性映射的零空间都是其定义空间的子空间. 特别地，每个线性映射的零空间都包含 0.

**3.13 零空间是子空间**

假设 $T \in \mathcal{L}(V, W)$. 那么 null $T$ 是 $V$ 的子空间.

♡

**证明** 因为 $T$ 是线性映射，所以 $T(0) = 0$（由 3.10）. 于是 $0 \in \operatorname{null} T$.

假定 $u, v \in \operatorname{null} T$. 那么

$$T(u + v) = Tu + Tv = 0 + 0 = 0.$$

所以 $u + v \in \operatorname{null} T$. 于是 null $T$ 对于加法封闭.

“null” 的意思是“零”. 于是，“null space” 这个术语应该能让你想起它和 0 的联系. 有些数学家使用术语**“核”（kernel）**而不是零空间.

假定 $u \in \operatorname{null} T$ 且 $\lambda \in \mathbf{F}$. 那么

$$T(\lambda u) = \lambda T u = \lambda 0 = 0.$$

所以 $\lambda u \in \operatorname{null} T$. 于是 null $T$ 对于标量乘法封闭.

我们证明了 null $T$ 含有 0，并且对于加法和标量乘法运算都封闭. 因此 null $T$ 是 $V$ 的子空间（由 1.34）. ■

我们很快就会看到，对于线性映射来说，接下来这条定义和零空间紧密相关.

**3.14 定义：单射（injective）**

对于函数 $T: V \to W$，若 $Tu = Tv$ 蕴涵 $u = v$，则称该函数是**单射**[4].

我们可以将上述定义重新表述为：如果 $u \neq v$ 蕴涵 $Tu \neq Tv$，那么 $T$ 是单射. 于是，当且仅当 $T$ 将不同的输入映射到不同的输出时，它才是单射.

> 术语 **“一对一的”（one-to-one）** 和单射的意思一样.

接下来的结果是说，我们可以通过检验 0 是否为唯一被映射为 0 的向量，来检验一个线性映射是否为单射. 在例 3.12 中，我们计算了一些线性映射的零空间，应用下面结论即可看出，这些线性映射中仅有“与 $x^2$ 相乘”映射是单射（在特殊情况 $V = \{0\}$ 下，零映射也是单射）.

**3.15 单射性 $\Longleftrightarrow$ 零空间等于 $\{0\}$**

令 $T \in \mathcal{L}(V, W)$. 那么 $T$ 是单射当且仅当 $\operatorname{null} T = \{0\}$.

**证明** 首先假设 $T$ 是单射. 我们想要证明 $\operatorname{null} T = \{0\}$. 我们已知 $\{0\} \subseteq \operatorname{null} T$（由 3.10）. 为了证明这个包含关系反过来也成立，假设 $v \in \operatorname{null} T$. 那么

$$T(v) = 0 = T(0).$$

因为 $T$ 是单射，所以上述等式表明 $v = 0$. 于是我们能得出 $\operatorname{null} T = \{0\}$，此方向得证.

为了证明另一方向的蕴涵关系，现在假设 $\operatorname{null} T = \{0\}$. 我们想要证明 $T$ 是单射. 为此，假设 $u, v \in V$ 且 $Tu = Tv$. 那么

$$0 = Tu - Tv = T(u - v).$$

于是 $u - v$ 在 $\operatorname{null} T$ 中，而 $\operatorname{null} T$ 等于 $\{0\}$. 因此 $u - v = 0$，则 $u = v$. 所以 $T$ 是单射，证毕. ■

## 值域和满射性

现在我们给线性映射的输出所构成的集合起个名字.

**3.16 定义：值域（range）**

对于 $T \in \mathcal{L}(V, W)$，$T$ 的**值域**是 $W$ 的子集，由所有等于 $Tv$（其中 $v \in V$）的向量构成：

$$\operatorname{range} T = \{Tv : v \in V\}.$$

**3.17 例：值域**

- 如果 $T$ 是从 $V$ 到 $W$ 的零映射，意即对于所有的 $v \in V$ 都有 $Tv = 0$，那么 $\operatorname{range} T = \{0\}$.
- 假设 $T \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ 定义为 $T(x, y) = (2x, 5y, x + y)$. 那么

$$\operatorname{range} T = \{(2x, 5y, x + y) : x, y \in \mathbf{R}\}.$$

转下页

[4]原书使用 injective，是个形容词. 在不影响句意的情况下，将其用名词“单射”替换.

注意到 range $T$ 是 $\mathbf{R}^3$ 的子空间. 我们很快将看到, $\mathcal{L}(V,W)$ 中每个线性映射的值域都是 $W$ 的子空间.

- 假设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是由 $Dp = p'$ 定义的微分映射. 因为对于每个多项式 $q \in \mathcal{P}(\mathbf{R})$ 都存在一个多项式 $p \in \mathcal{P}(\mathbf{R})$ 使得 $p' = q$, 所以 $D$ 的值域是 $\mathcal{P}(\mathbf{R})$.

接下来的结果表明, 每个线性映射的值域都是它所映射到[5]的向量空间的子空间.

**3.18 值域是子空间**

如果 $T \in \mathcal{L}(V,W)$, 那么 range $T$ 是 $W$ 的子空间.

**证明** 假设 $T \in \mathcal{L}(V,W)$. 那么 $T(0) = 0$ (根据 3.10), 意味着 $0 \in \operatorname{range} T$.

如果 $w_1, w_2 \in \operatorname{range} T$, 那么存在 $v_1, v_2 \in V$ 使 $Tv_1 = w_1$ 和 $Tv_2 = w_2$ 成立. 于是

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2.$$

所以 $w_1 + w_2 \in \operatorname{range} T$, 于是 range $T$ 对于加法封闭.

如果 $w \in \operatorname{range} T$ 且 $\lambda \in \mathbf{F}$, 那么存在 $v \in V$ 使 $Tv = w$ 成立. 于是

$$T(\lambda v) = \lambda T v = \lambda w.$$

所以 $\lambda w \in \operatorname{range} T$, 于是 range $T$ 对于标量乘法封闭.

我们证明了 range $T$ 含有 0 以及对于加法和标量乘法运算都封闭. 因此 range $T$ 是 $W$ 的子空间 (由 1.34). ■

**3.19 定义: 满射 (surjective)**

如果函数 $T: V \to W$ 的值域等于 $W$, 则称该函数为**满射**[6].

我们举些实例来阐释上述定义: 留意在 3.17 中计算过值域的那些映射, 其中只有微分映射是满射 (对于特殊情况 $W = \{0\}$, 零映射也是满射).

一个线性映射是否为满射, 取决于我们认为它映射到哪个向量空间.

有些人使用术语 **"映成"** (**onto**), 这和满射意思一样.

**3.20 例: 是否为满射取决于目标空间的选取**

定义为 $Dp = p'$ 的微分映射 $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ 不是满射, 因为多项式 $x^5$ 不在 $D$ 的值域中. 然而, 定义为 $Sp = p'$ 的微分映射 $S \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ 是满射, 因为它的值域等于 $\mathcal{P}_4(\mathbf{R})$, 恰是 $S$ 映射到的向量空间.

## 线性映射基本定理

下个结果十分重要, 所以它有个引人注目的名字.

[5]映射**到**的向量空间 (目标空间) 和映射**成**的向量空间 (满射, 目标空间等于值域) 是不同的.

[6]原书使用 surjective, 是个形容词. 在不影响句意的情况下, 将其用名词 "满射" 替换.

**3.21 线性映射基本定理（fundamental theorem of linear maps）**

假设 $V$ 是有限维的且 $T \in \mathcal{L}(V,W)$. 那么 range $T$ 是有限维的，且

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

**证明** 令 $u_1,\ldots,u_m$ 是 null $T$ 的基，于是 $\dim \operatorname{null} T = m$. 由 2.32，线性无关组 $u_1,\ldots,u_m$ 可被扩充成 $V$ 的基

$$u_1,\ldots,u_m,v_1,\ldots,v_n.$$

于是 $\dim V = m+n$. 为了完成证明，我们需要证明 range $T$ 是有限维的且 $\dim \operatorname{range} T = n$. 为此，我们证明 $Tv_1,\ldots,Tv_n$ 是 range $T$ 的基.

令 $v \in V$. 因为 $u_1,\ldots,u_m,v_1,\ldots,v_n$ 张成 $V$，所以我们可以写出

$$v = a_1u_1+\cdots+a_mu_m+b_1v_1+\cdots+b_nv_n,$$

其中各 $a$ 和 $b$ 都属于 $\mathbf{F}$. 将 $T$ 作用于等式两侧，可得

$$Tv = b_1Tv_1+\cdots+b_nTv_n,$$

其中形如 $Tu_k$ 的项消失了，因为各 $u_k$ 均属于 null $T$. 上述方程表明组 $Tv_1,\ldots,Tv_n$ 张成 range $T$. 特别地，range $T$ 是有限维的.

为了证明 $Tv_1,\ldots,Tv_n$ 是线性无关的，假设 $c_1,\ldots,c_n \in \mathbf{F}$ 且

$$c_1Tv_1+\cdots+c_nTv_n = 0.$$

那么，

$$T(c_1v_1+\cdots+c_nv_n) = 0.$$

因此，

$$c_1v_1+\cdots+c_nv_n \in \operatorname{null} T.$$

因为 $u_1,\ldots,u_m$ 张成 null $T$，所以我们可以写出

$$c_1v_1+\cdots+c_nv_n = d_1u_1+\cdots+d_mu_m,$$

其中各 $d$ 都属于 $\mathbf{F}$. 这个等式表明，所有各 $c$（以及各 $d$）都是 0（因为 $u_1,\ldots,u_m,v_1,\ldots,v_n$ 是线性无关的）. 于是 $Tv_1,\ldots,Tv_n$ 是线性无关的，进而它是 range $T$ 的基. 则原命题得证. ■

现在我们可以证明，将一个有限维向量空间映射到比它"更小"的向量空间的线性映射都不是单射，这里的"更小"是用维数来度量的.

**3.22 映到更低维空间上的线性映射不是单射**

假设 $V$ 和 $W$ 是有限维向量空间且满足 $\dim V > \dim W$. 那么从 $V$ 到 $W$ 的线性映射一定不是单射.

**证明** 令 $T \in \mathcal{L}(V, W)$. 那么

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &\geq \dim V - \dim W \\ &> 0, \end{aligned}$$

上述第一行来自线性映射基本定理（3.21），第二行来自 2.37. 上述不等式表明，$\dim \operatorname{null} T > 0$. 这就意味着 $\operatorname{null} T$ 包含除了 0 以外的其他向量. 于是 $T$ 不是单射（由 3.15）. ■

**3.23 例：从 $\mathbf{F}^4$ 到 $\mathbf{F}^3$ 的线性映射不是单射**

按下式定义一个线性映射 $T: \mathbf{F}^4 \to \mathbf{F}^3$：

$$T(z_1, z_2, z_3, z_4) = \left(\sqrt{7} z_1 + \pi z_2 + z_4, 97 z_1 + 3 z_2 + 2 z_3, z_2 + 6 z_3 + 7 z_4\right).$$

因为 $\dim \mathbf{F}^4 > \dim \mathbf{F}^3$，我们可以利用 3.22 来断言 $T$ 不是单射，而无需任何计算.

接下来的结果表明，从一个有限维向量空间映射到比它“更大”的向量空间的线性映射都不是满射，这里的“更大”是从维数上来度量的.

**3.24 映到更高维空间上的线性映射不是满射**

假设 $V$ 和 $W$ 是有限维向量空间且满足 $\dim V < \dim W$. 那么从 $V$ 到 $W$ 的线性映射一定不是满射.

**证明** 令 $T \in \mathcal{L}(V, W)$. 那么

$$\begin{aligned} \dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T \\ &\leq \dim V \\ &< \dim W, \end{aligned}$$

上述式子中的等号源于线性映射基本定理（3.21）. 该不等式表明，$\dim \operatorname{range} T < \dim W$. 这就意味着 $\operatorname{range} T$ 不会等于 $W$. 于是 $T$ 不是满射. ■

我们马上就会看到，3.22 和 3.24 在线性方程理论中有一些重要推论，其中的想法是将有关线性方程组的问题用线性映射的语言来表达. 首先，我们将齐次线性方程组是否有非零解这个问题用线性映射的语言来改写.

**齐次的（homogeneous）**，在这个语境里，意为下面每个方程右端的常数项都是 0.

固定正整数 $m$ 和 $n$，并令 $A_{j,k} \in \mathbf{F}$（$j = 1, \ldots, m$ 且 $k = 1, \ldots, n$）. 考虑齐次线性方程组

$$\begin{gathered} \sum_{k=1}^{n} A_{1,k} x_k = 0 \\ \vdots \\ \sum_{k=1}^{n} A_{m,k} x_k = 0. \end{gathered}$$

显然，$x_1 = \cdots = x_n = 0$ 是上述方程组的一个解. 此处的问题是，是否存在其他的解.

定义 $T:\mathbf{F}^n \to \mathbf{F}^m$ 为

$$T(x_1,\ldots,x_n)=\left(\sum_{k=1}^{n}A_{1,k}x_k,\ldots,\sum_{k=1}^{n}A_{m,k}x_k\right). \tag{3.25}$$

方程 $T(x_1,\ldots,x_n)=0$（此处的 0 是 $\mathbf{F}^m$ 中的加法恒等元，也就是长度为 $m$ 的全由 0 构成的组）与上面的齐次线性方程组是一样的．于是我们就想知道 $\operatorname{null}T$ 是否严格大于 $\{0\}$，这也就等价于 $T$ 不是单射（由 3.15）．接下来的结果给出了保证 $T$ 不是单射的条件．

**3.26　齐次线性方程组**

未知数个数多于方程个数的齐次线性方程组具有非零解．

**证明**　沿用前面讨论中所用的记号和结论．于是 $T$ 是从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的线性映射，且我们有共含 $m$ 个方程、$n$ 个未知数 $x_1,\ldots,x_n$ 的齐次线性方程组．由 3.22 即可知若 $n>m$ 则 $T$ 就不是单射．■

上述结果的一个实例：具有四个方程和五个未知数的齐次线性方程组具有非零解．

现在我们考虑这个问题：一个线性方程组是否在常数项取某些值时会是无解的．为将这个问题用线性映射的语言来描述，固定正整数 $m$ 和 $n$，并令 $A_{j,k}\in\mathbf{F}$（$j=1,\ldots,m$ 且 $k=1,\ldots,n$）．对于 $c_1,\ldots,c_m\in\mathbf{F}$，考虑线性方程组

$$\begin{aligned}\sum_{k=1}^{n}A_{1,k}x_k&=c_1\\&\vdots\\\sum_{k=1}^{n}A_{m,k}x_k&=c_m.\end{aligned} \tag{3.27}$$

用这样的记号，我们可将问题表述成：是否存在常数项 $c_1,\ldots,c_m\in\mathbf{F}$ 的某些取值，使得上述方程组无解．

如式 (3.25) 一样定义 $T:\mathbf{F}^n\to\mathbf{F}^m$．方程 $T(x_1,\ldots,x_n)=(c_1,\ldots,c_m)$ 与上面式 (3.27) 所示的方程组是一样的．于是我们想知道是否有 $\operatorname{range}T\neq\mathbf{F}^m$．所以我们可以将问题（$c_1,\ldots,c_m\in\mathbf{F}$ 的某些取值使得方程组无解）改写成这样：什么条件能使 $T$ 一定不是满射？接下来的结论就给出了一个这样的条件．

结论 3.26 和 3.28 都将未知数的个数和方程的个数作比较．它们也可利用高斯消元法证明，但此处所用的抽象方法似乎提供了更简洁的证明．

**3.28　方程个数多于未知数个数的线性方程组**

方程个数多于未知数个数的线性方程组当常数项取某些值时无解．

**证明**　沿用前面讨论中所用的记号和结论．于是 $T$ 是从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的线性映射，且我们有共含 $m$ 个方程、$n$ 个未知数 $x_1,\ldots,x_n$ 的线性方程组【见式 (3.27)】．若 $n<m$，由 3.24 即可知 $T$ 不是满射．按照上面的讨论，这表明，如果一个线性方程组中，方程个数多于未知数个数，那么当常数项取某些值时该方程组无解．■

上述结果的一个实例：具有五个方程和四个未知数的线性方程组在常数项取某些值时会是无解的．

## 习题 3B

**1** 给出一例：满足 $\dim \operatorname{null} T = 3$ 且 $\dim \operatorname{range} T = 2$ 的线性映射 $T$.

**2** 设 $S, T \in \mathcal{L}(V)$ 使得 $\operatorname{range} S \subseteq \operatorname{null} T$，证明 $(ST)^2 = 0$.

**3** 设 $v_1, \ldots, v_m$ 是 $V$ 中一组向量，定义 $T \in \mathcal{L}(\mathbf{F}^m, V)$ 为：

$$T(z_1, \ldots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

(a) $T$ 的什么性质对应于 $v_1, \ldots, v_m$ 张成 $V$？

(b) $T$ 的什么性质对应于 $v_1, \ldots, v_m$ 是线性无关组？

**4** 证明：$\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$ 不是 $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$ 的子空间.

**5** 给出一例：使得 $\operatorname{range} T = \operatorname{null} T$ 的 $T \in \mathcal{L}(\mathbf{R}^4)$.

**6** 证明不存在 $T \in \mathcal{L}(\mathbf{R}^5)$ 使得 $\operatorname{range} T = \operatorname{null} T$.

**7** 设 $V$ 和 $W$ 是有限维的（$2 \leq \dim V \leq \dim W$），证明：$\{T \in \mathcal{L}(V, W) : T \text{ 不是单射}\}$ 不是 $\mathcal{L}(V, W)$ 的子空间.

**8** 设 $V$ 和 $W$ 是有限维的（$\dim V \geq \dim W \geq 2$），证明：$\{T \in \mathcal{L}(V, W) : T \text{ 不是满射}\}$ 不是 $\mathcal{L}(V, W)$ 的子空间.

**9** 设 $T \in \mathcal{L}(V, W)$ 是单射，$v_1, \ldots, v_n$ 在 $V$ 中线性无关，证明 $Tv_1, \ldots, Tv_n$ 在 $W$ 中线性无关.

**10** 设 $v_1, \ldots, v_n$ 张成 $V$，$T \in \mathcal{L}(V, W)$，证明 $Tv_1, \ldots, Tv_n$ 张成 $\operatorname{range} T$.

**11** 设 $V$ 是有限维的，$T \in \mathcal{L}(V, W)$，证明存在 $V$ 的一子空间 $U$ 使得

$$U \cap \operatorname{null} T = \{0\} \quad \text{且} \quad \operatorname{range} T = \{Tu : u \in U\}.$$

**12** 设 $T$ 是从 $\mathbf{F}^4$ 到 $\mathbf{F}^2$ 的线性映射，使得

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ 且 } x_3 = 7x_4\}.$$

证明 $T$ 是满射.

**13** 设 $U$ 是 $\mathbf{R}^8$ 的三维子空间，$T$ 是从 $\mathbf{R}^8$ 到 $\mathbf{R}^5$ 的线性映射，$\operatorname{null} T = U$. 证明 $T$ 是满射.

**14** 证明：不存在从 $\mathbf{F}^5$ 到 $\mathbf{F}^2$ 的线性映射，满足其零空间等于

$$\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ 且 } x_3 = x_4 = x_5\}.$$

**15** 设存在 $V$ 上的线性映射，其零空间和值域都是有限维的. 证明 $V$ 是有限维的.

**16** 设 $V$ 和 $W$ 都是有限维的. 证明：存在从 $V$ 到 $W$ 的单的线性映射，当且仅当 $\dim V \leq \dim W$.

**17** 设 $V$ 和 $W$ 都是有限维的. 证明：存在从 $V$ 到 $W$ 的满的线性映射，当且仅当 $\dim V \geq \dim W$.

**18** 设 $V$ 和 $W$ 都是有限维的，$U$ 是 $V$ 的子空间. 证明：存在 $T \in \mathcal{L}(V, W)$ 使得 $\operatorname{null} T = U$，当且仅当 $\dim U \geq \dim V - \dim W$.

**19** 设 $W$ 是有限维的，$T \in \mathcal{L}(V, W)$. 证明：$T$ 是单射，当且仅当存在 $S \in \mathcal{L}(W, V)$ 使得 $ST$ 是 $V$ 上的恒等算子.

**20** 设 $W$ 是有限维的，$T \in \mathcal{L}(V, W)$. 证明：$T$ 是满射，当且仅当存在 $S \in \mathcal{L}(W, V)$ 使得 $TS$ 是 $W$ 上的恒等算子.

**21** 设 $V$ 是有限维的，$T \in \mathcal{L}(V, W)$，$U$ 是 $W$ 的子空间. 证明：$\{v \in V : Tv \in U\}$ 是 $V$ 的子空间，且

$$\dim \{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

**22** 设 $U$ 和 $V$ 都是有限维向量空间，$S \in \mathcal{L}(V,W)$ 和 $T \in \mathcal{L}(U,V)$，证明：

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

**23** 设 $U$ 和 $V$ 都是有限维向量空间，$S \in \mathcal{L}(V,W)$ 和 $T \in \mathcal{L}(U,V)$，证明：

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

**24** (a) 设 $\dim V = 5$，且 $S,T \in \mathcal{L}(V)$ 使得 $ST = 0$．证明 $\dim \operatorname{range} TS \leq 2$.

(b) 给出一例：$S,T \in \mathcal{L}(\mathbf{F}^5)$，$ST = 0$ 且 $\dim \operatorname{range} TS = 2$.

**25** 设 $W$ 是有限维的，$S,T \in \mathcal{L}(V,W)$．证明：$\operatorname{null} S \subseteq \operatorname{null} T$，当且仅当存在 $E \in \mathcal{L}(W)$ 使得 $T = ES$.

**26** 设 $V$ 是有限维的，$S,T \in \mathcal{L}(V,W)$．证明：$\operatorname{range} S \subseteq \operatorname{range} T$，当且仅当存在 $E \in \mathcal{L}(V)$ 使得 $S = TE$.

**27** 设 $P \in \mathcal{L}(V)$ 且 $P^2 = P$．证明 $V = \operatorname{null} P \oplus \operatorname{range} P$.

**28** 设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 使得对任一非常数多项式 $p \in \mathcal{P}(\mathbf{R})$ 有 $\deg Dp = (\deg p) - 1$．证明 $D$ 是满射.

注　上面的记号 $D$ 是用来让你想起微分映射，它将多项式 $p$ 对应到 $p'$.

**29** 设 $p \in \mathcal{P}(\mathbf{R})$．证明：存在多项式 $q \in \mathcal{P}(\mathbf{R})$，使得 $5q'' + 3q' = p$.

注　做这道习题可以不用到线性代数，但是用线性代数来做更有意思.

**30** 设 $\varphi \in \mathcal{L}(V,\mathbf{F})$（$\varphi \neq 0$），$u \in V$ 不在 $\operatorname{null}\varphi$ 里．证明：

$$V = \operatorname{null}\varphi \oplus \{au : a \in \mathbf{F}\}.$$

**31** 设 $V$ 是有限维的，$X$ 是 $V$ 的子空间，$Y$ 是 $W$ 的有限维子空间．证明：存在 $T \in \mathcal{L}(V,W)$ 使得 $\operatorname{null} T = X$ 且 $\operatorname{range} T = Y$，当且仅当 $\dim X + \dim Y = \dim V$.

**32** 设 $V$ 是有限维的（$\dim V > 1$），证明：如果 $\varphi : \mathcal{L}(V) \to \mathbf{F}$ 是线性映射，使得 $\varphi(ST) = \varphi(S)\varphi(T)$ 对所有 $S,T \in \mathcal{L}(V)$ 成立，那么 $\varphi = 0$.

提示　3A 节的习题 17 中给出了关于 $\mathcal{L}(V)$ 的双边理想的描述，或许有用.

**33** 设 $V$ 和 $W$ 是实向量空间，$T \in \mathcal{L}(V,W)$．定义 $T_{\mathbf{C}} : V_{\mathbf{C}} \to W_{\mathbf{C}}$：

$$T_{\mathbf{C}}(u + \mathrm{i}v) = Tu + \mathrm{i}Tv$$

对所有 $u,v \in V$ 成立.

(a) 证明：$T_{\mathbf{C}}$ 是从 $V_{\mathbf{C}}$ 到 $W_{\mathbf{C}}$ 的（复）线性映射.

(b) 证明：$T_{\mathbf{C}}$ 是单射，当且仅当 $T$ 是单射.

(c) 证明：$\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$，当且仅当 $\operatorname{range} T = W$.

注　复化 $V_{\mathbf{C}}$ 的定义见 1B 节的习题 8．线性映射 $T_{\mathbf{C}}$ 称为线性映射 $T$ 的**复化**.

# 3C 矩阵

## 用矩阵表示线性映射

我们知道，如果 $v_1, \ldots, v_n$ 是 $V$ 的基且 $T: V \to W$ 是线性的，那么 $Tv_1, \ldots, Tv_n$ 的取值决定了 $T$ 将 $V$ 中任意向量所映射到的值——见线性映射引理（3.4）．我们马上会看到，利用 $W$ 的基，矩阵可以高效地记录各 $Tv_k$ 的值．

**3.29 定义：矩阵（matrix）、$A_{j,k}$**

假设 $m$ 和 $n$ 是非负整数．$m \times n$ **矩阵** $A$ 是由 $\mathbf{F}$ 中元素构成的 $m$ 行 $n$ 列的矩形阵列：

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$

记号 $A_{j,k}$ 表示 $A$ 的第 $j$ 行第 $k$ 列中的元素．♣

**3.30 例：$A_{j,k}$ 等于 $A$ 的第 $j$ 行第 $k$ 列中的元素**

假设 $A = \begin{pmatrix} 8 & 4 & 5-3\mathrm{i} \\ 1 & 9 & 7 \end{pmatrix}$．于是 $A_{2,3}$ 表示 $A$ 的第二行第三列中的元素，意即 $A_{2,3} = 7$．

在讨论矩阵时，第一个下标代表第几行，第二个坐标代表第几列．

现在我们给出本节中的关键定义．

**3.31 定义：线性映射的矩阵（matrix of a linear map）、$\mathcal{M}(T)$**

假设 $T \in \mathcal{L}(V, W)$，$v_1, \ldots, v_n$ 是 $V$ 的基，$w_1, \ldots, w_m$ 是 $W$ 的基，$T$ 关于这些基的**矩阵**是 $m \times n$ 矩阵 $\mathcal{M}(T)$，其中各元素 $A_{j,k}$ 由下式定义：

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m.$$

如果从上下文无法明确得知基 $v_1, \ldots, v_n$ 和 $w_1, \ldots, w_m$ 取什么，那么就用 $\mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$ 这个记号．♣

线性映射 $T \in \mathcal{L}(V, W)$ 的矩阵 $\mathcal{M}(T)$ 决定于 $V$ 的基 $v_1, \ldots, v_n$，$W$ 的基 $w_1, \ldots, w_m$，以及 $T$．不过，从上下文中往往能明确得知这些基，因此常常把它们从记号里省去．

为了记住 $\mathcal{M}(T)$ 是如何由 $T$ 构造出来的，你可以在矩阵的上方横着标上定义空间的基向量 $v_1, \ldots, v_n$，并在矩阵的左侧竖着列出 $T$ 映射到的向量空间的基向量 $w_1, \ldots, w_m$，就像下面这样：

$$\mathcal{M}(T) = \begin{array}{c} \\ w_1 \\ \vdots \\ w_m \end{array} \!\!\!\begin{array}{c} \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_n \end{array} \\ \begin{pmatrix} & & A_{1,k} & & \\ & & \vdots & & \\ & & A_{m,k} & & \end{pmatrix} \end{array}.$$

我们只展示了上述矩阵的第 $k$ 列，于是此处矩阵各元素的第二个下标都是 $k$. 上面这种描述还会让你想出由 $\mathcal{M}(T)$ 计算 $Tv_k$ 的方法：先将第 $k$ 列中的各元素分别与它们左侧对应的 $w_j$ 相乘，再将所得各向量都加起来.

$\mathcal{M}(T)$ 的第 $k$ 列包含了将 $Tv_k$ 写成 $w_1, \ldots, w_m$ 的线性组合所需的各标量：

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j.$$

如果 $T$ 是从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的线性映射，那么除非另有说明，总假定所讨论的基都是标准基（其中第 $k$ 个向量的第 $k$ 个坐标是 1 而其他各坐标都是 0）. 如果你将 $\mathbf{F}^m$ 中的元素看成由 $m$ 个数组成的列，那么你可以将 $\mathcal{M}(T)$ 中的第 $k$ 列看成将 $T$ 作用于第 $k$ 个标准基向量所得.

如果 $T$ 是从 $n$ 维向量空间到 $m$ 维向量空间的线性映射，那么 $\mathcal{M}(T)$ 是 $m \times n$ 矩阵.

**3.32 例：从 $\mathbf{F}^2$ 到 $\mathbf{F}^3$ 的线性映射的矩阵**

假设 $T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$ 定义为

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

因为 $T(1, 0) = (1, 2, 7)$ 且 $T(0, 1) = (3, 5, 9)$，所以 $T$ 关于标准基的矩阵就是如下所示的 $3 \times 2$ 矩阵：

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}.$$

在研究 $\mathcal{P}_m(\mathbf{F})$ 时，除非上下文有另作说明，都取标准基 $1, x, x^2, \ldots, x^m$.

**3.33 例：从 $\mathcal{P}_3(\mathbf{R})$ 到 $\mathcal{P}_2(\mathbf{R})$ 的微分映射的矩阵**

假设 $D \in \mathcal{L}\big(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R})\big)$ 是由 $Dp = p'$ 所定义的微分映射. 因为 $(x^n)' = nx^{n-1}$，所以 $D$ 关于标准基的矩阵就是如下所示的 $3 \times 4$ 矩阵：

$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

## 矩阵的加法和标量乘法

在本节的剩余部分，假定 $U, V, W$ 是有限维的且对于其中每个向量空间都已选定了一个基. 于是对于每个从 $V$ 到 $W$ 的线性映射，我们都能够讨论它们（关于所选择的基）的矩阵.

两个线性映射之和的矩阵，是否等于这两个映射的矩阵之和呢？这个问题目前还没有意义，因为尽管我们已定义两个线性映射的和，但我们还没定义两个矩阵的和. 幸运的是，按一种很自然的方式来作出两矩阵之和的定义，它便具备我们期望的性质. 具体而言，我们给出以下定义.

**3.34　定义：矩阵加法（matrix addition）**

两个相同大小的矩阵之和，是将两矩阵对应位置上的元素相加所得的矩阵：

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix}$$
$$= \begin{pmatrix} A_{1,1}+C_{1,1} & \cdots & A_{1,n}+C_{1,n} \\ \vdots & & \vdots \\ A_{m,1}+C_{m,1} & \cdots & A_{m,n}+C_{m,n} \end{pmatrix}.$$

♣

在接下来这条结果中，假设对于 $S+T, S$ 和 $T$ 这三个线性映射都选取相同的基.

**3.35　线性映射之和的矩阵**

假设 $S,T \in \mathcal{L}(V,W)$. 那么 $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

♡

由定义即可验证上面这条结论，验证过程留给读者完成.

仍然假设我们已选取了某些基. 一标量与线性映射之积的矩阵，是否等于同一标量与该线性映射的矩阵之积？同样，这个问题现在也没有意义，因为我们还没有定义矩阵的标量乘法. 幸运的是，再次按自然的方式作出定义，所得定义仍有我们期望的性质.

**3.36　定义：矩阵的标量乘法（scalar multiplication of a matrix）**

一个标量和一个矩阵的乘积，是将该矩阵的各元素都乘以该标量所得的矩阵：

$$\lambda\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}.$$

♣

---

**3.37　例：矩阵的加法和标量乘法**

$$2\begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ -2 & 10 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ -1 & 16 \end{pmatrix}$$

---

在接下来的结果中，假设对于 $\lambda T$ 和 $T$ 这两个线性映射都选取相同的基.

**3.38　标量与线性映射之积的矩阵**

假设 $\lambda \in \mathbf{F}$ 且 $T \in \mathcal{L}(V,W)$. 那么 $\mathcal{M}(\lambda T) = \lambda\mathcal{M}(T)$.

♡

上述结论的验证同样留给读者完成.

因为现已定义了矩阵的加法和标量乘法，由此产生一个向量空间也就不足为奇了. 我们先引入个记号，给这个新的向量空间取个名字，然后得出这个新的向量空间的维数.

**3.39 记号：$\mathbf{F}^{m,n}$**

对于正整数 $m$ 和 $n$，各元素均属于 $\mathbf{F}$ 的所有 $m \times n$ 矩阵构成的集合记作 $\mathbf{F}^{m,n}$.

**3.40 $\dim \mathbf{F}^{m,n} = mn$**

假设 $m$ 和 $n$ 为正整数. 按上面定义的加法和标量乘法，$\mathbf{F}^{m,n}$ 是维数为 $mn$ 的向量空间.

**证明** 验证 $\mathbf{F}^{m,n}$ 是向量空间留给读者完成. 注意 $\mathbf{F}^{m,n}$ 的加法恒等元是所有元素都等于 0 的 $m \times n$ 矩阵.

读者也应自行验证，所有 $m \times n$ 矩阵中，那些除某元素等于 1 外其余各元素都等于 0 的互异矩阵，构成 $\mathbf{F}^{m,n}$ 的一个基. 这样的矩阵共有 $mn$ 个，因此 $\mathbf{F}^{m,n}$ 的维数等于 $mn$. ■

## 矩阵乘法

与之前一样，假设 $v_1, \ldots, v_n$ 是 $V$ 的基，$w_1, \ldots, w_m$ 是 $W$ 的基. 另设 $u_1, \ldots, u_p$ 是 $U$ 的基.

考虑线性映射 $T : U \to V$ 和 $S : V \to W$. 复合映射 $ST$ 是从 $U$ 到 $W$ 的映射. 那么 $\mathcal{M}(ST)$ 等于 $\mathcal{M}(S)\mathcal{M}(T)$ 吗？这个问题目前还没有意义，因为我们还没有定义两矩阵的乘积. 我们将选定一种矩阵乘法的定义，来让上面的问题必有肯定的回答. 现在来看该怎么做.

假设 $\mathcal{M}(S) = A$ 且 $\mathcal{M}(T) = B$. 对于 $1 \le k \le p$，我们有

$$\begin{aligned}
(ST)u_k &= S\left(\sum_{r=1}^{n} B_{r,k} v_r\right) \\
&= \sum_{r=1}^{n} B_{r,k} S v_r \\
&= \sum_{r=1}^{n} B_{r,k} \sum_{j=1}^{m} A_{j,r} w_j \\
&= \sum_{j=1}^{m} \left(\sum_{r=1}^{n} A_{j,r} B_{r,k}\right) w_j.
\end{aligned}$$

于是 $\mathcal{M}(ST)$ 是 $m \times p$ 矩阵，其中第 $j$ 行第 $k$ 列的元素等于

$$\sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

这样我们就明白了如何定义矩阵乘法才能使我们期望的等式 $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$ 成立.

**3.41 定义：矩阵乘法（matrix multiplication）**

假设 $A$ 是 $m \times n$ 矩阵且 $B$ 是 $n \times p$ 矩阵. 那么 $AB$ 定义为一个 $m \times p$ 矩阵，其中第 $j$ 行第 $k$ 列的元素由下式给出：

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

于是，取 $A$ 的第 $j$ 行和 $B$ 的第 $k$ 列，将它们对应位置上的元素分别相乘再相加，就得到了 $AB$ 第 $j$ 行第 $k$ 列的元素.

注意，只有当第一个矩阵的列数等于第二个矩阵的行数时，我们才能定义这两个矩阵的乘积.

你可能在之前的课程中已经学过矩阵乘法的定义，尽管你可能并没看出这里所说的定义它的动机.

**3.42 例：矩阵乘积**

这里，我们将一个 3×2 矩阵和一个 2×4 矩阵相乘，得到一个 3×4 矩阵：

$$\begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix}\begin{pmatrix}6 & 5 & 4 & 3\\ 2 & 1 & 0 & -1\end{pmatrix}=\begin{pmatrix}10 & 7 & 4 & 1\\ 26 & 19 & 12 & 5\\ 42 & 31 & 20 & 9\end{pmatrix}.$$

矩阵乘法不满足交换律——$AB$ 并不一定等于 $BA$，即使两个乘积都有定义（见习题 10）. 矩阵乘法满足分配律和结合律（见习题 11、12）.

在接下来的结果中，我们假设，考虑 $T \in \mathcal{L}(U,V)$ 和 $S \in \mathcal{L}(V,W)$ 时取 $V$ 中的同一个基，考虑 $S \in \mathcal{L}(V,W)$ 和 $ST \in \mathcal{L}(U,W)$ 时取 $W$ 中的同一个基，考虑 $T \in \mathcal{L}(U,V)$ 和 $ST \in \mathcal{L}(U,W)$ 时取 $U$ 中的同一个基.

**3.43 线性映射之积的矩阵**

如果 $T \in \mathcal{L}(U,V)$ 且 $S \in \mathcal{L}(V,W)$，那么 $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

上述结果的证明，就是在定义矩阵乘法之前，我们说明其动机时所做的计算.

在下面的记号中，注意，如往常一样，第一个下标代表行，第二个下标代表列；垂直居中的点“·”用于占位.

**3.44 记号：$A_{j,\cdot}$、$A_{\cdot,k}$**

假设 $A$ 是 $m \times n$ 矩阵.

- 如果 $1 \le j \le m$，那么 $A_{j,\cdot}$ 表示由 $A$ 的第 $j$ 行构成的 $1{\times}n$ 矩阵.
- 如果 $1 \le k \le n$，那么 $A_{\cdot,k}$ 表示由 $A$ 的第 $k$ 列构成的 $m{\times}1$ 矩阵.

**3.45 例：$A_{j,\cdot}$ 等于 $A$ 的第 $j$ 行且 $A_{\cdot,k}$ 等于 $A$ 的第 $k$ 列**

记号 $A_{2,\cdot}$ 表示 $A$ 的第 2 行，$A_{\cdot,2}$ 表示 $A$ 的第 2 列. 于是如果 $A=\begin{pmatrix}8 & 4 & 5\\ 1 & 9 & 7\end{pmatrix}$，那么 $A_{2,\cdot}=\begin{pmatrix}1 & 9 & 7\end{pmatrix}$，$A_{\cdot,2}=\begin{pmatrix}4\\ 9\end{pmatrix}$.

$1 \times n$ 矩阵和 $n \times 1$ 矩阵的乘积是 $1 \times 1$ 矩阵. 然而，我们常常会把 $1 \times 1$ 矩阵与其元素等同起来看. 例如，

$$\begin{pmatrix}3 & 4\end{pmatrix}\begin{pmatrix}6\\ 2\end{pmatrix}=\begin{pmatrix}26\end{pmatrix}$$

因为 $3\cdot 6+4\cdot 2=26$. 然而，我们可以把 $(26)$ 与 26 等同起来看，也就是将上式写成 $\begin{pmatrix}3 & 4\end{pmatrix}\begin{pmatrix}6\\ 2\end{pmatrix}=26$.

接下来的结论就采用了上段里讨论的这个约定，从而给出了另一种思考矩阵乘法的方式. 举一实例：将接下来的结论结合上段所作的计算，就能解释为什么例 3.42 中所得乘积的第 2 行第 1 列中的元素等于 26.

**3.46 矩阵之积的元素等于行乘以列**

假设 $A$ 是 $m \times n$ 矩阵且 $B$ 是 $n \times p$ 矩阵. 那么如果 $1 \le j \le m$ 且 $1 \le k \le p$，则

$$(AB)_{j,k} = A_{j,\cdot}B_{\cdot,k}.$$

换言之，$AB$ 中第 $j$ 行第 $k$ 列的元素等于：($A$ 的第 $j$ 行) 乘以 ($B$ 的第 $k$ 列).

**证明** 假设 $1 \le j \le m$ 且 $1 \le k \le p$. 矩阵乘法的定义表明

$$(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}. \tag{3.47}$$

矩阵乘法的定义还表明，$1 \times n$ 矩阵 $A_{j,\cdot}$ 和 $n \times 1$ 矩阵 $B_{\cdot,k}$ 的乘积是一个 $1 \times 1$ 矩阵，且该矩阵的元素就是上述等式的右端. 于是 $AB$ 中第 $j$ 行第 $k$ 列的元素就等于：($A$ 的第 $j$ 行) 乘以 ($B$ 的第 $k$ 列). ■

接下来的结论又给出了另一种思考矩阵乘法的方式. 在下面的结论中，$(AB)_{\cdot,k}$ 是 $m \times p$ 矩阵 $AB$ 的第 $k$ 列，于是 $(AB)_{\cdot,k}$ 是 $m \times 1$ 矩阵. 同时，因为 $AB_{\cdot,k}$ 是 $m \times n$ 矩阵和 $n \times 1$ 矩阵的乘积，所以它也是 $m \times 1$ 矩阵. 于是下述结论中等式两侧的矩阵大小相同，猜测它们相等就是合理的了.

**3.48 矩阵之积的列等于矩阵与列之积**

假设 $A$ 是 $m \times n$ 矩阵且 $B$ 是 $n \times p$ 矩阵. 那么如果 $1 \le k \le p$，则

$$(AB)_{\cdot,k} = AB_{\cdot,k}.$$

换言之，$AB$ 的第 $k$ 列等于 $A$ 乘以 $B$ 的第 $k$ 列.

**证明** 上面已讨论过，$(AB)_{\cdot,k}$ 和 $AB_{\cdot,k}$ 都是 $m \times 1$ 矩阵. 如果 $1 \le j \le m$，那么 $(AB)_{\cdot,k}$ 中第 $j$ 行的元素是式 (3.47) 的左端项，$AB_{\cdot,k}$ 中第 $j$ 行的元素是式 (3.47) 的右端项. 于是 $(AB)_{\cdot,k} = AB_{\cdot,k}$. ■

接下来的结论由下面例子所启发，给出了另一种思考 $m \times n$ 矩阵和 $n \times 1$ 矩阵之积的方式.

**3.49 例：$3 \times 2$ 矩阵和 $2 \times 1$ 矩阵的乘积**

利用定义和基础的算术运算，我们可验证下式：

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 19 \\ 31 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 1\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}.$$

于是在本例中，$3 \times 2$ 矩阵和 $2 \times 1$ 矩阵的乘积是 $3 \times 2$ 矩阵各列的线性组合，而与各列相乘的标量（5 和 1）则来自 $2 \times 1$ 矩阵.

下面的结论推广了上述例子.

**3.50 列的线性组合**

假设 $A$ 是 $m\times n$ 矩阵且 $b=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$ 是 $n\times 1$ 矩阵. 那么

$$Ab=b_1A_{\cdot,1}+\cdots+b_nA_{\cdot,n}.$$

换言之，$Ab$ 是 $A$ 中各列的线性组合，而与这些列相乘的标量则来自 $b$.

**证明** 如果 $k\in\{1,\ldots,m\}$，那么由矩阵乘法的定义知，所得 $m\times 1$ 矩阵 $Ab$ 的第 $k$ 行中的元素是

$$A_{k,1}b_1+\cdots+A_{k,n}b_n.$$

而 $b_1A_{\cdot,1}+\cdots+b_nA_{\cdot,n}$ 的第 $k$ 行中的元素也等于该数.

因为对于每个 $k\in\{1,\ldots,m\}$，$Ab$ 和 $b_1A_{\cdot,1}+\cdots+b_nA_{\cdot,n}$ 的第 $k$ 行元素都相等，所以我们得出结论 $Ab=b_1A_{\cdot,1}+\cdots+b_nA_{\cdot,n}$. ■

前述两个结论重点关注矩阵的列. 类似的结论对于矩阵的行也是成立的，详见本节习题 8、9. 只要适当修改 3.48 和 3.50 的证明过程，即可证明它们.

接下来的结论，是在下个小节中证明行列分解（3.56）及矩阵的列秩等于行秩（3.57）时，所用的主要工具. 为了和叙述行列分解时所常用的记号保持一致，下面结论乃至下个小节中的矩阵记作 $C$ 和 $R$ 而不是 $A$ 和 $B$.

**3.51 将矩阵乘法视为列或行的线性组合**

假设 $C$ 是 $m\times c$ 矩阵且 $R$ 是 $c\times n$ 矩阵.

(a) 如果 $k\in\{1,\ldots,n\}$，那么 $CR$ 的第 $k$ 列是 $C$ 的各列的线性组合，其中各系数来自 $R$ 的第 $k$ 列.

(b) 如果 $j\in\{1,\ldots,m\}$，那么 $CR$ 的第 $j$ 行是 $R$ 的各行的线性组合，其中各系数来自 $C$ 的第 $j$ 行.

**证明** 假设 $k\in\{1,\ldots,n\}$. 那么 $CR$ 的第 $k$ 列等于 $CR_{\cdot,k}$（由 3.48）. 而 $CR_{\cdot,k}$ 还等于 $C$ 中各列的线性组合，其中系数取自 $R_{\cdot,k}$（由 3.50）. 于是 (a) 成立.

要证明 (b)，仿照 (a) 的证明即可，只需用行代替列，并用本节习题 8、9 的结论代替 3.48 和 3.50. ■

## 行列分解和矩阵的秩

我们从定义与每个矩阵都相关的两个非负整数开始.

**3.52 定义：列秩（column rank）、行秩（row rank）**

假设 $A$ 是 $m\times n$ 矩阵，其各元素属于 $\mathbf{F}$.

- $A$ 的**列秩**是 $A$ 的各列在 $\mathbf{F}^{m,1}$ 中的张成空间的维数.
- $A$ 的**行秩**是 $A$ 的各行在 $\mathbf{F}^{1,n}$ 中的张成空间的维数.

如果 $A$ 是 $m \times n$ 矩阵，那么 $A$ 的列秩不超过 $n$（因为 $A$ 有 $n$ 列）同时不超过 $m$（因为 $\dim \mathbf{F}^{m,1} = m$）．类似地，$A$ 的行秩也不超过 $\min\{m, n\}$．

**3.53 例：一个 $2 \times 4$ 矩阵的列秩和行秩**

假设

$$A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 3 & 5 & 2 & 9 \end{pmatrix}.$$

$A$ 的列秩是 $\mathbf{F}^{2,1}$ 中的

$$\operatorname{span}\left(\begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix}\right)$$

的维数．上面列出的 $\mathbf{F}^{2,1}$ 中各向量的前两个不成标量倍数关系．于是这个长度为 4 的向量组的张成空间的维数至少是 2．而该向量组在 $\mathbf{F}^{2,1}$ 中，则其张成空间的维数不超过 2，因为 $\dim \mathbf{F}^{2,1} = 2$．于是这个组的张成空间的维数就是 2，意味着 $A$ 的列秩是 2.

$A$ 的行秩是 $\mathbf{F}^{1,4}$ 中的

$$\operatorname{span}\big((4 \quad 7 \quad 1 \quad 8), (3 \quad 5 \quad 2 \quad 9)\big)$$

的维数．上面列出的 $\mathbf{F}^{1,4}$ 中的两个向量不成标量倍数关系．于是这个长度为 2 的向量组的张成空间的维数是 2，这意味着 $A$ 的行秩是 2.

我们现在定义矩阵的转置.

**3.54 定义：转置（transpose）、$A^{\mathrm{t}}$**

矩阵 $A$ 的**转置**记为 $A^{\mathrm{t}}$，是互换 $A$ 的行和列所得的矩阵．具体地说，如果 $A$ 是 $m \times n$ 矩阵，那么 $A^{\mathrm{t}}$ 是 $n \times m$ 矩阵，其中各元素由下面等式给出：

$$(A^{\mathrm{t}})_{k,j} = A_{j,k}.$$

♣

**3.55 例：矩阵的转置**

如果 $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$，那么 $A^{\mathrm{t}} = \begin{pmatrix} 5 & 3 & -4 \\ -7 & 8 & 2 \end{pmatrix}$.

注意此处 $A$ 是 $3 \times 2$ 矩阵而 $A^{\mathrm{t}}$ 是 $2 \times 3$ 矩阵.

矩阵的转置有着很好的代数性质：对于所有 $m \times n$ 矩阵 $A, B$、所有 $\lambda \in \mathbf{F}$ 和所有 $n \times p$ 矩阵 $C$，都有 $(A + B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$，$(\lambda A)^{\mathrm{t}} = \lambda A^{\mathrm{t}}$ 以及 $(AC)^{\mathrm{t}} = C^{\mathrm{t}} A^{\mathrm{t}}$（见本节习题 14、15）.

接下来的结果将是用于证明列秩等于行秩的主要工具（见 3.57）.

**3.56 行列分解（column-row factorization）**

假设 $A$ 是 $m \times n$ 矩阵，其中各元素均在 $\mathbf{F}$ 中且列秩 $c \geq 1$．那么存在各元素均属于 $\mathbf{F}$ 的 $m \times c$ 矩阵 $C$ 和 $c \times n$ 矩阵 $R$，使得 $A = CR$ 成立.

♡

**证明** $A$ 的各列都是 $m \times 1$ 矩阵. 由 $A$ 的各列构成的组 $A_{\cdot,1}, \dots, A_{\cdot,n}$ 可以被削减成 $A$ 的各列的张成空间的基(由 2.30). 由列秩的定义, 这个基的长度是 $c$. 将该基中的 $c$ 个列向量合在一起就形成了 $m \times c$ 矩阵 $C$.

如果 $k \in \{1, \dots, n\}$, 那么 $A$ 的第 $k$ 列是 $C$ 的各列的线性组合. 令该线性组合中的系数组成一个 $c \times n$ 矩阵的第 $k$ 列, 并记该矩阵为 $R$. 那么由 3.51 (a) 得, $A = CR$. ■

例 3.53 中的矩阵满足列秩等于行秩. 接下来的结论表明, 这一点对于每个矩阵都成立.

**3.57 列秩等于行秩**

假设 $A \in \mathbf{F}^{m,n}$. 那么 $A$ 的列秩等于 $A$ 的行秩. ♡

**证明** 令 $c$ 表示 $A$ 的列秩. 令 $A = CR$ 是由 3.56 给出的 $A$ 的行列分解式, 其中 $C$ 是 $m \times c$ 矩阵且 $R$ 是 $c \times n$ 矩阵. 那么 3.51 (b) 告诉我们, $A$ 的每一行都是 $R$ 的各行的线性组合. 因为 $R$ 有 $c$ 行, 所以这表明 $A$ 的行秩小于或等于 $c$(也就是 $A$ 的列秩).

为了证明上面这个不等关系反过来也成立, 将上一段中的结论应用于 $A^{\mathrm{t}}$, 可得

$$\begin{aligned} A \text{ 的列秩} &= A^{\mathrm{t}} \text{ 的行秩} \\ &\leq A^{\mathrm{t}} \text{ 的列秩} \\ &= A \text{ 的行秩}. \end{aligned}$$

于是 $A$ 的列秩就等于 $A$ 的行秩. ■

由于列秩等于行秩, 所以我们无需使用"列秩"和"行秩"这两个术语, 而用更简单的术语"秩"代替它们, 其定义如下.

**3.58 定义: 秩(rank)**

矩阵 $A \in \mathbf{F}^{m,n}$ 的**秩**是 $A$ 的列秩. ♣

列秩等于行秩的其他证明见于 3.133 和 7A 节的习题 8.

## 习题 3C

**1** 设 $T \in \mathcal{L}(V, W)$. 证明: 任取 $V$ 和 $W$ 的基, $T$ 所对应的矩阵都至少有 $\dim \operatorname{range} T$ 个非零元素.

**2** 设 $T \in \mathcal{L}(V, W)$, 其中 $V$ 和 $W$ 都是有限维且非零的向量空间. 证明: $\dim \operatorname{range} T = 1$, 当且仅当存在 $V$ 的一个基和 $W$ 的一个基, 使得关于这些基的 $\mathcal{M}(T)$ 的所有元素都是 1.

**3** 设 $v_1, \dots, v_n$ 是 $V$ 的基, $w_1, \dots, w_m$ 是 $W$ 的基.

(a) 证明: 如果 $S, T \in \mathcal{L}(V, W)$, 那么 $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

(b) 证明: 如果 $\lambda \in \mathbf{F}$, $T \in \mathcal{L}(V, W)$, 那么 $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

**注** 本题是在让你验证 3.35 和 3.38.

**4** 设 $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ 是微分映射，定义为 $Dp = p'$. 求 $\mathcal{P}_3(\mathbf{R})$ 的一个基和 $\mathcal{P}_2(\mathbf{R})$ 的一个基，使得 $D$ 关于这些基的矩阵为：

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

注 和例 3.33 比较一下. 下一题推广了本题.

**5** 设 $V$ 和 $W$ 是有限维的，$T \in \mathcal{L}(V, W)$. 证明：存在 $V$ 的一个基和 $W$ 的一个基，使得关于这些基的 $\mathcal{M}(T)$，除了在第 $k$ 行第 $k$ 列（$1 \le k \le \dim \operatorname{range} T$）的元素等于 1 外，其他所有元素都是 0.

**6** 设 $v_1, \ldots, v_m$ 是 $V$ 的基，$W$ 是有限维的. 设 $T \in \mathcal{L}(V, W)$. 证明：存在 $W$ 的一个基 $w_1, \ldots, w_n$ 使得 $\mathcal{M}(T)$【关于基 $v_1, \ldots, v_m$ 和 $w_1, \ldots, w_n$】除了第一行第一列可能有个 1 以外，第一列所有元素都是 0.

注 不同于第 5 题，在本题中，$V$ 的基是给定的而不是由你选择的.

**7** 设 $w_1, \ldots, w_n$ 是 $W$ 的基，$V$ 是有限维的. $T \in \mathcal{L}(V, W)$. 证明：存在 $V$ 的一个基 $v_1, \ldots, v_m$ 使得 $\mathcal{M}(T)$【关于基 $v_1, \ldots, v_m$ 和 $w_1, \ldots, w_n$】除了第一行第一列可能有个 1 以外，第一行所有元素都是 0.

注 不同于第 5 题，在本题中，$W$ 的基是给定的而不是由你选择的.

**8** 设 $A$ 是 $m \times n$ 矩阵，$B$ 是 $n \times p$ 矩阵，证明：

$$(AB)_{j,\cdot} = A_{j,\cdot}B$$

对 $1 \le j \le m$ 均成立. 换句话说，证明 $AB$ 的第 $j$ 行等于 $A$ 的第 $j$ 行乘以 $B$.

注 本题给出了 3.48 的行版本.

**9** 设 $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ 是 $1 \times n$ 矩阵，$B$ 是 $n \times p$ 矩阵. 证明：

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}.$$

换句话说，证明：$aB$ 是 $B$ 各行的线性组合，各行所乘的标量来自 $a$.

注 本题给出了 3.50 的行版本.

**10** 给出一例：使得 $AB \ne BA$ 的 $2 \times 2$ 矩阵 $A$ 和 $B$.

**11** 证明分配性质对矩阵加法和矩阵乘法仍成立. 换句话说，设矩阵 $A, B, C, D, E, F$ 的大小使得 $A(B+C)$ 和 $(D+E)F$ 都有意义，解释为什么 $AB + AC$ 和 $DF + EF$ 都有意义，并证明：

$$A(B+C) = AB + AC \quad 且 \quad (D+E)F = DF + EF.$$

**12** 证明矩阵乘法是可结合的. 换句话说，设矩阵 $A, B, C$ 的大小使得 $(AB)C$ 有意义. 解释为什么 $A(BC)$ 有意义，并证明：

$$(AB)C = A(BC).$$

注 试着找到一个简洁的证明，来说明下面引用的埃米尔·阿廷（Emil Artin）这句话："我的经验是，如果抛开矩阵的话，可将一个原本用矩阵完成的证明缩短 50%."

**13** 设 $A$ 是 $n \times n$ 矩阵，$1 \le j, k \le n$. 证明：$A^3$（定义为 $AAA$）的第 $j$ 行第 $k$ 列的元素是

$$\sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k}.$$

**14** 设 $m$ 和 $n$ 是正整数. 证明：函数 $A \mapsto A^{\mathrm{t}}$ 是从 $\mathbf{F}^{m,n}$ 到 $\mathbf{F}^{n,m}$ 的线性映射.

**15** 证明：如果 $A$ 是 $m \times n$ 矩阵，$C$ 是 $n \times p$ 矩阵，那么

$$(AC)^{\mathrm{t}} = C^{\mathrm{t}} A^{\mathrm{t}}.$$

注 本题表明，两矩阵乘积的转置，等于两矩阵各自转置然后按反方向相乘.

**16** 设 $A$ 是 $m \times n$ 矩阵（$A \neq 0$），证明：$A$ 的秩为 1，当且仅当存在 $(c_1, \ldots, c_m) \in \mathbf{F}^m$ 和 $(d_1, \cdots, d_n) \in \mathbf{F}^n$ 使得 $A_{j,k} = c_j d_k$ 对任意 $j = 1, \ldots, m$ 和任意 $k = 1, \ldots, n$ 都成立.

**17** 设 $T \in \mathcal{L}(V)$，$u_1, \ldots, u_n$ 和 $v_1, \ldots, v_n$ 是 $V$ 的基. 证明下列是等价的.

(a) $T$ 是单射.

(b) $\mathcal{M}(T)$ 的列在 $\mathbf{F}^{n,1}$ 中线性无关.

(c) $\mathcal{M}(T)$ 的列张成 $\mathbf{F}^{n,1}$.

(d) $\mathcal{M}(T)$ 的行张成 $\mathbf{F}^{1,n}$.

(e) $\mathcal{M}(T)$ 的行在 $\mathbf{F}^{1,n}$ 中线性无关.

这里的 $\mathcal{M}(T)$ 意即 $\mathcal{M}\big(T, (u_1, \ldots, u_n), (v_1, \ldots, v_n)\big)$.

# 3D 可逆性和同构

## 可逆线性映射

我们从线性映射的可逆和逆的概念开启这一节.

**3.59 定义：可逆的（invertible）、逆（inverse）**

- 对于线性映射 $T \in \mathcal{L}(V,W)$，如果存在线性映射 $S \in \mathcal{L}(W,V)$，使得 $ST$ 等于 $V$ 上的恒等算子且 $TS$ 等于 $W$ 上的恒等算子，则称 $T$ 是**可逆的**.
- 一个满足 $ST = I$ 及 $TS = I$ 的线性映射 $S \in \mathcal{L}(W,V)$ 被称为 $T$ 的一个**逆**.（注意，第一个 $I$ 是 $V$ 上的恒等算子，第二个 $I$ 是 $W$ 上的恒等算子）

上面的定义中用的说法是线性映射的“一个逆”. 然而，接下来的结果表明，我们可以将这个说法换成线性映射的“逆”. [7]

**3.60 逆是唯一的**

可逆的线性映射具有唯一的逆.

**证明** 假设 $T \in \mathcal{L}(V,W)$ 是可逆的，且 $S_1$ 和 $S_2$ 是 $T$ 的逆. 那么

$$S_1 = S_1 I = S_1(TS_2) = (S_1T)S_2 = IS_2 = S_2.$$

于是 $S_1 = S_2$. ■

既然我们知道逆是唯一的，那就可以给它一个记号.

**3.61 记号：$T^{-1}$**

如果 $T$ 是可逆的，那么它的逆记作 $T^{-1}$. 换言之，如果 $T \in \mathcal{L}(V,W)$ 是可逆的，那么 $T^{-1}$ 是 $\mathcal{L}(W,V)$ 中唯一使得 $T^{-1}T = I$ 和 $TT^{-1} = I$ 成立的元素.

**3.62 例：一个从 $\mathbf{R}^3$ 到 $\mathbf{R}^3$ 的线性映射的逆**

假设 $T \in \mathcal{L}(\mathbf{R}^3)$ 定义为 $T(x,y,z) = (-y,x,4z)$. 于是 $T$ 就将向量在 $xy$ 平面上逆时针旋转 $90°$ 并将其沿 $z$ 轴的分量拉伸至原先长度的 4 倍.

因此，逆映射 $T^{-1} \in \mathcal{L}(\mathbf{R}^3)$ 就将向量在 $xy$ 平面上顺时针旋转 $90°$ 并将其沿 $z$ 轴的分量压缩至原先长度的 $\frac{1}{4}$ 倍：

$$T^{-1}(x,y,z) = \left(y,-x,\tfrac{1}{4}z\right).$$

下个结果表明，一个线性映射是可逆的，当且仅当它既是单射又是满射.

**3.63 可逆性 $\Longleftrightarrow$ 单射性和满射性**

一个线性映射是可逆的，当且仅当它既是单射又是满射.

**证明** 假设 $T \in \mathcal{L}(V,W)$. 我们需要证明，$T$ 是可逆的当且仅当它既是单射又是满射.

[7] 原文：The definition above mentions “an inverse”. However, the next result shows that we can change this terminology to “the inverse”. 意思就是，本来是说“一个逆”（an inverse），证明了逆的唯一性之后，就可以将不定冠词 an 换成定冠词 the 来特指这唯一的逆了.

首先假设 $T$ 是可逆的. 为了证明 $T$ 是单射, 假设 $u,v\in V$ 且 $Tu=Tv$. 那么

$$u=T^{-1}(Tu)=T^{-1}(Tv)=v,$$

所以 $u=v$. 所以 $T$ 是单射.

仍假设 $T$ 是可逆的, 现在我们欲证 $T$ 是满射. 为此, 令 $w\in W$. 那么 $w=T(T^{-1}w)$, 这就表明 $w$ 在 $T$ 的值域中. 于是 $\operatorname{range}T=W$. 所以 $T$ 是满射, 这就完成了此方向的证明.

现在假设 $T$ 既是单射又是满射. 我们欲证 $T$ 是可逆的. 对于每个 $w\in W$, 定义 $S(w)$ 是 $V$ 中唯一使得 $T\big(S(w)\big)=w$ 成立的元素 (这样一个元素的存在性和唯一性分别来自 $T$ 的满射性和单射性). 根据 $S$ 的定义可得, $T\circ S$ 就等于 $W$ 上的恒等算子.

为了证明 $S\circ T$ 等于 $V$ 上的恒等算子, 令 $v\in V$. 那么

$$T\big((S\circ T)v\big)=(T\circ S)(Tv)=I(Tv)=Tv.$$

此式表明 $(S\circ T)v=v$ (因为 $T$ 是单射). 从而 $S\circ T$ 等于 $V$ 上的恒等算子.

为了完成整个证明, 我们还需要证明 $S$ 是线性的. 为此, 假设 $w_1,w_2\in W$. 那么

$$T\big(S(w_1)+S(w_2)\big)=T\big(S(w_1)\big)+T\big(S(w_2)\big)=w_1+w_2.$$

于是 $S(w_1)+S(w_2)$ 是 $V$ 中唯一被 $T$ 映射到 $w_1+w_2$ 的元素. 根据 $S$ 的定义, 这意味着 $S(w_1+w_2)=S(w_1)+S(w_2)$. 所以 $S$ 满足可加性, 这是证明其线性所必需的.

齐次性的证明也是类似的. 具体而言, 如果 $w\in W$ 且 $\lambda\in\mathbf{F}$, 那么

$$T\big(\lambda S(w)\big)=\lambda T\big(S(w)\big)=\lambda w.$$

于是 $\lambda S(w)$ 是 $V$ 中唯一被 $T$ 映射到 $\lambda w$ 的元素. 由 $S$ 的定义, 这意味着 $S(\lambda w)=\lambda S(w)$. 所以 $S$ 是线性的, 则原命题得证. ■

你可能会好奇, 对于从某向量空间映射到该空间本身的线性映射, 是否仅凭单射性或满射性就可以推出可逆性. 在无限维向量空间中, 只凭两个条件之一不能推出可逆性. 接下来这个例子, 就用例 3.3 中我们所熟知的两个线性映射, 说明了这一点.

**3.64 例: 仅凭单射性或满射性不能推出可逆性**

- 从 $\mathcal{P}(\mathbf{R})$ 到 $\mathcal{P}(\mathbf{R})$ 的与 $x^2$ 相乘线性映射 (见例 3.3) 是单射, 但它不可逆, 因为它不是满射 (多项式 1 不在其值域内).
- 从 $\mathbf{F}^\infty$ 到 $\mathbf{F}^\infty$ 的后向移位线性映射 (见例 3.3) 是满射, 但它不可逆, 因为它不是单射【向量 $(1,0,0,0,\ldots)$ 在其零空间中】.

鉴于上面例子, 下个结果就非同寻常——它指出, 对于从某有限维向量空间映射到与之维数相同的向量空间的线性映射, 单射性和满射性可以互推. 注意, 对于 $V$ 是有限维的且 $W=V$ 这一特殊情形, 下面的前提条件 $\dim V=\dim W$ 自动满足.

**3.65 若 $\dim V=\dim W<\infty$, 则单射性与满射性等价**

假设 $V$ 和 $W$ 都是有限维向量空间, $\dim V=\dim W$, 且 $T\in\mathcal{L}(V,W)$. 那么

$$T\text{ 可逆}\iff T\text{ 是单射}\iff T\text{ 是满射}.$$

**证明** 线性映射基本定理（3.21）表明

$$\dim V=\dim\operatorname{null}T+\dim\operatorname{range}T. \tag{3.66}$$

如果 $T$ 是单射（由 3.15，这就等价于 $\dim\operatorname{null}T=0$），那么由上式可得

$$\dim\operatorname{range}T=\dim V-\dim\operatorname{null}T=\dim V=\dim W,$$

这就表明 $T$ 是满射（根据 2.39）.

反之，如果 $T$ 是满射，那么式 (3.66) 就表明

$$\dim\operatorname{null}T=\dim V-\dim\operatorname{range}T=\dim V-\dim W=0,$$

这就表明 $T$ 是单射.

于是，我们证明了 $T$ 是单射当且仅当 $T$ 是满射. 于是只要 $T$ 满足单射性或满射性之一，那么 $T$ 就既是单射又是满射，由此可推出 $T$ 是可逆的. 所以，$T$ 是可逆的当且仅当 $T$ 是单射当且仅当 $T$ 是满射. ■

下面的例子说明了上述结论有多么强大. 尽管不用线性代数也可以证明下例中的结论，但是利用线性代数能让证明更为简洁容易.

---

**3.67 例：存在多项式 $p$ 使得 $\big((x^2+5x+7)p\big)''=q$**

从 $\mathcal{P}(\mathbf{R})$ 到该空间本身的线性映射

$$p\mapsto\big((x^2+5x+7)p\big)''$$

是单射，你可自行证明. 于是我们想用 3.65 来证明这个映射是满射. 然而，例 3.64 表明，不能在无限维向量空间 $\mathcal{P}(\mathbf{R})$ 上发挥 3.65 之妙处. 我们通过将考虑的范围限制在有限维向量空间 $\mathcal{P}_m(\mathbf{R})$ 来绕开这个问题.

假设 $q\in\mathcal{P}(\mathbf{R})$. 存在非负整数 $m$ 使得 $q\in\mathcal{P}_m(\mathbf{R})$. 定义 $T:\mathcal{P}_m(\mathbf{R})\to\mathcal{P}_m(\mathbf{R})$ 为

$$Tp=\big((x^2+5x+7)p\big)''.$$

将一个非零的多项式乘以 $(x^2+5x+7)$ 会将其次数增加 2，再求两次微分又会将其次数降低 2. 于是 $T$ 的确是从 $\mathcal{P}_m(\mathbf{R})$ 到该空间本身的线性映射.

二阶导函数等于 0 的每个多项式都形如 $ax+b$，其中 $a,b\in\mathbf{R}$. 于是 $\operatorname{null}T=\{0\}$，所以 $T$ 是单射，进而 $T$ 是满射（由 3.65），这意味着存在一个多项式 $p\in\mathcal{P}_m(\mathbf{R})$ 使得 $\big((x^2+5x+7)p\big)''=q$，正如本例标题所言.

---

6A 节的习题 35 给出了类似但是更精彩的 3.65 应用实例.

下面结论中的前提 $\dim V=\dim W$，对于 $V$ 是有限维的且 $W=V$ 这一特殊情形，是自动满足的. 那么，在这种情况下，等式 $ST=I$ 蕴涵 $ST=TS$，即便乘法交换律并不对于从 $V$ 到 $V$ 的任意线性映射都成立.

**3.68 $ST=I\iff TS=I$（$S$ 和 $T$ 作用于维数相同的向量空间）**

假设 $V$ 和 $W$ 是维数相同的有限维向量空间，$S\in\mathcal{L}(W,V)$ 且 $T\in\mathcal{L}(V,W)$. 那么 $ST=I$ 当且仅当 $TS=I$.

**证明** 先设 $ST = I$. 如果 $v \in V$ 且 $Tv = 0$，那么

$$v = Iv = (ST)v = S(Tv) = S(0) = 0.$$

于是 $T$ 是单射（由 3.15）. 因为 $V$ 和 $W$ 维数相同，故可得 $T$ 是可逆的（由 3.65）.

现在将 $ST = I$ 的两侧都右乘 $T^{-1}$，得

$$S = T^{-1}.$$

于是 $TS = TT^{-1} = I$.

为了证明另一方向的蕴涵关系，只要将上面的证明中 $S$ 和 $T$ 以及 $V$ 和 $W$ 的角色互换，即可证明如果 $TS = I$ 那么 $ST = I$. ■

## 同构向量空间

接下来的定义有助于描述除了元素名称不同外本质上相同的两个向量空间.

**3.69 定义：同构（isomorphism）、同构的（isomorphic）**

- **同构**就是可逆线性映射.
- 对于两个向量空间，若存在将其中一个向量空间映成另一个向量空间的同构，则称它们是**同构的**[8].

可以将同构 $T : V \to W$ 理解为把 $v \in V$ 改写成 $Tv \in W$. 这个观点能解释为何两个同构的向量空间有相同的向量空间性质. “同构”和“可逆线性映射”这两个术语同义，当你想强调两个空间本质上相同时，就用“同构”这个词.

对于两个数学结构（例如群或拓扑空间），要判定它们（除了基础集合[9]中元素的名称不同外）在本质上相同，可能很困难. 然而，接下来的结论说明了，要判定两个向量空间是否同构，我们只需要关注一个数——维数——即可.

**3.70 维数表明了向量空间是否同构**

对于 $\mathbf{F}$ 上的两个有限维向量空间，当且仅当它们的维数相同时，它们才是同构的.

**证明** 先设 $V$ 和 $W$ 是同构的有限维向量空间. 于是，存在将 $V$ 映成 $W$ 的同构 $T$. 因为 $T$ 是可逆的，所以我们有 $\operatorname{null} T = \{0\}$ 且 $\operatorname{range} T = W$. 于是

$$\dim \operatorname{null} T = 0 \quad 且 \quad \dim \operatorname{range} T = \dim W.$$

公式

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

（线性映射基本定理，3.21）就可以化成 $\dim V = \dim W$. 这就完成了一个方向的证明.

---

[8] 我们说“同构”的时候，有时候是指“可逆线性映射”（isomorphism），有时候是指向量空间之间的关系（being isomorphic）. 两者词性不同，读者应该不难分辨.

[9] 这里的数学结构（mathematical structure）指的是带有特定运算的集合，而基础集合（underlying set）指的是未附加这些运算的集合本身.

为了证明另一方向，假设 $V$ 和 $W$ 是维数相同的有限维向量空间．令 $v_1,\ldots,v_n$ 是 $V$ 的基，$w_1,\ldots,w_n$ 是 $W$ 的基．令 $T\in\mathcal{L}(V,W)$ 定义为

$$T(c_1v_1+\cdots+c_nv_n)=c_1w_1+\cdots+c_nw_n.$$

那么由于 $v_1,\ldots,v_n$ 是 $V$ 的基，所以 $T$ 就是一个定义完善的线性映射．而且，由于 $w_1,\ldots,w_n$ 张成 $W$，所以 $T$ 是满射；由于 $w_1,\ldots,w_n$ 线性无关，可得 $\operatorname{null}T=\{0\}$，所以 $T$ 是单射．因为 $T$ 既是单射又是满射，所以它就是同构（见 3.63）．所以 $V$ 和 $W$ 是同构的． ■

上面结论表明，每个有限维向量空间 $V$ 都与 $\mathbf{F}^n$ 同构（其中 $n=\dim V$）．例如，如果 $m$ 是非负整数，那么 $\mathcal{P}_m(\mathbf{F})$ 与 $\mathbf{F}^{m+1}$ 同构．

回顾一下，记号 $\mathbf{F}^{m,n}$ 代表各元素都属于 $\mathbf{F}$ 的 $m\times n$ 矩阵所构成的向量空间．如果 $v_1,\ldots,v_n$ 是 $V$ 的基且 $w_1,\ldots,w_m$ 是 $W$ 的基，那么每个 $T\in\mathcal{L}(V,W)$ 都有一对应的矩阵 $\mathcal{M}(T)\in\mathbf{F}^{m,n}$．于是，一旦为 $V$ 和 $W$ 取定了基，$\mathcal{M}$ 就成为了一个从 $\mathcal{L}(V,W)$ 到 $\mathbf{F}^{m,n}$ 的函数．又注意到 3.35 和 3.38 表明 $\mathcal{M}$ 是一个线性映射．现在我们证明，这个线性映射其实是一个同构．

> 既然每个有限维向量空间都与某个 $\mathbf{F}^n$ 同构，那么为何不只研究 $\mathbf{F}^n$，而还要研究更一般的向量空间呢？要回答这个问题，就要看到对 $\mathbf{F}^n$ 的研究势必引入其他向量空间．例如，我们会碰上线性映射的零空间和值域．尽管这些向量空间都分别和某个 $\mathbf{F}^m$ 同构，但这样考虑问题往往更复杂且并不带来新的见解．

**3.71 $\mathcal{L}(V,W)$ 与 $\mathbf{F}^{m,n}$ 是同构的**

设 $v_1,\ldots,v_n$ 是 $V$ 的基且 $w_1,\ldots,w_m$ 是 $W$ 的基．那么 $\mathcal{M}$ 是 $\mathcal{L}(V,W)$ 与 $\mathbf{F}^{m,n}$ 间的同构．

**证明** 我们已经注意到 $\mathcal{M}$ 是线性的，则只需证明 $\mathcal{M}$ 既是单射又是满射．

我们从证明单射性开始．如果 $T\in\mathcal{L}(V,W)$ 且 $\mathcal{M}(T)=0$，那么对每个 $k=1,\ldots,n$ 均有 $Tv_k=0$．因为 $v_1,\ldots,v_n$ 是 $V$ 的基，故可得 $T=0$．于是 $\mathcal{M}$ 是单射（由 3.15）．

为了证明 $\mathcal{M}$ 是满射，假设 $A\in\mathbf{F}^{m,n}$．由线性映射引理（3.4），存在 $T\in\mathcal{L}(V,W)$ 使得对每个 $k=1,\ldots,n$ 均有

$$Tv_k=\sum_{j=1}^{m}A_{j,k}w_j.$$

因为 $\mathcal{M}(T)$ 等于 $A$，所以 $\mathcal{M}$ 的值域等于 $\mathbf{F}^{m,n}$，满射性得证． ■

现在我们可以确定两个有限维向量空间之间的线性映射所构成的向量空间的维数了．

**3.72 $\dim\mathcal{L}(V,W)=(\dim V)(\dim W)$**

假设 $V$ 和 $W$ 是有限维的．那么 $\mathcal{L}(V,W)$ 是有限维的，且

$$\dim\mathcal{L}(V,W)=(\dim V)(\dim W).$$

**证明** 由 3.71、3.70 和 3.40 可证． ■

## 将线性映射视为矩阵乘法

之前我们定义了线性映射的矩阵，现在我们定义向量的矩阵.

**3.73 定义：向量的矩阵（matrix of a vector）、$\mathcal{M}(v)$**

假设 $v \in V$ 且 $v_1, \ldots, v_n$ 是 $V$ 的基. $v$ 关于该基的**矩阵**是 $n \times 1$ 矩阵

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

其中 $b_1, \ldots, b_n$ 是使得下式成立的标量：

$$v = b_1 v_1 + \cdots + b_n v_n.$$

♣

向量 $v \in V$ 的矩阵 $\mathcal{M}(v)$ 取决于 $V$ 的基 $v_1, \ldots, v_n$，也取决于 $v$. 然而，由上下文应该可明确基取什么，因此它没被包含在记号里.

**3.74 例：向量的矩阵**

- 多项式 $2 - 7x + 5x^3 + x^4$ 关于 $\mathcal{P}_4(\mathbf{R})$ 的标准基的矩阵是

$$\begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$

- 一向量 $x \in \mathbf{F}^n$ 关于标准基的矩阵可通过将 $x$ 的各坐标写成 $n \times 1$ 矩阵的各元素来获得. 换言之，如果 $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$，那么

$$\mathcal{M}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

有时，我们想把 $V$ 中的元素改写成 $n \times 1$ 矩阵. 一旦取定一个基 $v_1, \ldots, v_n$，将 $v \in V$ 对应到 $\mathcal{M}(v)$ 的函数 $\mathcal{M}$ 就是（用以实现这种改写的）将 $V$ 映成 $\mathbf{F}^{n,1}$ 的同构.

回忆一下，如果 $A$ 是一个 $m \times n$ 矩阵，那么 $A_{\cdot,k}$ 代表 $A$ 的第 $k$ 列，我们可将它看成一个 $m \times 1$ 矩阵. 下面结论中，$\mathcal{M}(Tv_k)$ 是关于 $W$ 的基 $w_1, \ldots, w_m$ 来计算的.

**3.75 $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$**

设 $T \in \mathcal{L}(V, W)$，$v_1, \ldots, v_n$ 是 $V$ 的基且 $w_1, \ldots, w_m$ 是 $W$ 的基. 令 $1 \le k \le n$. 那么 $\mathcal{M}(T)$ 的第 $k$ 列，记作 $\mathcal{M}(T)_{\cdot,k}$，就等于 $\mathcal{M}(Tv_k)$.

♡

**证明** 由 $\mathcal{M}(T)$ 和 $\mathcal{M}(Tv_k)$ 的定义可立刻推得上述结论. ■

下面结果展示了线性映射的矩阵、向量的矩阵以及矩阵乘法的概念是怎样融合在一起的.

**3.76 线性映射的作用就像矩阵乘法**

假设 $T \in \mathcal{L}(V,W)$ 且 $v \in V$. 假设 $v_1,\dots,v_n$ 是 $V$ 的基且 $w_1,\dots,w_m$ 是 $W$ 的基. 那么

$$\mathcal{M}(Tv)=\mathcal{M}(T)\mathcal{M}(v).$$

**证明** 假设 $v=b_1v_1+\cdots+b_nv_n$，其中 $b_1,\dots,b_n\in\mathbf{F}$. 于是

$$Tv=b_1Tv_1+\cdots+b_nTv_n. \tag{3.77}$$

因此

$$\begin{aligned}\mathcal{M}(Tv)&=b_1\mathcal{M}(Tv_1)+\cdots+b_n\mathcal{M}(Tv_n)\\&=b_1\mathcal{M}(T)_{\cdot,1}+\cdots+b_n\mathcal{M}(T)_{\cdot,n}\\&=\mathcal{M}(T)\mathcal{M}(v),\end{aligned}$$

其中第一个等号来自式 (3.77) 和 $\mathcal{M}$ 的线性，第二个等号来自 3.75，最后一个等号来自 3.50. ■

每个 $m\times n$ 矩阵 $A$ 都能诱导出一个从 $\mathbf{F}^{n,1}$ 到 $\mathbf{F}^{m,1}$ 的线性映射，也就是将 $x\in\mathbf{F}^{n,1}$ 对应至 $Ax\in\mathbf{F}^{m,1}$ 的矩阵乘法函数. 利用上述结论，通过同构 $\mathcal{M}$ 作适当的改写后，可以将任何（从一个有限维向量空间到另一个有限维向量空间的）线性映射看成矩阵乘法映射. 具体而言，如果 $T\in\mathcal{L}(V,W)$ 且我们把 $v\in V$ 与 $\mathcal{M}(v)\in\mathbf{F}^{n,1}$ 等同起来看，那么上述结论就告诉我们，可以把 $Tv$ 和 $\mathcal{M}(T)\mathcal{M}(v)$ 等同起来看.

因为上述结论使我们得以（通过同构）将每个线性映射都看成在 $\mathbf{F}^{n,1}$ 上与某矩阵 $A$ 相乘，所以要谨记这个矩阵 $A$ 不仅依赖于该线性映射，还依赖于基的选取. 后续章节中，许多至关重要的结论的主旨之一，就是选取使矩阵 $A$ 尽可能简单的基.

在本书中，我们专注于线性映射而不是矩阵. 然而，有时将线性映射看成矩阵（或将矩阵看成线性映射）能得出许多重要且有用的见解.

注意，下面结论的表述中没出现“基”. 尽管下面结论中的 $\mathcal{M}(T)$ 依赖于 $V$ 和 $W$ 的基的选取，但是该结论表明，对于所有的选取方式来说，$\mathcal{M}(T)$ 的列秩总是相同的（因为 $\operatorname{range}T$ 并不依赖于基的选取）.

**3.78 $\operatorname{range}T$ 的维数等于 $\mathcal{M}(T)$ 的列秩**

假设 $V$ 和 $W$ 是有限维的，$T\in\mathcal{L}(V,W)$. 那么 $\dim\operatorname{range}T$ 等于 $\mathcal{M}(T)$ 的列秩.

**证明** 假设 $v_1,\dots,v_n$ 是 $V$ 的基且 $w_1,\dots,w_m$ 是 $W$ 的基. 将 $w\in W$ 对应到 $\mathcal{M}(w)$ 的线性映射是一个同构，它将 $W$ 映成由 $m\times1$ 向量构成的空间 $\mathbf{F}^{m,1}$. 将该同构限制于 $\operatorname{range}T$【由 3B 节的习题 10，它等于 $\operatorname{span}(Tv_1,\dots,Tv_n)$】，该同构就成为将 $\operatorname{range}T$ 映成 $\operatorname{span}(\mathcal{M}(Tv_1),\dots,\mathcal{M}(Tv_n))$ 的同构. 对于每个 $k\in\{1,\dots,n\}$，$m\times1$ 矩阵 $\mathcal{M}(Tv_k)$ 就等于 $\mathcal{M}(T)$ 的第 $k$ 列. 于是

$$\dim\operatorname{range}T=\mathcal{M}(T)\text{的列秩},$$

原命题得证. ■

## 换基

在 3C 节中，我们定义了由 $V$ 到可能与之不同的向量空间 $W$ 的线性映射 $T$ 的矩阵

$$\mathcal{M}\big(T,(v_1,\ldots,v_n),(w_1,\ldots,w_m)\big),$$

其中 $v_1,\ldots,v_n$ 是 $V$ 的基且 $w_1,\ldots,w_m$ 是 $W$ 的基．对于从一个向量空间到其本身的线性映射，我们通常为其定义空间和目标空间选择相同的基．这时，我们通常只将所选的基写出一次．换言之，如果 $T\in\mathcal{L}(V)$ 且 $v_1,\ldots,v_n$ 是 $V$ 的基，那么记号 $\mathcal{M}\big(T,(v_1,\ldots,v_n)\big)$ 定义如下式：

$$\mathcal{M}\big(T,(v_1,\ldots,v_n)\big)=\mathcal{M}\big(T,(v_1,\ldots,v_n),(v_1,\ldots,v_n)\big).$$

如果从上下文能确知基 $v_1,\ldots,v_n$ 取什么，那么我们只写 $\mathcal{M}(T)$ 即可．

**3.79 定义：恒等矩阵（identity matrix）、$I$**

设 $n$ 为正整数．仅对角线上（即那些行号和列号相等的位置）的元素为 1 而其他各元素均为 0 的 $n\times n$ 矩阵

$$\begin{pmatrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix}$$

就称为**恒等矩阵**，记作 $I$. ♣

在上述定义中，矩阵左下角的 0 表示对角线下方的所有元素都是 0，右上角的 0 则表示对角线上方的元素都是 0.

恒等算子 $I\in\mathcal{L}(V)$ 关于 $V$ 的每个基的矩阵都是恒等矩阵 $I$．注意，符号 $I$ 既被用于表示恒等算子，又被用于表示恒等矩阵．根据上下文即可知 $I$ 指的是什么意思．例如，等式 $\mathcal{M}(I)=I$ 中，左侧的 $I$ 表示恒等算子，右侧的 $I$ 则表示恒等矩阵.

如果 $A$ 是与 $I$ 大小相同的方阵（行数等于列数的矩阵），那么 $AI=IA=A$（你应自行验证）.

**3.80 定义：可逆的（invertible），逆（inverse）、$A^{-1}$**

称方阵 $A$ 是**可逆的**，如果存在与之大小相等的方阵 $B$ 使得 $AB=BA=I$．我们称 $B$ 是 $A$ 的**逆**且将其记为 $A^{-1}$. ♣

利用与证明 3.60 相同的方法，可以证明如果 $A$ 是可逆的方阵，那么就存在唯一的矩阵 $B$ 使得 $AB=BA=I$ 成立（由此亦得 $B=A^{-1}$ 这个记号是合理的）.

有些数学家使用术语 **“非奇异”（nonsingular）** 和 **“奇异”（singular）**，它们分别与“可逆”和“不可逆”同义.

如果 $A$ 是可逆矩阵，那么 $(A^{-1})^{-1}=A$，这是因为

$$A^{-1}A=AA^{-1}=I.$$

并且，如果 $A$ 和 $C$ 是相同大小的可逆方阵，那么 $AC$ 也可逆，且 $(AC)^{-1} = C^{-1}A^{-1}$，这是因为

$$\begin{aligned}(AC)(C^{-1}A^{-1}) &= A(CC^{-1})A^{-1}\\ &= AIA^{-1}\\ &= AA^{-1}\\ &= I,\end{aligned}$$

类似地，可得 $(C^{-1}A^{-1})(AC) = I$.

下面结果成立，是因为我们定义矩阵乘法就是为了使其成立——见 3.43 及其前面的内容. 现在我们只是把与之有关的基写得更明确些.

**3.81 线性映射之积的矩阵**

设 $T \in \mathcal{L}(U,V)$ 且 $S \in \mathcal{L}(V,W)$. 如果 $u_1,\dots,u_m$ 是 $U$ 的基，$v_1,\dots,v_n$ 是 $V$ 的基且 $w_1,\dots,w_p$ 是 $W$ 的基，那么

$$\begin{aligned}&\mathcal{M}\big(ST,(u_1,\dots,u_m),(w_1,\dots,w_p)\big)\\ &= \mathcal{M}\big(S,(v_1,\dots,v_n),(w_1,\dots,w_p)\big)\mathcal{M}\big(T,(u_1,\dots,u_m),(v_1,\dots,v_n)\big).\end{aligned}$$

接下来的结论讨论的是恒等算子 $I$ 关于两个不同的基的矩阵. 注意到

$$\mathcal{M}\big(I,(u_1,\dots,u_n),(v_1,\dots,v_n)\big)$$

的第 $k$ 列是由将 $u_k$ 写成基 $v_1,\dots,v_n$ 的线性组合所需的标量构成的.

下面结论的表述中，$I$ 表示由 $V$ 到 $V$ 的恒等算子. 在证明中，$I$ 也表示 $n\times n$ 恒等矩阵.

**3.82 恒等算子关于两个基的矩阵**

假设 $u_1,\dots,u_n$ 和 $v_1,\dots,v_n$ 是 $V$ 的两个基. 那么矩阵

$$\mathcal{M}\big(I,(u_1,\dots,u_n),(v_1,\dots,v_n)\big) \quad 和 \quad \mathcal{M}\big(I,(v_1,\dots,v_n),(u_1,\dots,u_n)\big)$$

都是可逆的，且互为对方的逆.

**证明** 在 3.81 中，将各 $w_k$ 替换成 $u_k$ 并将 $S$ 和 $T$ 换成 $I$，可得

$$I = \mathcal{M}\big(I,(v_1,\dots,v_n),(u_1,\dots,u_n)\big)\mathcal{M}\big(I,(u_1,\dots,u_n),(v_1,\dots,v_n)\big).$$

现在将 $u$ 和 $v$ 的角色互换，可得

$$I = \mathcal{M}\big(I,(u_1,\dots,u_n),(v_1,\dots,v_n)\big)\mathcal{M}\big(I,(v_1,\dots,v_n),(u_1,\dots,u_n)\big).$$

上面两个等式就给出了我们想要的结果. ■

---

**3.83 例：$\mathbf{F}^2$ 上的恒等算子关于两个基的矩阵**

考虑 $\mathbf{F}^2$ 的基 $(4,2),(5,3)$ 和 $(1,0),(0,1)$. 因为 $I(4,2) = 4(1,0)+2(0,1)$ 且 $I(5,3) = 5(1,0)+3(0,1)$，所以我们有

$$\mathcal{M}\Big(I,\big((4,2),(5,3)\big),\big((1,0),(0,1)\big)\Big) = \begin{pmatrix}4 & 5\\ 2 & 3\end{pmatrix}$$

转下页

你应自行验证，上述矩阵的逆是

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

于是 3.82 就表明

$$\mathcal{M}\Big(I, \big((1,0),(0,1)\big), \big((4,2),(5,3)\big)\Big) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

---

下面的结果展示了当我们改变所选的基时，$T$ 的矩阵是如何改变的．在下面结果中，我们取 $V$ 的两个不同的基，其中每个基都被同时用作 $T$ 的定义空间和目标空间的基．回忆一下，这种情况下，使用下面的简短记号就只需将基写出一次：

$$\mathcal{M}\big(T,(u_1,\ldots,u_n)\big) = \mathcal{M}\big(T,(u_1,\ldots,u_n),(u_1,\ldots,u_n)\big).$$

**3.84 换基公式（change-of-basis formula）**

设 $T \in \mathcal{L}(V)$．假设 $u_1,\ldots,u_n$ 和 $v_1,\ldots,v_n$ 都是 $V$ 的基．令

$$A = \mathcal{M}\big(T,(u_1,\ldots,u_n)\big) \quad 且 \quad B = \mathcal{M}\big(T,(v_1,\ldots,v_n)\big)$$

且 $C = \mathcal{M}\big(I,(u_1,\ldots,u_n),(v_1,\ldots,v_n)\big)$．那么

$$A = C^{-1}BC.$$

**证明** 在 3.81 中，将各 $w_k$ 替换成 $u_k$ 并将 $S$ 替换为 $I$，可得

$$A = C^{-1}\mathcal{M}\big(T,(u_1,\ldots,u_n),(v_1,\ldots,v_n)\big), \tag{3.85}$$

其中用到了 3.82.

再次利用 3.81，这次将各 $w_k$ 替换成 $v_k$，并将 $T$ 替换为 $I$、$S$ 替换为 $T$，可得

$$\mathcal{M}\big(T,(u_1,\ldots,u_n),(v_1,\ldots,v_n)\big) = BC.$$

将上式代入式 (3.85)，即得等式 $A = C^{-1}BC$. ■

下面结论的证明留作习题.

**3.86 逆的矩阵等于矩阵的逆**

设 $v_1,\ldots,v_n$ 是 $V$ 的基且 $T \in \mathcal{L}(V)$ 是可逆的．那么 $\mathcal{M}(T^{-1}) = \big(\mathcal{M}(T)\big)^{-1}$，式中两个矩阵均是关于基 $v_1,\ldots,v_n$ 的.

## 习题 3D

1 设 $T \in \mathcal{L}(V,W)$ 可逆，证明 $T^{-1}$ 可逆，且 $(T^{-1})^{-1} = T$.

2 设 $T \in \mathcal{L}(U,V)$ 和 $S \in \mathcal{L}(V,W)$ 都是可逆线性映射，证明：$ST \in \mathcal{L}(U,W)$ 可逆，且 $(ST)^{-1} = T^{-1}S^{-1}$.

3 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$．证明下列是等价的.

(a) $T$ 可逆.

(b) 对于 $T$ 的任一个基 $v_1,\dots,v_n$，有 $Tv_1,\dots,Tv_n$ 是 $V$ 的基.

(c) 对于 $T$ 的某一个基 $v_1,\dots,v_n$，有 $Tv_1,\dots,Tv_n$ 是 $V$ 的基.

**4** 设 $V$ 是有限维的，且 $\dim V>1$. 证明：从 $V$ 到自身的不可逆线性映射构成的集合不是 $\mathcal{L}(V)$ 的子空间.

**5** 设 $V$ 是有限维的，$U$ 是 $V$ 的子空间，$S\in\mathcal{L}(U,V)$. 证明：存在从 $V$ 到自身的一可逆映射 $T$ 使得对每个 $u\in U$ 有 $Tu=Su$，当且仅当 $S$ 是单射.

**6** 设 $W$ 是有限维的，且 $S,T\in\mathcal{L}(V,W)$. 证明：$\operatorname{null}S=\operatorname{null}T$，当且仅当存在一可逆的 $E\in\mathcal{L}(W)$ 使得 $S=ET$.

**7** 设 $V$ 是有限维的，且 $S,T\in\mathcal{L}(V,W)$. 证明：$\operatorname{range}S=\operatorname{range}T$，当且仅当存在一可逆的 $E\in\mathcal{L}(V)$ 使得 $S=TE$.

**8** 设 $V$ 和 $W$ 是有限维的，且 $S,T\in\mathcal{L}(V,W)$. 证明：存在可逆的 $E_1\in\mathcal{L}(V)$ 和 $E_2\in\mathcal{L}(W)$ 使得 $S=E_2TE_1$，当且仅当 $\dim\operatorname{null}S=\dim\operatorname{null}T$.

**9** 设 $V$ 是有限维的，且 $T:V\to W$ 是由 $V$ 映成 $W$ 的满的线性映射. 证明：存在 $V$ 的一子空间 $W$ 使得 $T|_U$ 是由 $U$ 映成 $W$ 的同构.

注 这里 $T|_U$ 意即限制在 $U$ 中的函数 $T$. 因此，$T|_U$ 是以 $U$ 为定义域的函数，其定义是对任一 $u\in U$ 都有 $T|_U(u)=Tu$.

**10** 设 $V$ 和 $W$ 是有限维的，且 $U$ 是 $V$ 的子空间. 令

$$\mathcal{E}=\{T\in\mathcal{L}(V,W):U\subseteq\operatorname{null}T\}.$$

(a) 证明 $\mathcal{E}$ 是 $\mathcal{L}(V,W)$ 的子空间.

(b) 求 $\dim\mathcal{E}$ 关于 $\dim V$、$\dim W$ 和 $\dim U$ 的表达式.

提示 定义 $\Phi:\mathcal{L}(V,W)\to\mathcal{L}(U,W)$ 为 $\Phi(T)=T|_U$. $\operatorname{null}\Phi$ 是什么？$\operatorname{range}\Phi$ 是什么？

**11** 设 $V$ 是有限维的，且 $S,T\in\mathcal{L}(V)$. 证明：

$$ST\text{ 可逆}\iff S\text{ 和 }T\text{ 可逆}.$$

**12** 设 $V$ 是有限维的，且 $S,T,U\in\mathcal{L}(V)$，且 $STU=I$. 证明：$T$ 可逆，且 $T^{-1}=US$.

**13** 证明：没有“$V$ 是有限维的”这一假设，第 12 题的结果可能不成立.

**14** 证明或给出一反例：如果 $V$ 是有限维向量空间且 $R,S,T\in\mathcal{L}(V)$ 使得 $RST$ 是满射，那么 $S$ 是单射.

**15** 设 $T\in\mathcal{L}(V)$，且 $v_1,\dots,v_m$ 是 $V$ 中一向量组，使得 $Tv_1,\dots,Tv_m$ 张成 $V$. 证明 $v_1,\dots,v_m$ 张成 $V$.

**16** 证明：从 $\mathbf{F}^{n,1}$ 到 $\mathbf{F}^{m,1}$ 的每个线性映射都能由一矩阵乘法给出. 换句话说，证明：如果 $T\in\mathcal{L}(\mathbf{F}^{n,1},\mathbf{F}^{m,1})$，那么存在一 $m\times n$ 矩阵 $A$ 使得 $Tx=Ax$ 对每个 $x\in\mathbf{F}^{n,1}$ 成立.

**17** 设 $V$ 是有限维的，且 $S\in\mathcal{L}(V)$. 定义 $\mathcal{A}\in\mathcal{L}(\mathcal{L}(V))$ 如下：

$$\mathcal{A}(T)=ST$$

对 $T\in\mathcal{L}(V)$ 成立.

(a) 证明 $\dim\operatorname{null}\mathcal{A}=(\dim V)(\dim\operatorname{null}S)$.

(b) 证明 $\dim\operatorname{range}\mathcal{A}=(\dim V)(\dim\operatorname{range}S)$.

**18** 证明 $V$ 和 $\mathcal{L}(\mathbf{F}, V)$ 是同构向量空间. [10]

**19** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明：$T$ 关于 $V$ 每一个基的矩阵都相同，当且仅当 $T$ 是恒等算子的标量倍.

**20** 设 $q \in \mathcal{P}(\mathbf{R})$，证明：存在一个多项式 $p \in \mathcal{P}(\mathbf{R})$ 使得

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

对所有 $x \in \mathbf{R}$ 成立.

**21** 设 $n$ 是一正整数，且 $A_{j,k} \in \mathbf{F}$（$j, k = 1, \ldots, n$）. 证明下列是等价的. 注意，在以下两点结论中，方程和未知数的个数相等.

(a) 平凡解 $x_1 = \cdots = x_n = 0$ 是下列齐次方程组的唯一解.

$$\begin{gathered}\sum_{k=1}^{n} A_{1,k} x_k = 0 \\ \vdots \\ \sum_{k=1}^{n} A_{n,k} x_k = 0.\end{gathered}$$

(b) 对任意 $c_1, \ldots, c_n \in \mathbf{F}$，下列方程组都有解.

$$\begin{gathered}\sum_{k=1}^{n} A_{1,k} x_k = c_1 \\ \vdots \\ \sum_{k=1}^{n} A_{n,k} x_k = c_n.\end{gathered}$$

**22** 设 $T \in \mathcal{L}(V)$ 且 $v_1, \ldots, v_n$ 是 $V$ 的基，证明：

$$\mathcal{M}\big(T, (v_1, \ldots, v_n)\big) \text{ 可逆} \iff T \text{ 可逆}.$$

**23** 设 $u_1, \ldots, u_n$ 和 $v_1, \ldots, v_n$ 是 $V$ 的基. 令 $T \in \mathcal{L}(V)$ 使得 $Tv_k = u_k$ 对任意 $k = 1, \ldots, n$ 成立，证明：

$$\mathcal{M}\big(T, (v_1, \ldots, v_n)\big) = \mathcal{M}\big(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n)\big).$$

**24** 设 $A$ 和 $B$ 是相同大小的方阵且 $AB = I$，证明 $BA = I$.

[10]想到个笑话：先证明 $V$ 是同构的，再证明 $\mathcal{L}(\mathbf{F}, V)$ 是同构的，证毕.

# 3E 向量空间的积和商

## 向量空间的积

通常，与多个向量空间打交道时，这些向量空间都应在相同的域上.

**3.87 定义：向量空间的积（product of vector spaces）**

设 $V_1, \ldots, V_m$ 都是 $\mathbf{F}$ 上的向量空间.

- **乘积** $V_1 \times \cdots \times V_m$ 定义为

$$V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) : v_1 \in V_1, \ldots, v_m \in V_m\}.$$

- $V_1 \times \cdots \times V_m$ 上的加法定义为

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, u_m + v_m).$$

- $V_1 \times \cdots \times V_m$ 上的标量乘法定义为

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m).$$

♣

**3.88 例：向量空间 $\mathcal{P}_5(\mathbf{R})$ 和 $\mathbf{R}^3$ 的乘积**

$\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ 的元素是长度为 2 的组，组中第一项是 $\mathcal{P}_5(\mathbf{R})$ 的元素，第二项是 $\mathbf{R}^3$ 的元素.

例如，$\big(5 - 6x + 4x^2, (3, 8, 7)\big)$ 和 $\big(x + 9x^5, (2, 2, 2)\big)$ 都是 $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ 的元素. 它们的和定义为

$$\begin{aligned}\big(5 - 6x + 4x^2, (3, 8, 7)\big) &+ \big(x + 9x^5, (2, 2, 2)\big) \\ &= \big(5 - 5x + 4x^2 + 9x^5, (5, 10, 9)\big).\end{aligned}$$

此外，$2\big(5 - 6x + 4x^2, (3, 8, 7)\big) = \big(10 - 12x + 8x^2, (6, 16, 14)\big)$.

下面结论应解释成：在按 3.87 定义的加法和标量乘法运算下，向量空间的积是向量空间.

**3.89 向量空间的积是向量空间**

设 $V_1, \ldots, V_m$ 都是 $\mathbf{F}$ 上的向量空间. 那么 $V_1 \times \cdots \times V_m$ 是 $\mathbf{F}$ 上的向量空间.

♡

上述结论的证明留给读者完成. 注意，$V_1 \times \cdots \times V_m$ 的加法恒等元是 $(0, \ldots, 0)$，其中第 $k$ 个坐标中的 0 是 $V_k$ 的加法恒等元. $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$ 的加法逆元是 $(-v_1, \ldots, -v_m)$.

**3.90 例：$\mathbf{R}^2 \times \mathbf{R}^3 \neq \mathbf{R}^5$，但 $\mathbf{R}^2 \times \mathbf{R}^3$ 与 $\mathbf{R}^5$ 是同构的**

向量空间 $\mathbf{R}^2 \times \mathbf{R}^3$ 的元素是这样的组：

$$\big((x_1, x_2), (x_3, x_4, x_5)\big),$$

其中 $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$. $\mathbf{R}^5$ 的元素是这样的组：

$$(x_1, x_2, x_3, x_4, x_5),$$

其中 $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$.

转下页

尽管 $\mathbf{R}^2 \times \mathbf{R}^3$ 的元素和 $\mathbf{R}^5$ 的元素看起来很像，它们却不是同类对象：$\mathbf{R}^2 \times \mathbf{R}^3$ 的元素是长度为 2 的组（其中第一项本身就是长度为 2 的组，而第二项是长度为 3 的组），但 $\mathbf{R}^5$ 的元素是长度为 5 的组．因此，$\mathbf{R}^2 \times \mathbf{R}^3$ 不等于 $\mathbf{R}^5$.

线性映射

$$\big((x_1, x_2), (x_3, x_4, x_5)\big) \mapsto (x_1, x_2, x_3, x_4, x_5)$$

是将向量空间 $\mathbf{R}^2 \times \mathbf{R}^3$ 映成向量空间 $\mathbf{R}^5$ 的同构．于是，这两个向量空间是同构的，尽管它们并不相等.

这个同构非常自然，我们把它看成改写即可．有些人通俗地说"$\mathbf{R}^2 \times \mathbf{R}^3$ 等于 $\mathbf{R}^5$"，尽管严格来讲这说法并不正确，但它体现了通过改写将两个向量空间等同起来看的思想.

下面例子形象地展示了我们将在 3.92 的证明中使用的思想.

**3.91 例：$\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 的一个基**

考虑 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 中这个长度为 5 的向量组：

$$\big(1, (0, 0)\big), \big(x, (0, 0)\big), \big(x^2, (0, 0)\big), \big(0, (1, 0)\big), \big(0, (0, 1)\big).$$

上述组是线性无关的，且张成 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$．所以它是 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 的一个基.

**3.92 向量空间之积的维数是各向量空间维数之和**

设 $V_1, \ldots, V_m$ 都是有限维向量空间．那么 $V_1 \times \cdots \times V_m$ 是有限维的，且

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

**证明** 为每个 $V_k$ 选定一个基. 对于每个 $V_k$ 的每个基向量，考虑 $V_1 \times \cdots \times V_m$ 中这样一个元素：它的第 $k$ 个坐标等于这个基向量，而其他坐标都等于 0. 所有这样的元素组成的向量组，是线性无关的且张成 $V_1 \times \cdots \times V_m$，从而该组是 $V_1 \times \cdots \times V_m$ 的一个基. 这个基的长度是 $\dim V_1 + \cdots + \dim V_m$，原命题得证. ■

在下面结果中，由 $V_1 + \cdots + V_m$ 的定义可得映射 $\Gamma$ 是满射．于是下述结论的最后一个词可以由"单射"替换成"可逆的".

**3.93 积与直和**

设 $V_1, \ldots, V_m$ 都是 $V$ 的子空间．由下式定义线性映射 $\Gamma : V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$：

$$\Gamma(v_1, \ldots, v_m) = v_1 + \cdots + v_m.$$

那么 $V_1 + \cdots + V_m$ 是直和，当且仅当 $\Gamma$ 是单射.

**证明** 由 3.15 可得，$\Gamma$ 是单射，当且仅当以 $v_1 + \cdots + v_m$（式中各向量 $v_k$ 属于 $V_k$）表达 0 的唯一方式是将其中的各 $v_k$ 都取 0．于是 1.45 表明，$\Gamma$ 是单射，当且仅当 $V_1 + \cdots + V_m$ 是直和，此即待证命题. ■

**3.94 向量空间的和是直和，当且仅当该和的维数等于各求和项维数之和**

设 $V$ 是有限维的，$V_1, \ldots, V_m$ 都是 $V$ 的子空间．那么 $V_1 + \cdots + V_m$ 是直和，当且仅当

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m.$$

**证明** 3.93 中的映射 Γ 是满射. 于是根据线性映射基本定理 ( 3.21 ), Γ 是单射，当且仅当

$$\dim(V_1 + \cdots + V_m) = \dim(V_1 \times \cdots \times V_m).$$

将 3.93 和 3.92 相结合，就证明了 $V_1 + \cdots + V_m$ 是直和，当且仅当

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m,$$

原命题得证. ■

在 $m = 2$ 这个特殊情形下，“$V_1 + V_2$ 是直和当且仅当 $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$”这个命题还可以通过将 1.46 与 2.43 相结合来证明.

## 商空间

我们从定义向量与子集之和开启迈向商空间的第一步.

**3.95 记号：$v + U$**

设 $v \in V$ 且 $U \subseteq V$. 那么 $v + U$ 是 $V$ 的一个由下式定义的子集：

$$v + U = \{v + u : u \in U\}.$$

♣

**3.96 例：一向量与 $\mathbf{R}^2$ 的一个一维子空间之和**

假设

$$U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}.$$

那么 $U$ 是 $\mathbf{R}^2$ 中过原点且斜率为 2 的直线. 从而

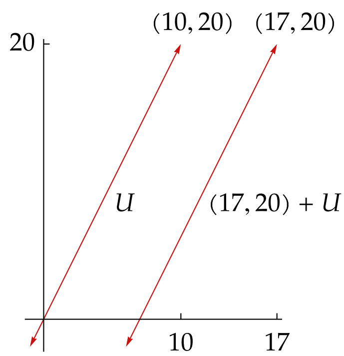
$$(17, 20) + U$$

是 $\mathbf{R}^2$ 中过点 $(17, 20)$ 且斜率为 2 的直线.

因为

$$(10, 20) \in U \quad 且 \quad (17, 20) \in (17, 20) + U,$$

我们可知 $(17, 20) + U$ 是由 $U$ 向右平移 7 个单位所得到的.




$(17, 20) + U$ 平行于子空间 $U$.

**3.97 定义：平移 ( translate )**

对于 $v \in V$ 和 $V$ 的一个子集 $U$，称集合 $v + U$ 是 $U$ 的一个**平移**.

♣

**3.98 例：平移**

- 如果 $U$ 是 $\mathbf{R}^2$ 中定义为 $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$ 的直线，那么 $\mathbf{R}^2$ 中所有斜率为 2 的直线都是 $U$ 的平移. $U$ 和其一个平移的图形如上面的例 3.96 所示.
- 更一般地说，如果 $U$ 是 $\mathbf{R}^2$ 中的直线，那么 $U$ 的所有平移构成的集合就是 $\mathbf{R}^2$ 中所有与 $U$ 平行的直线构成的集合.
- 如果 $U = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$，那么 $U$ 的平移是 $\mathbf{R}^3$ 中所有与 $xy$ 平面 ( 即 $U$ ) 平行的平面.

转下页

- 更一般地说，如果 $U$ 是 $\mathbf{R}^3$ 中的平面，那么 $U$ 的所有平移构成的集合就是 $\mathbf{R}^3$ 中所有与 $U$ 平行的平面构成的集合（实例见本节习题 7）.

**3.99 定义：商空间（quotient space）、$V/U$**

设 $U$ 是 $V$ 的子空间．那么**商空间** $V/U$ 是由 $U$ 的所有平移构成的集合．从而

$$V/U = \{v + U : v \in V\}.$$

**3.100 例：商空间**

- 如果 $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$，那么 $\mathbf{R}^2/U$ 是 $\mathbf{R}^2$ 中所有斜率为 2 的直线所构成的集合.
- 如果 $U$ 是 $\mathbf{R}^3$ 中过原点的直线，那么 $\mathbf{R}^3/U$ 是 $\mathbf{R}^3$ 中所有与 $U$ 平行的直线所构成的集合.
- 如果 $U$ 是 $\mathbf{R}^3$ 中过原点的平面，那么 $\mathbf{R}^3/U$ 是 $\mathbf{R}^3$ 中所有与 $U$ 平行的平面所构成的集合.

我们接下来的目标，是让 $V/U$ 成为向量空间．为此，我们需要下面这个结果.

**3.101 子空间的两个平移要么相等要么不相交**

设 $U$ 是 $V$ 的子空间且 $v, w \in V$．那么

$$v - w \in U \iff v + U = w + U \iff (v + U) \cap (w + U) \neq \varnothing.$$

**证明** 先设 $v - w \in U$．如果 $u \in U$，那么

$$v + u = w + \big((v - w) + u\big) \in w + U.$$

从而 $v + U \subseteq w + U$．类似地，$w + U \subseteq v + U$．于是 $v + U = w + U$，也就证明了 $v - w \in U$ 蕴涵 $v + U = w + U$.

等式 $v + U = w + U$ 蕴涵 $(v + U) \cap (w + U) \neq \varnothing$.

现在设 $(v + U) \cap (w + U) \neq \varnothing$．于是存在 $u_1, u_2 \in U$ 使得

$$v + u_1 = w + u_2.$$

于是 $v - w = u_2 - u_1$．所以 $v - w \in U$，这表明 $(v + U) \cap (w + U) \neq \varnothing$ 蕴涵 $v - w \in U$，证毕. ■

现在我们可以定义 $V/U$ 上的加法和标量乘法了.

**3.102 定义：$V/U$ 上的加法和标量乘法（addition and scalar multiplication on $V/U$）**

设 $U$ 是 $V$ 的子空间.那么 $V/U$ 上的**加法**和**标量乘法**分别由下面两式定义：对所有 $v, w \in V$ 和所有 $\lambda \in \mathbf{F}$，

$$(v + U) + (w + U) = (v + w) + U$$

$$\lambda(v + U) = (\lambda v) + U.$$

以上定义是合理的——对这一点的证明，是对接下来这条结果的证明中的一部分.

**3.103 商空间是向量空间**

假设 $U$ 是 $V$ 的子空间．那么带有定义如上的加法和标量乘法的 $V/U$ 就是向量空间.

**证明** 上述 $V/U$ 上加法和标量乘法的定义可能存在一个问题，就是 $U$ 的同一个平移存在不唯一的表达法. 更具体地说，为了证明上面给出的 $V/U$ 上的加法定义是合理的，假设有向量 $v_1, v_2, w_1, w_2 \in V$ 满足

$$v_1 + U = v_2 + U \quad 且 \quad w_1 + U = w_2 + U,$$

我们必须证明 $(v_1 + w_1) + U = (v_2 + w_2) + U$.

由 3.101，我们有

$$v_1 - v_2 \in U \quad 且 \quad w_1 - w_2 \in U.$$

因为 $U$ 是 $V$ 的子空间，从而对加法封闭，所以此式表明 $(v_1 - v_2) + (w_1 - w_2) \in U$. 于是 $(v_1 + w_1) - (v_2 + w_2) \in U$. 再次利用 3.101 可见

$$(v_1 + w_1) + U = (v_2 + w_2) + U,$$

即我们欲证的结论成立. 于是 $V/U$ 上加法的定义是合理的.

类似地，假设 $\lambda \in \mathbf{F}$. 仍假设 $v_1 + U = v_2 + U$. 因为 $U$ 是 $V$ 的子空间，从而对标量乘法封闭，所以我们有 $\lambda(v_1 - v_2) \in U$. 于是 $\lambda v_1 - \lambda v_2 \in U$. 所以 3.101 表明

$$(\lambda v_1) + U = (\lambda v_2) + U.$$

于是 $V/U$ 上加法和标量乘法的定义都是合理的.

既然 $V/U$ 上的加法和标量乘法运算已定义清楚，那么验证这些运算使 $V/U$ 成为向量空间就很简单了，请读者自行完成. 注意，$V/U$ 的加法恒等元是 $0 + U$（也就等于 $U$），$v + U$ 的加法逆元是 $(-v) + U$. ■

接下来的概念将引出 $V/U$ 的维数计算方法.

**3.104 定义：商映射（quotient map）、$\pi$**

设 $U$ 是 $V$ 的子空间. **商映射** $\pi : V \to V/U$ 是由下式定义的线性映射：对每个 $v \in V$,

$$\pi(v) = v + U.$$

♣

读者应自行验证 $\pi$ 的确是一个线性映射. 尽管 $\pi$ 既依赖于 $U$ 又依赖于 $V$，记号中却省略了这些向量空间，因为根据上下文就能明确它们是什么.

**3.105 商空间的维数**

设 $V$ 是有限维的，$U$ 是 $V$ 的子空间. 那么

$$\dim V/U = \dim V - \dim U.$$

♡

**证明** 令 $\pi$ 表示从 $V$ 到 $V/U$ 的线性映射. 如果 $v \in V$，那么 $v + U = 0 + U$ 当且仅当 $v \in U$（由 3.101），这表明 $\operatorname{null} \pi = U$. $\pi$ 的定义表明 $\operatorname{range} \pi = V/U$. 这样一来，由线性映射基本定理（3.21）可知 $\dim V = \dim U + \dim V/U$，这正是我们想要的结果. ■

$V$ 上的每个线性映射 $T$ 都能在 $V/\operatorname{null} T$ 上诱导出一个线性映射 $\widetilde{T}$. 我们现在给出其定义.

**3.106 记号：$\widetilde{T}$**

设 $T \in \mathcal{L}(V,W)$. $\widetilde{T} : V/(\operatorname{null} T) \to W$ 由下式定义：

$$\widetilde{T}(v + \operatorname{null} T) = Tv.$$

♣

为了说明上述 $\widetilde{T}$ 的定义是合理的，假设 $u, v \in V$ 满足 $u + \operatorname{null} T = v + \operatorname{null} T$. 由 3.101，我们有 $u - v \in \operatorname{null} T$. 于是 $T(u - v) = 0$. 所以 $Tu = Tv$. 于是 $\widetilde{T}$ 的定义确实是合理的. 验证 $\widetilde{T}$ 是一个从 $V/\operatorname{null} T$ 到 $W$ 的线性映射的过程很常规，留给读者完成.

下面的结果说明，我们可以将 $\widetilde{T}$ 看成 $T$ 的修改版——换了定义空间，从而变为单射.

**3.107 $\widetilde{T}$ 的零空间和值域**

设 $T \in \mathcal{L}(V,W)$. 那么

(a) $\widetilde{T} \circ \pi = T$，其中 $\pi$ 是将 $V$ 映成 $V/(\operatorname{null} T)$ 的商映射；

(b) $\widetilde{T}$ 是单射；

(c) $\operatorname{range} \widetilde{T} = \operatorname{range} T$；

(d) $V/(\operatorname{null} T)$ 和 $\operatorname{range} T$ 是同构的向量空间.

♡

**证明**

(a) 如果 $v \in V$，那么 $(\widetilde{T} \circ \pi)(v) = \widetilde{T}(\pi(v)) = \widetilde{T}(v + \operatorname{null} T) = Tv$，命题得证；

(b) 设 $v \in V$ 且 $\widetilde{T}(v+\operatorname{null} T) = 0$. 那么 $Tv = 0$. 于是 $v \in \operatorname{null} T$. 所以 3.101 表明 $v+\operatorname{null} T = 0+\operatorname{null} T$. 由此推出 $\operatorname{null} \widetilde{T} = \{0 + \operatorname{null} T\}$. 所以 $\widetilde{T}$ 是单射，命题得证；

(c) $\widetilde{T}$ 的定义即表明 $\operatorname{range} \widetilde{T} = \operatorname{range} T$；

(d) 现在由 (b) 和 (c) 可得，如果我们将 $\widetilde{T}$ 视作映射到 $\operatorname{range} T$，那么 $\widetilde{T}$ 就是将 $V/(\operatorname{null} T)$ 映成 $\operatorname{range} T$ 的同构. ■

## 习题 3E

**1** 设 $T$ 是从 $V$ 到 $W$ 的函数. $T$ 的**图**（**graph**）是 $V \times W$ 的子集，定义为

$$T \text{ 的图 } = \{(v, Tv) \in V \times W : v \in V\}.$$

证明：$T$ 是线性映射，当且仅当 $T$ 的图是 $V \times W$ 的子空间.

注　正式而言，从 $V$ 到 $W$ 的函数 $T$，是 $V \times W$ 的一个子集 $T$，使得对每个 $v \in V$ 存在恰好一个元素 $(v, w) \in T$. 换句话说，函数正式而言就是上面所谓的图. 我们通常不把函数看成这种正式形式，然而，如果我们确实要正式些的话，那么这道习题可以重写如下：证明从 $V$ 到 $W$ 的函数 $T$ 是线性映射当且仅当 $T$ 是 $V \times W$ 的子空间.

**2** 设 $V_1, \dots, V_m$ 是向量空间，使得 $V_1 \times \cdots \times V_m$ 是有限维的. 证明：$V_k$（$k = 1, \dots, m$）是有限维的.

**3** 设 $V_1, \dots, V_m$ 是向量空间. 证明 $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ 和 $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ 是同构向量空间.

注　上面这道题和接下来两道题中，并未假定各向量空间是有限维的.

**4** 设 $W_1,\dots,W_m$ 是向量空间．证明 $\mathcal{L}(V,W_1\times\cdots\times W_m)$ 和 $\mathcal{L}(V,W_1)\times\cdots\times\mathcal{L}(V,W_m)$ 是同构向量空间．

**5** 对于正整数 $m$，定义 $V^m$ 为：

$$V^m=\underbrace{V\times\cdots\times V}_{m\text{个}}.$$

证明：$V^m$ 和 $\mathcal{L}(\mathbf{F}^m,V)$ 是同构向量空间．

**6** 设 $v,x$ 是 $V$ 中的向量，$U,W$ 是 $V$ 的子空间，使得 $v+U=x+W$．证明 $U=W$．

**7** 令 $U=\{(x,y,z)\in\mathbf{R}^3:2x+3y+5z=0\}$．设 $A\subseteq\mathbf{R}^3$．证明：$A$ 是 $U$ 的平移，当且仅当存在 $c\in\mathbf{R}$ 使得

$$A=\{(x,y,z)\in\mathbf{R}^3:2x+3y+5z=c\}.$$

**8** (a) 设 $T\in\mathcal{L}(V,W)$，$c\in W$．证明：$\{x\in V:Tx=c\}$ 或是空集或是 $\operatorname{null}T$ 的平移．

(b) 解释为什么线性方程组【如式 (3.27)】的解集或是空集或是 $\mathbf{F}^n$ 的某个子空间的平移．

**9** 证明：$V$ 的一非空子集 $A$ 是 $V$ 的某个子空间的平移，当且仅当 $\lambda v+(1-\lambda)w\in A$ 对所有 $v,w\in A$ 和所有 $\lambda\in\mathbf{F}$ 成立．

**10** 设 $A_1=v+U_1$ 且 $A_2=w+U_2$，其中 $v,w\in V$，$U_1,U_2$ 是 $V$ 的子空间．证明交集 $A_1\cap A_2$ 是 $V$ 的某个子空间的平移或者空集．

**11** 设 $U=\{(x_1,x_2,\dots)\in\mathbf{F}^\infty:x_k\neq 0\text{ 仅对有限多个 }k\text{ 成立}\}$．

(a) 证明 $U$ 是 $\mathbf{F}^\infty$ 的子空间．

(b) 证明 $\mathbf{F}^\infty/U$ 是无限维的．

**12** 设 $v_1,\dots,v_m\in V$．令

$$A=\{\lambda_1v_1+\cdots+\lambda_mv_m:\lambda_1,\dots,\lambda_m\in\mathbf{F}\text{ 且 }\lambda_1+\cdots+\lambda_m=1\}.$$

(a) 证明：$A$ 是 $V$ 的某个子空间的平移．

(b) 证明：如果 $B$ 是 $V$ 的某个子空间的平移且 $\{v_1,\dots,v_m\}\subseteq B$，那么 $A\subseteq B$．

(c) 证明：$A$ 是 $V$ 的某个子空间的平移，且该子空间的维数小于 $m$．

**13** 设 $U$ 是 $V$ 的子空间，使得 $V/U$ 是有限维的．证明 $V$ 和 $U\times(V/U)$ 同构．

**14** 设 $U$ 和 $W$ 是 $V$ 的子空间，且 $V=U\oplus W$．设 $w_1,\dots,w_m$ 是 $W$ 的基．证明 $w_1+U,\dots,w_m+U$ 是 $V/U$ 的基．

**15** 设 $U$ 是 $V$ 的子空间，$v_1+U,\dots,v_m+U$ 是 $V/U$ 的基，$u_1,\dots,u_n$ 是 $U$ 的基．证明 $v_1,\dots,v_m,u_1,\dots,u_n$ 是 $V$ 的基．

**16** 设 $\varphi\in\mathcal{L}(V,\mathbf{F})$，$\varphi\neq 0$．证明 $\dim V/(\operatorname{null}\varphi)=1$．

**17** 设 $U$ 是 $V$ 的子空间，使得 $\dim V/U=1$．证明：存在 $\varphi\in\mathcal{L}(V,\mathbf{F})$ 使得 $\operatorname{null}\varphi=U$．

**18** 设 $U$ 是 $V$ 的子空间，使得 $V/U$ 是有限维的．

(a) 证明：如果 $W$ 是 $V$ 的有限维子空间且 $V=U+W$，那么 $\dim W\geq\dim V/U$．

(b) 证明：存在 $V$ 的一有限维子空间 $W$，使得 $\dim W=\dim V/U$ 且 $V=U\oplus W$．

**19** 设 $T\in\mathcal{L}(V,W)$ 且 $U$ 是 $V$ 的子空间．令 $\pi$ 表示由 $V$ 映成 $V/U$ 的商映射．证明：存在 $S\in\mathcal{L}(V/U,W)$ 使得 $T=S\circ\pi$，当且仅当 $U\subseteq\operatorname{null}T$．

# 3F 对偶

## 对偶空间和对偶映射

映射到标量域 $\mathbf{F}$ 中的线性映射在线性代数中扮演着特殊的角色，于是，我们赋予它们特殊的名称.

**3.108 定义：线性泛函（linear functional）**

$V$ 上的**线性泛函**是从 $V$ 到 $\mathbf{F}$ 的线性映射．换言之，线性泛函是 $\mathcal{L}(V,\mathbf{F})$ 的元素.

**3.109 例：线性泛函**

- 定义 $\varphi:\mathbf{R}^3\to\mathbf{R}$ 为 $\varphi(x,y,z)=4x-5y+2z$．那么 $\varphi$ 是 $\mathbf{R}^3$ 上的线性泛函.
- 固定 $(c_1,\ldots,c_n)\in\mathbf{F}^n$．定义 $\varphi:\mathbf{F}^n\to\mathbf{F}$ 为 $\varphi(x_1,\ldots,x_n)=c_1x_1+\cdots+c_nx_n$．那么 $\varphi$ 是 $\mathbf{F}^n$ 上的线性泛函.
- 定义 $\varphi:\mathcal{P}(\mathbf{R})\to\mathbf{R}$ 如下：
$$\varphi(p)=3p''(5)+7p(4).$$
那么 $\varphi$ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函.
- 定义 $\varphi:\mathcal{P}(\mathbf{R})\to\mathbf{R}$ 如下：对每个 $p\in\mathcal{P}(\mathbf{R})$，
$$\varphi(p)=\int_0^1 p.$$
那么 $\varphi$ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函.

向量空间 $\mathcal{L}(V,\mathbf{F})$ 同样有特殊的名称和特殊的记号.

**3.110 定义：对偶空间（dual space）、$V'$**

$V$ 的**对偶空间**记作 $V'$，是 $V$ 上全体线性泛函所构成的向量空间．换言之，$V'=\mathcal{L}(V,\mathbf{F})$.

**3.111 $\dim V'=\dim V$**

假设 $V$ 是有限维的．那么 $V'$ 也是有限维的，且
$$\dim V'=\dim V.$$

**证明** 由 3.72 我们有
$$\dim V'=\dim\mathcal{L}(V,\mathbf{F})=(\dim V)(\dim\mathbf{F})=\dim V,$$
命题得证. ∎

在下面定义中，线性映射引理（3.4）表明每个 $\varphi_j$ 都具有明确的定义.

**3.112 定义：对偶基（dual basis）**

如果 $v_1,\ldots,v_n$ 是 $V$ 的基，那么 $v_1,\ldots,v_n$ 的**对偶基**是 $V'$ 中的元素 $\varphi_1,\ldots,\varphi_n$ 所构成的组，其中各 $\varphi_j$ 是 $V$ 上满足下式的线性泛函：

$$\varphi_j(v_k)=\begin{cases}1 & 若\ k=j,\\ 0 & 若\ k\neq j.\end{cases}$$

**3.113 例：$\mathbf{F}^n$ 的标准基的对偶基**

设 $n$ 是正整数．对 $1\leq j\leq n$，将 $\varphi_j$ 定义为将 $\mathbf{F}^n$ 中的向量对应至它的第 $j$ 个坐标的 $\mathbf{F}^n$ 上的线性泛函．于是对于各 $(x_1,\ldots,x_n)\in\mathbf{F}^n$，

$$\varphi_j(x_1,\ldots,x_n)=x_j.$$

令 $e_1,\ldots,e_n$ 是 $\mathbf{F}^n$ 的标准基．那么

$$\varphi_j(e_k)=\begin{cases}1 & 若\ k=j,\\ 0 & 若\ k\neq j.\end{cases}$$

于是 $\varphi_1,\ldots,\varphi_n$ 就是 $\mathbf{F}^n$ 的标准基 $e_1,\ldots,e_n$ 的对偶基．

下面结论说明了，对于 $V$ 的基而言，组成其对偶基的线性泛函，给出了用基向量的线性组合来表示 $V$ 中向量时所取的系数．

**3.114 对偶基给出了线性组合的系数**

假设 $v_1,\ldots,v_n$ 是 $V$ 的基，且 $\varphi_1,\ldots,\varphi_n$ 是其对偶基．那么对每个 $v\in V$，有

$$v=\varphi_1(v)v_1+\cdots+\varphi_n(v)v_n.$$

**证明** 设 $v\in V$．那么存在 $c_1,\ldots,c_n\in\mathbf{F}$ 使得

$$v=c_1v_1+\cdots+c_nv_n. \tag{3.115}$$

如果 $j\in\{1,\ldots,n\}$，那么在等式两端同时作用 $\varphi_j$ 可得

$$\varphi_j(v)=c_j.$$

将由上式给出的 $c_1,\ldots,c_n$ 表达式代入式 (3.115) 即可知道 $v=\varphi_1(v)v_1+\cdots+\varphi_n(v)v_n$． ■

接下来的结论表明对偶基的确是对偶空间的基．于是使用术语“对偶基”就是恰当的．

**3.116 对偶基是对偶空间的基**

假设 $V$ 是有限维的．那么 $V$ 的基的对偶基是 $V'$ 的基．

**证明** 设 $v_1,\ldots,v_n$ 是 $V$ 的基．令 $\varphi_1,\ldots,\varphi_n$ 表示其对偶基．

为了证明 $\varphi_1,\ldots,\varphi_n$ 是由 $V'$ 中元素构成的线性无关组，设 $a_1,\ldots,a_n\in\mathbf{F}$ 满足

$$a_1\varphi_1+\cdots+a_n\varphi_n=0. \tag{3.117}$$

而对于各 $k\in 1,\ldots,n$，有

$$(a_1\varphi_1+\cdots+a_n\varphi_n)(v_k)=a_k.$$

于是式 (3.117) 表明 $a_1 = \cdots = a_n = 0$. 所以 $\varphi_1, \ldots, \varphi_n$ 是线性无关的.

因为 $\varphi_1, \ldots, \varphi_n$ 是 $V'$ 中的线性无关组且其长度等于 $\dim V'$（由 3.111），我们可得出结论 $\varphi_1, \ldots, \varphi_n$ 是 $V'$ 的一个基（见 2.38）. ■

在接下来的定义中，注意如果 $T$ 是从 $V$ 到 $W$ 的线性映射，那么 $T'$ 是从 $W'$ 到 $V'$ 的线性映射.

**3.118 定义：对偶映射（dual map）、$T'$**

设 $T \in \mathcal{L}(V, W)$. $T$ 的**对偶映射**是由下式定义的线性映射 $T' \in \mathcal{L}(W', V')$：对每个 $\varphi \in W'$，

$$T'(\varphi) = \varphi \circ T.$$

♣

如果 $T \in \mathcal{L}(V, W)$ 且 $\varphi \in W'$，那么按上述定义，$T'(\varphi)$ 就是 $\varphi$ 和 $T$ 这两个线性映射的复合. 于是 $T'(\varphi)$ 的确是从 $V$ 到 $\mathbf{F}$ 的线性映射；换言之，$T'(\varphi) \in V'$.

有些书籍中用于表示对偶的记号是 $V^*$ 和 $T^*$ 而不是 $V'$ 和 $T'$. 但我们将 $T^*$ 这个记号预留给伴随——到了第 7 章，我们学习内积空间上的线性映射时，将引入伴随这个概念.

下面两点结论表明 $T'$ 是从 $W'$ 到 $V'$ 的线性映射：

- 若 $\varphi, \psi \in W'$，那么
$$T'(\varphi + \psi) = (\varphi + \psi) \circ T = \varphi \circ T + \psi \circ T = T'(\varphi) + T'(\psi).$$
- 若 $\lambda \in \mathbf{F}$ 且 $\varphi \in W'$，那么
$$T'(\lambda\varphi) = (\lambda\varphi) \circ T = \lambda(\varphi \circ T) = \lambda T'(\varphi).$$

下例中撇号（$'$）有两种不相关的含义：在 $D'$ 中用于表示线性映射 $D$ 的对偶，在 $p'$ 中则用于表示多项式 $p$ 的导数.

---

**3.119 例：微分线性映射的对偶映射**

定义 $D : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ 为 $Dp = p'$.

- 设 $\varphi$ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函，定义为 $\varphi(p) = p(3)$. 那么 $D'(\varphi)$ 是由下式定义的 $\mathcal{P}(\mathbf{R})$ 上的线性泛函：
$$\big(D'(\varphi)\big)(p) = (\varphi \circ D)(p) = \varphi(Dp) = \varphi(p') = p'(3).$$
于是 $D'(\varphi)$ 是 $\mathcal{P}(\mathbf{R})$ 上将 $p$ 对应到 $p'(3)$ 的线性泛函.
- 设 $\varphi$ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函，定义为 $\varphi(p) = \int_0^1 p$. 那么 $D'(\varphi)$ 是由下式定义的 $\mathcal{P}(\mathbf{R})$ 上的线性泛函：
$$\begin{aligned}\big(D'(\varphi)\big)(p) &= (\varphi \circ D)(p)\\ &= \varphi(Dp)\\ &= \varphi(p')\\ &= \int_0^1 p'\\ &= p(1) - p(0).\end{aligned}$$
于是 $D'(\varphi)$ 是 $\mathcal{P}(\mathbf{R})$ 上将 $p$ 对应到 $p(1) - p(0)$ 的线性泛函.

---

在下面结果中，(a) 和 (b) 表明将 $T$ 对应到 $T'$ 的函数是一个从 $\mathcal{L}(V,W)$ 到 $\mathcal{L}(W',V')$ 的线性映射.

在下述 (c) 中，注意从等式左侧的 $ST$ 到右侧的 $T'S'$，$S$ 和 $T$ 的顺序是相反的.

**3.120 对偶映射的代数性质**

设 $T \in \mathcal{L}(V,W)$. 那么

(a) 对所有 $S \in \mathcal{L}(V,W)$，均有 $(S+T)' = S' + T'$；

(b) 对所有 $\lambda \in \mathbf{F}$，均有 $(\lambda T)' = \lambda T'$；

(c) 对所有 $S \in \mathcal{L}(W,U)$，均有 $(ST)' = T'S'$.

**证明** (a) 和 (b) 的证明留给读者完成.

为证明 (c)，设 $\varphi \in U'$. 那么

$$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'\big(S'(\varphi)\big) = (T'S')(\varphi),$$

其中第 1、3、4 个等号成立是因为对偶映射的定义，第 2 个等号成立是因为函数的复合满足结合律，而最后一个等号是由复合的定义得来.

上式表明对于所有 $\varphi \in U'$，均有 $(ST)'(\varphi) = (T'S')(\varphi)$. 于是 $(ST)' = T'S'$. ■

## 线性映射的对偶的零空间和值域

本小节中，我们的目标是用 $\operatorname{range} T$ 和 $\operatorname{null} T$ 来刻画 $\operatorname{null} T'$ 和 $\operatorname{range} T'$. 为达成这个目标，我们需要给出下面这个定义.

**3.121 定义：零化子（annihilator）、$U^0$**

对 $U \subseteq V$，$U$ 的**零化子**，记作 $U^0$，定义为

$$U^0 = \{\varphi \in V' : 对所有\ u \in U, \varphi(u) = 0\}.$$

---

**3.122 例：一个零化子的元素**

设 $U$ 是 $\mathcal{P}(\mathbf{R})$ 的子空间，其由 $x^2$ 的多项式倍[11]所构成. 如果 $\varphi$ 是 $\mathcal{P}(\mathbf{R})$ 上定义为 $\varphi(p) = p'(0)$ 的线性泛函，那么 $\varphi \in U^0$.

---

对 $U \subseteq V$，零化子 $U^0$ 是对偶空间 $V'$ 的子集. 于是 $U^0$ 依赖于包含 $U$ 的那个向量空间 $V$，所以形如 $U_V^0$ 的记号会更准确些. 然而，根据上下文我们总能明确包含 $U$ 的那个空间是什么，所以我们常用 $U^0$ 这个更简单的记号.

---

**3.123 例：$\mathbf{R}^5$ 的一个二维子空间的零化子**

令 $e_1, e_2, e_3, e_4, e_5$ 表示 $\mathbf{R}^5$ 的标准基，并令 $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5 \in (\mathbf{R}^5)'$ 表示 $e_1, e_2, e_3, e_4, e_5$ 的对偶基. 设

$$U = \operatorname{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbf{R}^5 : x_1, x_2 \in \mathbf{R}\}.$$

转下页

[11] $x^2$ 的多项式倍，就是指形如 $x^2 p(x)$【其中 $p(x)$ 也为多项式】的多项式.

我们想要证明 $U^0=\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)$.

回忆一下（参见例 3.113），$\varphi_j$ 是 $\mathbf{R}^5$ 上将向量对应到其第 $j$ 个坐标的线性泛函：$\varphi_j(x_1,x_2,x_3,x_4,x_5)=x_j$.

先设 $\varphi\in\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)$. 那么存在 $c_3,c_4,c_5\in\mathbf{R}$ 使得 $\varphi=c_3\varphi_3+c_4\varphi_4+c_5\varphi_5$ 成立. 如果 $(x_1,x_2,0,0,0)\in U$，那么

$$\varphi(x_1,x_2,0,0,0)=(c_3\varphi_3+c_4\varphi_4+c_5\varphi_5)(x_1,x_2,0,0,0)=0.$$

于是 $\varphi\in U^0$. 由此我们证明了 $\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)\subseteq U^0$.

为了证明这个包含关系反过来也成立，设 $\varphi\in U^0$. 因为对偶基是 $(\mathbf{R}^5)'$ 的基，所以存在 $c_1,c_2,c_3,c_4,c_5\in\mathbf{R}$ 使得 $\varphi=c_1\varphi_1+c_2\varphi_2+c_3\varphi_3+c_4\varphi_4+c_5\varphi_5$ 成立. 因为 $e_1\in U$ 且 $\varphi\in U^0$，所以我们有

$$0=\varphi(e_1)=(c_1\varphi_1+c_2\varphi_2+c_3\varphi_3+c_4\varphi_4+c_5\varphi_5)(e_1)=c_1.$$

类似地，$e_2\in U$ 从而 $c_2=0$. 所以 $\varphi=c_3\varphi_3+c_4\varphi_4+c_5\varphi_5$. 于是 $\varphi\in\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)$，说明 $U^0\subseteq\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)$.

于是，$U^0=\operatorname{span}(\varphi_3,\varphi_4,\varphi_5)$.

**3.124 零化子是子空间**

设 $U\subseteq V$. 那么 $U^0$ 是 $V'$ 的子空间.

**证明** 注意到 $0\in U^0$（此处 $0$ 是 $V$ 上的零线性泛函），因为将零线性泛函作用于 $U$ 中每个向量都会得到 $0\in\mathbf{F}$.

设 $\varphi,\psi\in U^0$. 于是 $\varphi,\psi\in V'$ 且对每个 $u\in U$ 都有 $\varphi(u)=\psi(u)=0$. 如果 $u\in U$，那么

$$(\varphi+\psi)(u)=\varphi(u)+\psi(u)=0+0=0.$$

于是 $\varphi+\psi\in U^0$.

类似可证 $U^0$ 对于标量乘法封闭. 于是 1.34 表明 $U^0$ 是 $V'$ 的子空间. ■

下面结果说明 $\dim U^0$ 是 $\dim V$ 和 $\dim U$ 之差. 例如，若 $U$ 是 $\mathbf{R}^5$ 的二维子空间，那么 $U^0$ 就是 $(\mathbf{R}^5)'$ 的三维子空间——见例 3.123.

下面的结论可以遵循例 3.123 的思考方式来证明：选取 $U$ 的基 $u_1,\ldots,u_m$，将其扩充成 $V$ 的基 $u_1,\ldots,u_m,\ldots,u_n$，令 $\varphi_1,\ldots,\varphi_m,\ldots,\varphi_n$ 为 $V'$ 的对偶基，然后证明 $\varphi_{m+1},\ldots,\varphi_n$ 是 $U^0$ 的基，即可得到欲证的结论. 即便下面给出了一个更取巧的证明，你还是应该按上面概述的思路自行构造一个证明.

**3.125 零化子的维数**

设 $V$ 是有限维的且 $U$ 是 $V$ 的子空间. 那么

$$\dim U^0=\dim V-\dim U.$$

**证明** 令 $i \in \mathcal{L}(U,V)$ 是包含映射（inclusion map）[12]，定义为 $i(u) = u$（对于任一 $u \in U$ 均成立）. 于是 $i'$ 是从 $V'$ 到 $U'$ 的线性映射. 将线性映射基本定理（3.21）应用于 $i'$ 可得

$$\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'.$$

而 $\operatorname{null} i' = U^0$（思考定义即可想出此式）且 $\dim V' = \dim V$（由 3.111），所以我们可以将上述等式改写成

$$\dim \operatorname{range} i' + \dim U^0 = \dim V. \tag{3.126}$$

如果 $\varphi \in U'$，那么 $\varphi$ 可以被扩充为 $V$ 上的一个线性泛函 $\psi$（实例见于 3A 节习题 13）. $i'$ 的定义表明 $i'(\psi) = \varphi$. 由此 $\varphi \in \operatorname{range} i'$，这表明 $\operatorname{range} i' = U'$，所以

$$\dim \operatorname{range} i' = \dim U' = \dim U,$$

那么式 (3.126) 就可改写成 $\dim U + \dim U^0 = \dim V$，原命题得证. ■

下面结果可作一个有力工具，用以证明某个子空间是最大子空间【见 (a)】或是最小子空间【见 (b)】.

**3.127 零化子等于 $\{0\}$ 或整个空间的条件**

设 $V$ 是有限维的，且 $U$ 是 $V$ 的子空间. 那么

(a) $U^0 = \{0\} \iff U = V$;

(b) $U^0 = V' \iff U = \{0\}$.

**证明** 为证明 (a)，我们有

$$\begin{aligned} U^0 = \{0\} &\iff \dim U^0 = 0 \\ &\iff \dim U = \dim V \\ &\iff U = V, \end{aligned}$$

其中，第二个等价关系来自 3.125 而第三个等价关系来自 2.39.

类似地，为证明 (b)，我们有

$$\begin{aligned} U^0 = V' &\iff \dim U^0 = \dim V' \\ &\iff \dim U^0 = \dim V \\ &\iff \dim U = 0 \\ &\iff U = \{0\}, \end{aligned}$$

其中，第一个等价关系中的一个方向源于 2.39，第二个等价关系来自 3.111，而第三个等价关系来自 3.125. ■

下面结果中 (a) 的证明无需假设 $V$ 和 $W$ 是有限维的.

**3.128 $T'$ 的零空间**

设 $V$ 和 $W$ 是有限维的且 $T \in \mathcal{L}(V,W)$. 那么

(a) $\operatorname{null} T' = (\operatorname{range} T)^0$;

(b) $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

[12] 对比一下恒等映射 $I : U \to U, I(u) = u$，不难看出两者的联系与区别：都将定义空间中的元素映到同一元素，但是目标空间不同.

**证明**

(a) 先设 $\varphi \in \operatorname{null} T'$. 于是 $0 = T'(\varphi) = \varphi \circ T$. 因此

$$\text{对每个 } v \in V, \text{ 均有 } 0 = (\varphi \circ T)(v) = \varphi(Tv).$$

于是 $\varphi \in (\operatorname{range} T)^0$. 这意味着 $\operatorname{null} T' \subseteq (\operatorname{range} T)^0$.

为了证明这个包含关系反过来也成立，现设 $\varphi \in (\operatorname{range} T)^0$. 于是对每个 $v \in V$ 都有 $\varphi(Tv) = 0$. 所以 $0 = \varphi \circ T = T'(\varphi)$. 换言之，$\varphi \in \operatorname{null} T'$，这表明 $(\operatorname{range} T)^0 \subseteq \operatorname{null} T'$, 这样就完成了 (a) 的证明.

(b) 我们有

$$\begin{aligned}\dim \operatorname{null} T' &= \dim (\operatorname{range} T)^0 \\ &= \dim W - \dim \operatorname{range} T \\ &= \dim W - (\dim V - \dim \operatorname{null} T) \\ &= \dim \operatorname{null} T + \dim W - \dim V,\end{aligned}$$

其中第一个等号源于 (a)，第二个等号源于 3.125，第三个等号则源于线性映射基本定理 ( 3.21 ). ■

下面结果在某些情况下有用处，因为有时验证 $T'$ 是单射比直接验证 $T$ 是满射更容易.

**3.129 $T$ 是满射等价于 $T'$ 是单射**

设 $V$ 和 $W$ 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

$$T \text{ 是满射} \iff T' \text{ 是单射}.$$

**证明** 我们有

$$\begin{aligned}T \in \mathcal{L}(V, W) \text{是满射} &\iff \operatorname{range} T = W \\ &\iff (\operatorname{range} T)^0 = \{0\} \\ &\iff \operatorname{null} T' = \{0\} \\ &\iff T' \text{是单射},\end{aligned}$$

其中第二个等价关系来源于 3.127 (a) 而第三个等价关系源于 3.128 (a). ■

**3.130 $T'$ 的值域**

设 $V$ 和 $W$ 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

(a) $\dim \operatorname{range} T' = \dim \operatorname{range} T$;

(b) $\operatorname{range} T' = (\operatorname{null} T)^0$.

**证明**

(a) 我们有

$$\begin{aligned}\dim \operatorname{range} T' &= \dim W' - \dim \operatorname{null} T' \\ &= \dim W - \dim (\operatorname{range} T)^0 \\ &= \dim \operatorname{range} T,\end{aligned}$$

其中第一个等号源于 3.21，第二个等号源于 3.111 和 3.128 (a)，第三个等号则源于 3.125.

(b) 先设 $\varphi \in \operatorname{range} T'$. 于是存在 $\psi \in W'$ 使得 $\varphi = T'(\psi)$ 成立. 如果 $v \in \operatorname{null} T$, 那么

$$\varphi(v) = \big(T'(\psi)\big)v = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0.$$

因此 $\varphi \in (\operatorname{null} T)^0$. 这意味着 $\operatorname{range} T' \subseteq (\operatorname{null} T)^0$.

我们将通过说明 $\operatorname{range} T'$ 和 $(\operatorname{null} T)^0$ 的维数相同来完成证明. 为此, 注意到

$$\begin{aligned}\dim \operatorname{range} T' &= \dim \operatorname{range} T\\ &= \dim V - \dim \operatorname{null} T\\ &= \dim (\operatorname{null} T)^0,\end{aligned}$$

其中第一个等号来源于 (a), 第二个等号来源于 3.21, 第三个等号则源于 3.125. ■

下个结果应和 3.129 对照起来看.

**3.131 $T$ 是单射等价于 $T'$ 是满射**

设 $V$ 和 $W$ 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

$$T \text{ 是单射} \iff T' \text{ 是满射}.$$

**证明** 我们有

$$\begin{aligned} T \text{ 是单射} &\iff \operatorname{null} T = \{0\}\\ &\iff (\operatorname{null} T)^0 = V'\\ &\iff \operatorname{range} T' = V',\end{aligned}$$

其中第二个等价关系来源于 3.127 (b) 而第三个等价关系源于 3.130 (b). ■

## 线性映射的对偶的矩阵

接下来的结论基于以下假设: 假设有 $V$ 的基 $v_1, \ldots, v_n$ 及相应的 $V'$ 的对偶基 $\varphi_1, \ldots, \varphi_n$, 也有 $W$ 的基 $w_1, \ldots, w_m$ 及相应的 $W'$ 的对偶基 $\psi_1, \ldots, \psi_m$. 于是可按上述 $V$ 和 $W$ 的基计算出 $\mathcal{M}(T)$, 按上述 $W'$ 和 $V'$ 的对偶基计算出 $\mathcal{M}(T')$. 利用这些基可以得出下面这个漂亮的结论.

**3.132 $T'$ 的矩阵是 $T$ 的矩阵的转置**

设 $V$ 和 $W$ 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

$$\mathcal{M}(T') = \big(\mathcal{M}(T)\big)^{\mathrm{t}}.$$

**证明** 令 $A = \mathcal{M}(T)$ 且 $C = \mathcal{M}(T')$. 设 $1 \le j \le m$ 且 $1 \le k \le n$.

由 $\mathcal{M}(T')$ 的定义, 我们有

$$T'(\psi_j) = \sum_{r=1}^{n} C_{r,j}\varphi_r.$$

上式左侧等于 $\psi_j \circ T$. 于是将上式两侧均作用于 $v_k$ 上, 可得

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j}\varphi_r(v_k)\\ &= C_{k,j}.\end{aligned}$$

我们还可写出

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \psi_j(Tv_k)\\ &= \psi_j\left(\sum_{r=1}^{m} A_{r,k}w_r\right)\\ &= \sum_{r=1}^{m} A_{r,k}\psi_j(w_r)\\ &= A_{j,k}.\end{aligned}$$

比照上面两组等式的最后一行，我们有 $C_{k,j} = A_{j,k}$. 于是 $C = A^{\mathrm{t}}$. 换言之，$\mathcal{M}(T') = \big(\mathcal{M}(T)\big)^{\mathrm{t}}$. ■

现在，我们利用对偶给出矩阵的列秩等于行秩的另一种证法. 我们之前用别的工具证明了该结论——见 3.57.

**3.133 列秩等于行秩**

设 $A \in \mathbf{F}^{m,n}$. 那么 $A$ 的列秩等于 $A$ 的行秩.

**证明** 定义 $T : \mathbf{F}^{n,1} \to \mathbf{F}^{m,1}$ 为 $Tx = Ax$. 于是 $\mathcal{M}(T) = A$，其中 $\mathcal{M}(T)$ 是关于 $\mathbf{F}^{n,1}$ 和 $\mathbf{F}^{m,1}$ 的标准基计算得到的. 这样一来，

$$\begin{aligned}A\text{ 的列秩} &= \mathcal{M}(T)\text{ 的列秩}\\ &= \dim \operatorname{range} T\\ &= \dim \operatorname{range} T'\\ &= \mathcal{M}(T')\text{ 的列秩}\\ &= A^{\mathrm{t}}\text{ 的列秩}\\ &= A\text{ 的行秩},\end{aligned}$$

其中第二个等号源于 3.78，第三个等号源于 3.130 (a)，第四个等号源于 3.78，第五个等号源于 3.132，最后一个等号则源于列秩和行秩的定义. ■

上述结果的另一种证明方法见于 7A 节的习题 8.

## 习题 3F

**1** 解释为什么线性泛函不是满射就是零映射.

**2** 给出三个不同的例子：$\mathbf{R}^{[0,1]}$ 上的线性泛函.

**3** 设 $V$ 是有限维的，且 $v \in V$（$v \neq 0$）. 证明存在 $\varphi \in V'$ 使得 $\varphi(v) = 1$.

**4** 设 $V$ 是有限维的，且 $U$ 是 $V$ 的子空间，$U \neq V$. 证明：存在 $\varphi \in V'$ 使得 $\varphi(u) = 0$ 对任意 $u \in U$ 成立但 $\varphi \neq 0$.

**5** 设 $T \in \mathcal{L}(V,W)$, $w_1,\dots,w_m$ 是 range $T$ 的基. 于是对每个 $v \in V$, 存在唯一的 $\varphi_1(v),\dots,\varphi_m(v)$ 使得

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

从而定义了从 $V$ 到 $\mathbf{F}$ 的函数 $\varphi_1,\dots,\varphi_m$. 证明函数 $\varphi_1,\dots,\varphi_m$ 都是 $V$ 上的线性泛函.

**6** 设 $\varphi,\beta \in V'$. 证明：$\operatorname{null}\varphi \subseteq \operatorname{null}\beta$ 当且仅当存在 $c \in \mathbf{F}$ 使得 $\beta = c\varphi$.

**7** 设 $V_1,\dots,V_m$ 是向量空间. 证明：$(V_1 \times \cdots \times V_m)'$ 和 $V_1{}' \times \cdots \times V_m{}'$ 是同构向量空间.

**8** 设 $v_1,\dots,v_n$ 是 $V$ 的基，$\varphi_1,\dots,\varphi_n$ 是 $V'$ 的对偶基. 定义 $\Gamma : V \to \mathbf{F}^n$ 和 $\Lambda : \mathbf{F}^n \to V$ 为

$$\Gamma(v) = \bigl(\varphi_1(v),\dots,\varphi_n(v)\bigr) \quad \text{和} \quad \Lambda(a_1,\dots,a_n) = a_1v_1 + \cdots + a_nv_n.$$

解释为什么 $\Gamma$ 和 $\Lambda$ 是彼此的逆.

**9** 设 $m$ 是一正整数. 证明：$\mathcal{P}_m(\mathbf{R})$ 的基 $1,x,\dots,x^m$ 的对偶基是 $\varphi_0,\varphi_1,\dots,\varphi_m$，其中

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

注　这里 $p^{(k)}$ 表示 $p$ 的 $k$ 阶导数，$p$ 的 0 阶导数应理解为 $p$ 本身.

**10** 设 $m$ 是一正整数.

(a) 证明 $1, x-5,\dots,(x-5)^m$ 是 $\mathcal{P}_m(\mathbf{R})$ 的基.

(b) (a) 中的基的对偶基是什么？

**11** 设 $v_1,\dots,v_n$ 是 $V$ 的基，$\varphi_1,\dots,\varphi_n$ 是 $V'$ 相应的对偶基. 设 $\psi \in V'$. 证明

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

**12** 设 $S,T \in \mathcal{L}(V,W)$.

(a) 证明 $(S+T)' = S' + T'$.

(b) 证明 $(\lambda T)' = \lambda T'$ 对所有 $\lambda \in \mathbf{F}$ 成立.

注　这道习题是让你验证 3.120 的 (a) 和 (b).

**13** 证明：$V$ 上的恒等算子的对偶映射是 $V'$ 上的恒等算子.

**14** 定义 $T : \mathbf{R}^3 \to \mathbf{R}^2$ 为

$$T(x,y,z) = (4x + 5y + 6z, 7x + 8y + 9z).$$

设 $\varphi_1,\varphi_2$ 表示 $\mathbf{R}^2$ 的标准基的对偶基，$\psi_1,\psi_2,\psi_3$ 表示 $\mathbf{R}^3$ 的标准基的对偶基.

(a) 描述 $T'(\varphi_1)$ 和 $T'(\varphi_2)$ 这两个线性泛函.

(b) 将 $T'(\varphi_1)$ 和 $T'(\varphi_2)$ 都写成 $\psi_1,\psi_2,\psi_3$ 的线性组合.

**15** 定义 $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ 为

$$(Tp)(x) = x^2p(x) + p''(x)$$

对每个 $x \in \mathbf{R}$ 成立.

(a) 设 $\varphi \in \mathcal{P}(\mathbf{R})'$ 定义为 $\varphi(p) = p'(4)$. 描述 $\mathcal{P}(\mathbf{R})$ 上的线性泛函 $T'(\varphi)$.

(b) 设 $\varphi \in \mathcal{P}(\mathbf{R})'$ 定义为 $\varphi(p) = \int_0^1 p$. 计算 $\bigl(T'(\varphi)\bigr)(x^3)$.

**16** 设 $W$ 是有限维的，$T \in \mathcal{L}(V,W)$. 证明

$$T' = 0 \iff T = 0.$$

**17** 设 $V$ 和 $W$ 是有限维的，$T \in \mathcal{L}(V,W)$. 证明：$T$ 可逆，当且仅当 $T' \in \mathcal{L}(W',V')$ 可逆.

**18** 设 $V$ 和 $W$ 是有限维的，证明：将 $T \in \mathcal{L}(V,W)$ 映到 $T' \in \mathcal{L}(W',V')$ 的映射是将 $\mathcal{L}(V,W)$ 映成 $\mathcal{L}(W',V')$ 的同构.

**19** 设 $U \subseteq V$，解释为什么

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null}\varphi\}.$$

**20** 设 $V$ 是有限维的，且 $U$ 是 $V$ 的子空间．证明：

$$U = \big\{v \in V : \varphi(v) = 0 \text{ 对任一 } \varphi \in U^0 \text{ 都成立}\big\}.$$

**21** 设 $V$ 是有限维的，且 $U$ 和 $W$ 是 $V$ 的子空间.

(a) 证明：$W^0 \subseteq U^0$ 当且仅当 $U \subseteq W$.

(b) 证明：$W^0 = U^0$ 当且仅当 $U = W$.

**22** 设 $V$ 是有限维的，且 $U$ 和 $W$ 是 $V$ 的子空间.

(a) 证明 $(U+W)^0 = U^0 \cap W^0$.

(b) 证明 $(U \cap W)^0 = U^0 + W^0$.

**23** 设 $V$ 是有限维的，且 $\varphi_1,\dots,\varphi_m \in V'$．证明以下三个集合彼此相等：

(a) $\operatorname{span}(\varphi_1,\dots,\varphi_m)$

(b) $\big((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m)\big)^0$

(c) $\{\varphi \in V' : (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m) \subseteq \operatorname{null}\varphi\}$

**24** 设 $V$ 是有限维的，且 $v_1,\dots,v_m \in V$.定义一线性映射 $\Gamma : V' \to \mathbf{F}^m$ 为 $\Gamma(\varphi) = \big(\varphi(v_1),\dots,\varphi(v_m)\big)$.

(a) 证明：$v_1,\dots,v_m$ 张成 $V$，当且仅当 $\Gamma$ 是单射.

(b) 证明：$v_1,\dots,v_m$ 线性无关，当且仅当 $\Gamma$ 是满射.

**25** 设 $V$ 是有限维的，且 $\varphi_1,\dots,\varphi_m \in V'$.定义一线性映射 $\Gamma : V \to \mathbf{F}^m$ 为 $\Gamma(v) = \big(\varphi_1(v),\dots,\varphi_m(v)\big)$.

(a) 证明：$\varphi_1,\dots,\varphi_m$ 张成 $V'$，当且仅当 $\Gamma$ 是单射.

(b) 证明：$\varphi_1,\dots,\varphi_m$ 线性无关，当且仅当 $\Gamma$ 是满射.

**26** 设 $V$ 是有限维的，且 $\Omega$ 是 $V'$ 的子空间．证明

$$\Omega = \big\{v \in V : \varphi(v) = 0 \text{ 对任一 } \varphi \in \Omega \text{ 都成立}\big\}^0.$$

**27** 设 $T \in \mathcal{L}\big(\mathcal{P}_5(\mathbf{R})\big)$ 且 $\operatorname{null}T' = \operatorname{span}(\varphi)$，其中 $\varphi$ 是 $\mathcal{P}_5(\mathbf{R})$ 上的线性泛函，定义为 $\varphi(p) = p(8)$. 证明

$$\operatorname{range}T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.$$

**28** 设 $V$ 是有限维的，且 $\varphi_1,\dots,\varphi_m$ 是 $V'$ 中的线性无关组．证明

$$\dim\big((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m)\big) = (\dim V) - m.$$

**29** 设 $V$ 和 $W$ 是有限维的，且 $T \in \mathcal{L}(V,W)$.

(a) 证明：如果 $\varphi \in W'$ 且 $\operatorname{null}T' = \operatorname{span}(\varphi)$，那么 $\operatorname{range}T = \operatorname{null}\varphi$.

(b) 证明：如果 $\psi \in V'$ 且 $\operatorname{range}T' = \operatorname{span}(\psi)$，那么 $\operatorname{null}T = \operatorname{null}\psi$.

**30** 设 $V$ 是有限维的，且 $\varphi_1,\dots,\varphi_n$ 是 $V'$ 的基．证明 $V$ 有一个基的对偶基为 $\varphi_1,\dots,\varphi_n$.

**31** 设 $U$ 是 $V$ 的子空间，令 $i: U \to V$ 是定义为 $i(u) = u$ 的包含映射．从而 $i' \in \mathcal{L}(V', U')$．

(a) 证明 $\operatorname{null} i' = U^0$．

(b) 证明：如果 $V$ 是有限维的，那么 $\operatorname{range} i' = U'$．

(c) 证明：如果 $V$ 是有限维的，那么 $\tilde{i'}$ 是将 $V'/U^0$ 映成 $U'$ 的同构．

注　(c) 中的同构是自然的，因为它不依赖于其中任一向量空间的基的选取．

**32** $V$ 的**双重对偶空间**（**double dual space**），记为 $V''$，定义为 $V'$ 的对偶空间．换句话说，$V'' = (V')'$．定义 $\Lambda: V \to V''$ 为

$$(\Lambda v)(\varphi) = \varphi(v)$$

对任一 $v \in V$ 和任一 $\varphi \in V'$ 成立．

(a) 证明 $\Lambda$ 是从 $V$ 到 $V''$ 的线性映射．

(b) 证明：如果 $T \in \mathcal{L}(V)$，那么 $T'' \circ \Lambda = \Lambda \circ T$，其中 $T'' = (T')'$．

(c) 证明：如果 $V$ 是有限维的，那么 $\Lambda$ 是将 $V$ 映成 $V''$ 的同构．

注　设 $V$ 是有限维的．那么 $V$ 和 $V'$ 同构，但是找到从 $V$ 到 $V'$ 的同构一般需要选取 $V$ 的基．相反，从 $V$ 到 $V''$ 的同构 $\Lambda$ 不需要选取基，因而被认为是更加自然的．

**33** 设 $U$ 是 $V$ 的子空间．令 $\pi: V \to V/U$ 为通常的商映射．从而 $\pi' \in \mathcal{L}\big((V/U)', V'\big)$．

(a) 证明 $\pi'$ 是单射．

(b) 证明 $\operatorname{range} \pi' = U^0$．

(c) 得出结论：$\pi'$ 是将 $(V/U)'$ 映成 $U^0$ 的同构．

注　(c) 中的同构是自然的，因为它不依赖于其中任一向量空间的基的选取．实际上，此处还没有假设这些向量空间里的任何一个是有限维的．

# 第 4 章 多项式

本章包含有关多项式的内容，我们将利用它们来研究从一个向量空间映射到其自身的线性映射. 本章中有许多结果可能已在其他课程中为你所熟悉，仍将它们纳入本书中是为了内容完整起见.

因为本章主题与线性代数无关，所以你的老师可能会很快讲完它，并可能不会要求你细究所有证明. 但是，请确保你至少阅读并理解本章中所有结果的表述——后续章节中会用到它们.

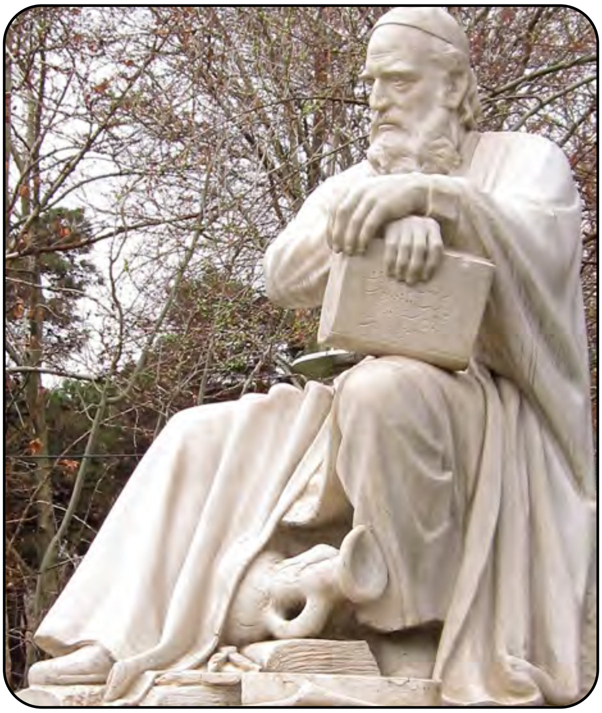
本章从简要讨论复数的代数性质开始. 然后，我们证明了不为常值的多项式的零点个数不超过它的次数. 我们还基于线性代数证明了多项式带余除法，即使你已经熟悉不用线性代数的证法，这也值得一读.

我们将看到，由代数基本定理，可将每个多项式分解为一次因式之积（若标量域为 $\mathbf{C}$）或不高于二次的因式之积（若标量域为 $\mathbf{R}$）.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.

♣





图为数学家和诗人奥马尔·海亚姆（Omar Khayyam，1048-1131）的雕像. 由他著于 1070 年的代数书首度包含了对三次多项式的严谨研究.

在讨论复系数多项式或实系数多项式之前，我们需要再学习一点有关复数的知识.

**4.1 定义：实部（real part）、$\mathrm{Re}\, z$，虚部（imaginary part）、$\mathrm{Im}\, z$**

设 $z = a + b\mathrm{i}$，其中 $a$ 和 $b$ 为实数.

- $z$ 的**实部**，记作 $\mathrm{Re}\, z$，定义为 $\mathrm{Re}\, z = a$；
- $z$ 的**虚部**，记作 $\mathrm{Im}\, z$，定义为 $\mathrm{Im}\, z = b$.

于是，对于每个复数 $z$，我们有

$$z = \mathrm{Re}\, z + (\mathrm{Im}\, z)\mathrm{i}.$$

**4.2 定义：复共轭（complex conjugate）、$\bar{z}$，绝对值（absolute value）、$|z|$**

设 $z \in \mathbf{C}$.

- $z \in \mathbf{C}$ 的**复共轭**，记作 $\bar{z}$，定义为

$$\bar{z} = \mathrm{Re}\, z - (\mathrm{Im}\, z)\mathrm{i}.$$

- 复数 $z$ 的**绝对值**，记作 $|z|$，定义为

$$|z| = \sqrt{(\mathrm{Re}\, z)^2 + (\mathrm{Im}\, z)^2}.$$

**4.3 例：实部和虚部、复共轭、绝对值**

设 $z = 3 + 2\mathrm{i}$. 那么

- $\mathrm{Re}\, z = 3$，$\mathrm{Im}\, z = 2$；
- $\bar{z} = 3 - 2\mathrm{i}$；
- $|z| = \sqrt{3^2 + 2^2} = \sqrt{13}$.

将复数 $z \in \mathbf{C}$ 与有序对 $(\mathrm{Re}\, z, \mathrm{Im}\, z) \in \mathbf{R}^2$ 等同起来看，就是把 $\mathbf{C}$ 和 $\mathbf{R}^2$ 等同起来看. 注意，$\mathbf{C}$ 是一个 1 维复向量空间，但我们也可以把 $\mathbf{C}$（将其等同于 $\mathbf{R}^2$）视作一个 2 维的实向量空间.

每个复数的绝对值都是非负数. 具体而言，如果 $z \in \mathbf{C}$，那么 $|z|$ 等于由 $\mathbf{R}^2$ 中的原点到 $\mathbf{R}^2$ 中的点 $(\mathrm{Re}\, z, \mathrm{Im}\, z)$ 的距离.

你应该验证一下，$z = \bar{z}$ 当且仅当 $z$ 为实数.

实部和虚部、复共轭和绝对值的若干性质列在下面的结果框中.

**4.4 复数的性质**

设 $w, z \in \mathbf{C}$. 那么有下面等式和不等式成立.

$z$ 与 $\bar{z}$ 之和（**sum of $z$ and $\bar{z}$**）

$$z + \bar{z} = 2\,\mathrm{Re}\, z.$$

$z$ 与 $\bar{z}$ 之差（**difference of $z$ and $\bar{z}$**）

$$z - \bar{z} = 2(\mathrm{Im}\, z)\mathrm{i}.$$

$z$ 与 $\bar{z}$ 之积（**product of $z$ and $\bar{z}$**）

$$z\bar{z} = |z|^2.$$

转下页

**复共轭的可加性和可乘性（additivity and multiplicativity of complex conjugate）**

$\overline{w+z} = \overline{w} + \overline{z}$ 且 $\overline{wz} = \overline{w}\,\overline{z}$.

**复共轭的复共轭（double complex conjugate）**

$\overline{\overline{z}} = z$.

**实部和虚部以 $|z|$ 为界（real and imaginary parts are bounded by $|z|$）**

$|\operatorname{Re} z| \leq |z|$ 且 $|\operatorname{Im} z| \leq |z|$.

**复共轭的绝对值（absolute value of the complex conjugate）**

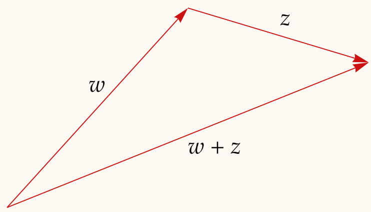
$|\overline{z}| = |z|$.

**绝对值的可乘性（multiplicativity of absolute value）**

$|wz| = |w||z|$.

**三角不等式（triangle inequality）**

$|w+z| \leq |w| + |z|$.




证明 除去上述最后一项以外，其余结论的验证都很常规，留给读者完成. 为验证三角不等式，我们有

三角不等式的几何解释：三角形的一边长小于或等于另两边长之和.

$$\begin{aligned}
|w+z|^2 &= (w+z)(\overline{w}+\overline{z}) \\
&= w\overline{w} + z\overline{z} + w\overline{z} + z\overline{w} \\
&= |w|^2 + |z|^2 + w\overline{z} + \overline{w\overline{z}} \\
&= |w|^2 + |z|^2 + 2\operatorname{Re}(w\overline{z}) \\
&\leq |w|^2 + |z|^2 + 2|w\overline{z}| \\
&= |w|^2 + |z|^2 + 2|w||z| \\
&= (|w|+|z|)^2.
\end{aligned}$$

两边开平方根即得我们欲证的不等式 $|w+z| \leq |w| + |z|$. ■

反向三角不等式见于章末习题 2.

## 多项式的零点

回忆一下，对于一个函数 $p: \mathbf{F} \to \mathbf{F}$，如果存在 $a_0, \ldots, a_m \in \mathbf{F}$ 且 $a_m \neq 0$ 使得对所有 $z \in \mathbf{F}$ 都有

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

则称 $p$ 是次数为 $m$ 的多项式. 如果按上述形式表达 $p$ 的方式不止一种，那么多项式的次数可能不唯一. 我们的首项任务就是证明这不可能发生.

方程 $p(z) = 0$ 的解在对多项式 $p \in \mathcal{P}(\mathbf{F})$ 的研究中扮演关键角色. 于是，我们给这些解赋予特殊的名称.

**4.5 定义：多项式的零点（zero of a polynomial）**

称一个数 $\lambda \in \mathbf{F}$ 为多项式 $p \in \mathcal{P}(\mathbf{F})$ 的零点【或根（root）】，若

$$p(\lambda) = 0.$$

下面结论将成为我们证明多项式的次数唯一性时所用的关键工具.

**4.6 多项式的每个零点都对应一个一次因式**

设 $m$ 是正整数且 $p \in \mathcal{P}(\mathbf{F})$ 是次数为 $m$ 的多项式. 设 $\lambda \in \mathbf{F}$. 那么 $p(\lambda) = 0$ 当且仅当存在一个次数为 $m-1$ 的多项式 $q \in \mathcal{P}(\mathbf{F})$ 使得对每个 $z \in \mathbf{F}$ 都有

$$p(z) = (z-\lambda)q(z).$$

**证明** 先设 $p(\lambda) = 0$. 令 $a_0, a_1, \ldots, a_m \in \mathbf{F}$ 使得对全体 $z \in \mathbf{F}$ 都有

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m.$$

那么对全体 $z \in \mathbf{F}$ 都有

$$p(z) = p(z) - p(\lambda) = a_1(z-\lambda) + \cdots + a_m(z^m - \lambda^m). \tag{4.7}$$

对每个 $k \in \{1, \ldots, m\}$，等式

$$z^k - \lambda^k = (z-\lambda)\sum_{j=1}^{k} \lambda^{j-1} z^{k-j}$$

表明，$z^k - \lambda^k$ 等于 $z-\lambda$ 乘以某个次数为 $k-1$ 的多项式. 于是式 (4.7) 表明 $p$ 等于 $z-\lambda$ 乘以某个次数为 $m-1$ 的多项式，该方向得证.

为了证明另一方向的蕴涵关系，现假设存在多项式 $q \in \mathcal{P}(\mathbf{F})$ 使得对每个 $z \in \mathbf{F}$ 都有 $p(z) = (z-\lambda)q(z)$. 那么 $p(\lambda) = (\lambda-\lambda)q(\lambda) = 0$，该方向得证. ■

现在我们就能证明多项式不会有太多零点了.

**4.8 次数为 $m$ 表明最多有 $m$ 个零点**

假定 $m$ 是正整数且 $p \in \mathcal{P}(\mathbf{F})$ 是次数为 $m$ 的多项式. 那么 $p$ 在 $\mathbf{F}$ 中最多有 $m$ 个零点.

**证明** 对 $m$ 用归纳法. 我们欲证的结论当 $m=1$ 时是成立的，因为若 $a_1 \neq 0$ 那么多项式 $a_0 + a_1 z$ 仅有一个零点（等于 $-a_0/a_1$）. 从而，假定 $m > 1$ 且欲证的结论对于 $m-1$ 情形成立.

如果 $p$ 在 $\mathbf{F}$ 中无零点，那么欲证结论成立，证明完成. 于是假设 $p$ 有一个零点 $\lambda \in \mathbf{F}$. 由 4.6，存在次数为 $m-1$ 的多项式 $q \in \mathcal{P}(\mathbf{F})$ 使得对于每个 $z \in \mathbf{F}$ 都有

$$p(z) = (z-\lambda)q(z).$$

我们所作的归纳假设表明，$q$ 在 $\mathbf{F}$ 中至多有 $m-1$ 个零点. 而上述等式表明，$p$ 在 $\mathbf{F}$ 中的零点恰为 $q$ 在 $\mathbf{F}$ 中的零点以及 $\lambda$. 于是 $p$ 在 $\mathbf{F}$ 中至多有 $m$ 个零点. ■

上述结论表明，一个多项式的系数是唯一确定的（因为如果一个多项式有两组不同的系数，那么将该多项式的这两种表示形式相减，就会得到一个某些系数非零却有无穷多个零点的多项式）. 特别地，多项式的次数是唯一确定的.

回忆一下，多项式 0 的次数定义为 $-\infty$. 必要情况下，应用有关 $-\infty$ 的一般算术规则即可. 例如，对每个整数 $m$ 有 $-\infty < m$ 且 $-\infty + m = -\infty$.

规定多项式 0 的次数为 $-\infty$，于是对于许多结论【例如 $\deg(pq) = \deg p + \deg q$】就不需要考虑例外情况.

## 多项式的带余除法

如果 $p$ 和 $s$ 是非负整数且 $s \neq 0$，那么存在非负整数 $q$ 和 $r$ 满足

$$p = sq + r$$

且 $r < s$. 可把此式看成将 $p$ 除以 $s$ 并得到商 $q$ 和余数 $r$. 接下来的结果给出了适用于多项式的类似结论. 因而，接下来的结果常被称为多项式的带余除法，尽管这里所说的并不真是一种计算法，而仅仅是个有用的结论.

多项式的带余除法可以不用任何线性代数知识来证明. 然而，此处的证明使用了线性代数的技巧，并巧妙利用了 $\mathcal{P}_n(\mathbf{F})$ 的基【$\mathcal{P}_n(\mathbf{F})$ 是由系数在 $\mathbf{F}$ 中且次数至多为 $n$ 的多项式所构成的 $n+1$ 维向量空间】. 这种证法对一本线性代数教材来说很合适.

可把多项式的带余除法看成：将多项式 $p$ 除以多项式 $s$，并得到余多项式 $r$.

**4.9 多项式的带余除法（division algorithm for polynomials）**

设 $p, s \in \mathcal{P}(\mathbf{F})$，其中 $s \neq 0$. 那么存在唯一的多项式 $q, r \in \mathcal{P}(\mathbf{F})$ 满足

$$p = sq + r$$

且 $\deg r < \deg s$.

**证明** 令 $n = \deg p$ 且 $m = \deg s$. 如果 $n < m$，那么取 $q = 0$ 与 $r = p$ 即可得欲证等式 $p = sq + r$ 并有 $\deg r < \deg s$. 从而我们现在假设 $n \geq m$.

向量组

$$1, z, \ldots, z^{m-1}, s, zs, \ldots, z^{n-m}s \tag{4.10}$$

在 $\mathcal{P}_n(\mathbf{F})$ 中线性无关，因为该组中每个多项式都有不同的次数. 同时，组 (4.10) 的长度是 $n+1$，这等于 $\dim \mathcal{P}_n(\mathbf{F})$. 所以，(4.10) 就是 $\mathcal{P}_n(\mathbf{F})$ 的基（由 2.38）.

因为 $p \in \mathcal{P}_n(\mathbf{F})$ 且 (4.10) 是 $\mathcal{P}_n(\mathbf{F})$ 的基，故存在唯一的常数 $a_0, a_1, \ldots, a_{m-1} \in \mathbf{F}$ 和 $b_0, b_1, \ldots, b_{n-m} \in \mathbf{F}$ 使得

$$\begin{aligned} p &= a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + b_0 s + b_1 z s + \cdots + b_{n-m} z^{n-m} s \\ &= \underbrace{a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}}_{r} + s\big(\underbrace{b_0 + b_1 z + \cdots + b_{n-m} z^{n-m}}_{q}\big). \end{aligned} \tag{4.11}$$

按照 $r$ 和 $q$ 的上述定义，我们发现 $p$ 可被写成 $p = sq + r$ 且 $\deg r < \deg s$，存在性得证.

满足这些条件的 $q, r \in \mathcal{P}(\mathbf{F})$ 的唯一性，源于满足式 (4.11) 的常数 $a_0, a_1, \ldots, a_{m-1} \in \mathbf{F}$ 和 $b_0, b_1, \ldots, b_{n-m} \in \mathbf{F}$ 的唯一性. ■

## 多项式在 C 上的分解

通过令 $\mathbf{F}$ 指代 $\mathbf{R}$ 或 $\mathbf{C}$，我们一直在同时跟实系数多项式与复系数多项式打交道.现在我们将看到这两种情形的区别. 我们先讨论复系数多项式，然后用适于复系数多项式的结论来证明适于实系数多项式的对应结论.

代数基本定理是个存在性定理. 它的证明并不能引出求解零点的方法. 二次多项式的零点由二次求根公式明确给出；对于三次和四次多项式也有与之相似但更为复杂的公式. 对于五次及以上的多项式，这样的求根公式就不存在了.

我们对代数基本定理的证明隐式地用到了“连续实值函数在 $\mathbf{R}^2$ 中的闭圆盘上取得最小值”这一结论. 通过上网搜索，你还能看到代数基本定理的一些其他证法. 如果你对解析函数得心应手，那么利用刘维尔定理（Liouville's theorem）的证法尤其好用. 代数基本定理的所有证明都需要用到一些分析学的知识，因为如果 $\mathbf{C}$ 被替换掉【例如换成形如 $c+d\mathrm{i}$（$c$ 和 $d$ 为有理数）的数的集合】，那么这个结论就不再成立了.

**4.12 代数基本定理，版本一（fundamental theorem of algebra, first version）**

每个不是常值的复系数多项式都在 $\mathbf{C}$ 中有零点.

**证明** 棣莫弗定理（De Moivre's theorem）的内容是，如果 $k$ 为正整数且 $\theta \in \mathbf{R}$，那么

$$(\cos\theta + \mathrm{i}\sin\theta)^k = \cos k\theta + \mathrm{i}\sin k\theta.$$

该定理可通过对 $k$ 作数学归纳法并用余弦和正弦的加法公式来证明.

假设 $w \in \mathbf{C}$ 且 $k$ 为正整数. 利用极坐标，我们可知存在 $r \geq 0$ 且 $\theta \in \mathbf{R}$ 使得

$$r(\cos\theta + \mathrm{i}\sin\theta) = w.$$

棣莫弗定理表明

$$\left(r^{1/k}\left(\cos\tfrac{\theta}{k} + \mathrm{i}\sin\tfrac{\theta}{k}\right)\right)^k = w.$$

于是每个复数都有 $k$ 次方根. 我们很快就会利用到这个结论.

假设 $p$ 是不为常值的复系数多项式且最高次非零项为 $c_m z^m$. 那么当 $|z| \to \infty$ 时 $|p(z)| \to \infty$（因为当 $|z| \to \infty$ 时 $|p(z)|/|z^m| \to |c_m|$）. 于是连续函数 $z \mapsto |p(z)|$ 在某点 $\zeta \in \mathbf{C}$ 处取得全局最小值. 为了说明 $p(\zeta) = 0$，先假设 $p(\zeta) \neq 0$.

定义一个新的多项式 $q$ 为

$$q(z) = \frac{p(z+\zeta)}{p(\zeta)}.$$

函数 $z \mapsto |q(z)|$ 在 $z = 0$ 处有全局最小值 1. 将 $q(z)$ 写成

$$q(z) = 1 + a_k z^k + \cdots + a_m z^m,$$

其中 $k$ 是满足 $z^k$ 的系数非零的最小正整数，换言之，$a_k \neq 0$.

令 $\beta \in \mathbf{C}$ 使得 $\beta^k = -\dfrac{1}{a_k}$. 那么存在常数 $c > 1$ 满足：若 $t \in (0,1)$，那么

$$\begin{aligned}|q(t\beta)| &\leq |1 + a_k t^k \beta^k| + t^{k+1}c \\ &= 1 - t^k(1 - tc).\end{aligned}$$

从而，在上述不等式中取 $t$ 为 $1/(2c)$，我们就可得 $|q(t\beta)| < 1$，与 $z \mapsto |q(z)|$ 的全局最小值为 1 的假设相矛盾. 由该矛盾就可推出 $p(\zeta) = 0$，证明了 $p$ 有一个零点，则原命题得证. ■

计算机能通过精巧的数值计算方法求出任何多项式的零点的充分近似值，即便确切的零点是求不出来的. 例如，对于定义为

$$p(x) = x^5 - 5x^4 - 6x^3 + 17x^2 + 4x - 7$$

的多项式 $p$，人们就无法给出其零点的确切求解公式. 但计算机可以求得 $p$ 的零点大致是 $-1.87, -0.74, 0.62, 1.47, 5.51$ 这五个数.

代数基本定理的版本一引导我们得到复系数多项式的如下分解定理. 注意，在该分解式中，$p$ 的零点恰为数 $\lambda_1, \ldots, \lambda_m$，只有当 $z$ 取这些数时才能让下面结论中等式的右侧等于 0.

**4.13 代数基本定理，版本二（fundamental theorem of algebra, second version）**

如果 $p \in \mathcal{P}(\mathbf{C})$ 是不恒为常数的多项式，那么 $p$ 可被唯一分解为（不计因式的顺序）

$$p(z) = c(z - \lambda_1)\cdots(z - \lambda_m),$$

其中 $c, \lambda_1, \ldots, \lambda_m \in \mathbf{C}$.

**证明** 设 $p \in \mathcal{P}(\mathbf{C})$ 且 $m = \deg p$. 我们对 $m$ 用数学归纳法. 若 $m = 1$，那么我们所期望的分解式存在且唯一. 因此，设 $m > 1$，并设对于所有 $m-1$ 次多项式，我们所期望的分解式都存在且唯一.

首先我们证明，我们所期望的 $p$ 的分解式是存在的. 由代数基本定理的版本一（4.12），$p$ 具有零点 $\lambda \in \mathbf{C}$. 由 4.6，存在 $m-1$ 次多项式 $q$ 使得对所有 $z \in \mathbf{C}$ 都有

$$p(z) = (z - \lambda)q(z).$$

我们的归纳假设表明，$q$ 可分解为我们所期望的形式，将其分解式代入上式即得到我们所期望的 $p$ 的分解式.

现在我们转而考虑唯一性问题. $c$ 就是 $p$ 中 $z^m$ 前的系数，因此是唯一确定的. 所以我们只需证明 $\lambda_1, \ldots, \lambda_m$ 仅有一种选取方法（不考虑次序）. 若

$$(z - \lambda_1)\cdots(z - \lambda_m) = (z - \tau_1)\cdots(z - \tau_m)$$

对所有 $z \in \mathbf{C}$ 都成立，那么由于当 $z = \lambda_1$ 时此式左侧等于 0，所以此式右侧会有某个 $\tau$ 等于 $\lambda_1$. 我们把下标改一下，即可假设 $\tau_1 = \lambda_1$. 那么如果 $z \neq \lambda_1$，我们就可以将上式两侧同除以 $z - \lambda_1$，得到

$$(z - \lambda_2)\cdots(z - \lambda_m) = (z - \tau_2)\cdots(z - \tau_m).$$

此式除了在 $z = \lambda_1$ 可能不成立外，对其他所有 $z \in \mathbf{C}$ 都成立. 但其实此式一定对所有 $z \in \mathbf{C}$ 都成立，否则将上式左端减去右端，就会得到一个非零却有无限多个零点的多项式. 上面等式和我们的归纳假设表明，各 $\lambda$ 与各 $\tau$ 相同（不考虑次序），这就完成了唯一性的证明. ■

## 多项式在 R 上的分解

一个实系数多项式可能没有实数零点. 例如，多项式 $1+x^2$ 没有实数零点.

我们会利用 $\mathbf{C}$ 上的分解定理来得出 $\mathbf{R}$ 上的分解定理. 我们从下面结论开始.

代数基本定理在 $\mathbf{R}$ 上不成立，由此可解释线性代数在实向量空间和复向量空间上的差异. 在之后的章节中我们将对此作进一步了解.

**4.14 实系数多项式的非实数零点成对出现**

设 $p \in \mathcal{P}(\mathbf{C})$ 为实系数多项式. 若 $\lambda \in \mathbf{C}$ 是 $p$ 的零点，那么 $\bar{\lambda}$ 也是 $p$ 的零点.

**证明** 令

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

其中 $a_0, \ldots, a_m$ 是实数. 设 $\lambda \in \mathbf{C}$ 是 $p$ 的零点，那么

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0.$$

等式两边同取复共轭得

$$a_0 + a_1\bar{\lambda} + \cdots + a_m\bar{\lambda}^m = 0.$$

其中我们用到了复共轭的基本性质（见 4.4）. 上式就表明 $\bar{\lambda}$ 是 $p$ 的零点. ■

我们想得到实系数多项式的分解定理. 我们从下面的结论出发.

将二次求根公式和下面结论结合起来思考.

**4.15 二次多项式的分解**

设 $b, c \in \mathbf{R}$. 那么当且仅当 $b^2 \geq 4c$ 时，存在 $\lambda_1, \lambda_2 \in \mathbf{R}$ 使得形如

$$x^2 + bx + c = (x-\lambda_1)(x-\lambda_2)$$

的分解式成立.

**证明** 注意到

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right).$$

上式是配方法这个技巧的基础.

先设 $b^2 < 4c$. 于是对每个 $x \in \mathbf{R}$，上式右侧均为正. 因此多项式 $x^2+bx+c$ 没有实数零点，从而也就不能被分解成 $(x-\lambda_1)(x-\lambda_2)$（其中 $\lambda_1, \lambda_2 \in \mathbf{R}$）的形式.

反之，现假设 $b^2 \geq 4c$. 那么存在实数 $d$ 使得 $d^2 = \frac{b^2}{4} - c$. 根据上面单行列出的公式有，

$$\begin{aligned} x^2 + bx + c &= \left(x + \frac{b}{2}\right)^2 - d^2 \\ &= \left(x + \frac{b}{2} + d\right)\left(x + \frac{b}{2} - d\right), \end{aligned}$$

即得我们想要的分解式. ■

接下来的结论给出了多项式在 $\mathbf{R}$ 上的分解式. 该结论的证明思想是利用代数基本定理的版本二（4.13），这个定理给出的是将 $p$ 视作复系数多项式时的分解式. 由 4.14 知，$p$ 的非实

数复零点总是成对出现．于是，如果将 $p$ 视作 $\mathcal{P}(\mathbf{C})$ 的元素时，所得分解式含有形如 $(x-\lambda)$ 的项（其中 $\lambda$ 为非实数的复数），那么 $(x-\bar{\lambda})$ 也是分解式的一项．将这两项相乘可得

$$x^2-2(\operatorname{Re}\lambda)x+|\lambda|^2,$$

这个二次的项正具备我们所需要的形式．

根据上段概述的思路基本上能证明我们所期望的分解式是存在的．然而，有一点需要注意：设 $\lambda$ 是非实数的复数，且 $(x-\lambda)$ 是将 $p$ 视作 $\mathcal{P}(\mathbf{C})$ 的元素时所得分解式的其中一项，那么根据 4.14 我们可以保证 $(x-\bar{\lambda})$ 肯定也是分解式中的一项，但是 4.14 并没说这两个因式出现的次数相同，而要使上段的思路可行，这一点是必需的．不过我们将证明这一点的确成立．

在下面结果中，$m$ 和 $M$ 中可能会有一个等于 0．数 $\lambda_1,\dots,\lambda_m$ 恰是 $p$ 的实数零点，因为 $x$ 仅有取这些实数值时才能使下面结论中等式的右侧等于 0．

**4.16 多项式在 R 上的分解**

设 $p\in\mathcal{P}(\mathbf{R})$ 是一个不恒为常数的多项式．那么 $p$ 可被唯一分解为（不计因式的顺序）

$$p(x)=c(x-\lambda_1)\cdots(x-\lambda_m)(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M),$$

其中 $c,\lambda_1,\dots,\lambda_m,b_1,\dots,b_M,c_1,\dots,c_M\in\mathbf{R}$ 且对各 $k=1,\dots,M$ 均有 $b_k^2<4c_k$．

**证明** 首先我们证明我们所期望的分解式是存在的，然后证明其唯一性．

将 $p$ 视作 $\mathcal{P}(\mathbf{C})$ 的一个元素．如果 $p$ 的所有（复）零点都是实的，那么根据 4.13 就得到了我们想要的分解式．于是，设 $p$ 存在零点 $\lambda\in\mathbf{C}$ 且 $\lambda\notin\mathbf{R}$．由 4.14 可得 $\bar{\lambda}$ 也是 $p$ 的零点．于是我们可写出

$$\begin{aligned}p(x)&=(x-\lambda)(x-\bar{\lambda})q(x)\\&=\left(x^2-2(\operatorname{Re}\lambda)x+|\lambda|^2\right)q(x),\end{aligned}$$

其中 $q\in\mathcal{P}(\mathbf{C})$ 是比 $p$ 低 2 次的多项式．如果我们可以证明 $q$ 的系数为实数，那么对 $p$ 的次数用归纳法即可完成存在性部分的证明．

为了证明 $q$ 的系数是实数，通过上述方程解出 $q$，可得对于所有 $x\in\mathbf{R}$，

$$q(x)=\frac{p(x)}{x^2-2(\operatorname{Re}\lambda)x+|\lambda|^2}.$$

上式意味着对于所有 $x\in\mathbf{R}$ 都有 $q(x)\in\mathbf{R}$．将 $q$ 写成

$$q(x)=a_0+a_1x+\cdots+a_{n-2}x^{n-2},$$

其中 $n=\deg p$ 且 $a_0,\dots,a_{n-2}\in\mathbf{C}$．由此我们可得

$$0=\operatorname{Im}q(x)=(\operatorname{Im}a_0)+(\operatorname{Im}a_1)x+\cdots+(\operatorname{Im}a_{n-2})x^{n-2}$$

对所有 $x\in\mathbf{R}$ 都成立．这就意味着 $\operatorname{Im}a_0,\dots,\operatorname{Im}a_{n-2}$ 都等于 0（由 4.8）．于是 $q$ 的所有系数都是实数．因此，我们所期望的分解式存在．

现在我们转而考虑分解式的唯一性问题．$p$ 的一个形如 $x^2+b_kx+c_k$ 且满足 $b_k^2<4c_k$ 的因式，可以被唯一地表示为 $(x-\lambda_k)(x-\overline{\lambda_k})$，其中 $\lambda_k\in\mathbf{C}$．不难想到，若将 $p$ 作为 $\mathcal{P}(\mathbf{R})$ 中元素来看时存在两个不同的分解式，会导致将其作为 $\mathcal{P}(\mathbf{C})$ 中元素来看时，也有两个不同的分解式，而这与 4.13 矛盾． ■

# 习题 4

**1** 设 $w, z \in \mathbf{C}$. 验证下列等式和不等式：

(a) $z + \bar{z} = 2\operatorname{Re} z$

(b) $z - \bar{z} = 2(\operatorname{Im} z)\mathrm{i}$

(c) $z\bar{z} = |z|^2$

(d) $\overline{w+z} = \bar{w} + \bar{z}$ 且 $\overline{wz} = \bar{w}\,\bar{z}$

(e) $\bar{\bar{z}} = z$

(f) $|\operatorname{Re} z| \le |z|$ 且 $|\operatorname{Im} z| \le |z|$

(g) $|\bar{z}| = |z|$

(h) $|wz| = |w|\,|z|$

注 以上结果是 4.4 留给读者的部分.

**2** 证明：如果 $w, z \in \mathbf{C}$，那么 $\big||w| - |z|\big| \le |w - z|$.

注 以上不等式称为**反向三角不等式（reverse triangle inequality）**.

**3** 设 $V$ 是复向量空间且 $\varphi \in V'$. 定义 $\sigma : V \to \mathbf{R}$ 为 $\sigma(v) = \operatorname{Re}\varphi(v)$ 对任一 $v \in V$ 都成立. 证明

$$\varphi(v) = \sigma(v) - \mathrm{i}\sigma(\mathrm{i}v)$$

对所有 $v \in V$ 成立.

**4** 设 $m$ 是一正整数，集合

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p = m\}$$

是 $\mathcal{P}(\mathbf{F})$ 的子空间吗?

**5** 集合

$$\{0\} \cup \left\{p \in \mathcal{P}(\mathbf{F}) : \deg p \text{ 为偶数}\right\}$$

是 $\mathcal{P}(\mathbf{F})$ 的子空间吗?

**6** 设 $m$ 和 $n$ 是正整数（$m \le n$），$\lambda_1, \dots, \lambda_m \in \mathbf{F}$. 证明：存在多项式 $p \in \mathcal{P}(\mathbf{F})$（$\deg p = n$），使得 $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ 且 $p$ 没有其他零点.

**7** 设 $m$ 是一非负整数，$z_1, \dots, z_{m+1}$ 是 $\mathbf{F}$ 中的不同元素，$w_1, \dots, w_{m+1} \in \mathbf{F}$. 证明：存在唯一的多项式 $p \in \mathcal{P}_m(\mathbf{F})$ 使得

$$p(z_k) = w_k$$

对每个 $k = 1, \dots, m+1$ 成立.

注 这个结果可以不用线性代数证明. 但是试试看，用上些线性代数，找到更简洁明了的证明.

**8** 设 $p \in \mathcal{P}(\mathbf{C})$ 的次数为 $m$. 证明：$p$ 有 $m$ 个不同零点，当且仅当 $p$ 和其导数 $p'$ 没有共同的零点.

**9** 证明：每个实系数奇次多项式都有实零点.

**10** 对于 $p \in \mathcal{P}(\mathbf{R})$，定义 $Tp : \mathbf{R} \to \mathbf{R}$ 为

$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(3)}{x - 3} & \text{若 } x \neq 3, \\ p'(3) & \text{若 } x = 3 \end{cases}$$

对每个 $x \in \mathbf{R}$ 都成立. 证明 $Tp \in \mathcal{P}(\mathbf{R})$ 对每个多项式 $p \in \mathcal{P}(\mathbf{R})$ 成立, 以及 $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ 是线性映射.

**11** 设 $p \in \mathcal{P}(\mathbf{C})$. 定义 $q : \mathbf{C} \to \mathbf{C}$ 为

$$q(z) = p(z)\overline{p(\bar{z})}.$$

证明 $q$ 是实系数多项式.

**12** 设 $m$ 是一非负整数，$p \in \mathcal{P}_m(\mathbf{C})$，存在不同的实数 $x_0, x_1, \ldots, x_m$ 使得 $p(x_k) \in \mathbf{R}$ 对每个 $k = 0, 1, \ldots, m$ 都成立. 证明 $p$ 的所有系数都是实的.

**13** 设 $p \in \mathcal{P}(\mathbf{F})$（$p \neq 0$），令 $U = \{pq : q \in \mathcal{P}(\mathbf{F})\}$.

(a) 证明 $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

(b) 求 $\mathcal{P}(\mathbf{F})/U$ 的一个基.

**14** 设 $p, q \in \mathcal{P}(\mathbf{C})$ 是不为常值的多项式，且没有共同的零点. 令 $m = \deg p$，$n = \deg q$. 按照下面 (a)–(c) 的方式，使用线性代数证明：存在 $r \in \mathcal{P}_{n-1}(\mathbf{C})$ 和 $s \in \mathcal{P}_{m-1}(\mathbf{C})$ 使得

$$rp + sq = 1.$$

(a) 定义 $T : \mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C}) \to \mathcal{P}_{m+n-1}(\mathbf{C})$ 为

$$T(r, s) = rp + sq.$$

证明线性映射 $T$ 是单射.

(b) 证明 (a) 中的线性映射 $T$ 是满射.

(c) 利用 (b) 得出结论：存在 $r \in \mathcal{P}_{n-1}(\mathbf{C})$ 和 $s \in \mathcal{P}_{m-1}(\mathbf{C})$ 使得 $rp + sq = 1$.

# 第 5 章 特征值和特征向量

由一个向量空间到另一个向量空间的线性映射是第 3 章的研究对象. 现在我们开始研究算子, 也就是由一个向量空间到其本身的线性映射. 对于算子的研究构成了线性代数中最为重要的组成部分.

为了研究一个算子, 我们尝试将其限制在一个更小的子空间中. 讨论将该算子限制于什么样的子空间内能使其仍为算子, 将引导我们得到不变子空间的概念. 每个一维不变子空间, 都产生于一个这样的向量——它被该算子映射为自身的标量倍. 沿着这条研究路径, 我们就会得出特征向量和特征值的概念.

接着, 我们将证明线性代数中最为重要的结论之一: 非零有限维复向量空间上的每个算子都有特征值. 这条结论使我们得以证明, 对于有限维复向量空间上的每个算子, 都存在该向量空间的一个基, 使得该算子关于该基的矩阵至少有近一半的元素等于 0.

**以下假设在本章中总是成立:**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 代表 $\mathbf{F}$ 上的向量空间.

♣


图为比萨的列奥纳多 (Leonardo of Pisa, 约 1170-约 1250) 之雕像, 他亦被称为斐波那契 (Fibonacci). 5D 节的习题 21 展示了如何用线性代数求出斐波那契数列的通项公式 (示于封面).

# 5A 不变子空间

## 特征值

**5.1 定义：算子（operator）**

称从一个向量空间到其本身的线性映射为**算子**.

回忆一下，我们将 $\mathcal{L}(V)$ 定义成 $\mathcal{L}(V,V)$.

设 $T \in \mathcal{L}(V)$. 若 $m \geq 2$ 且

$$V = V_1 \oplus \cdots \oplus V_m,$$

其中各 $V_k$ 都是 $V$ 的非零子空间，那么要理解 $T$ 的作用，我们只需要理解各 $T|_{V_k}$ 的作用即可. 此处 $T|_{V_k}$ 表示将 $T$ 限于更小的定义空间 $V_k$ 上. 因为 $V_k$ 是比 $V$ 小的向量空间，所以讨论 $T|_{V_k}$ 要比讨论 $T$ 简单些.

然而，我们如果想应用算子研究中一些有用的工具（例如取算子的幂），那么就会遇到这个问题：$T|_{V_k}$ 可能不会将 $V_k$ 映射至它本身；换句话说，$T|_{V_k}$ 可能不是 $V_k$ 上的算子. 由此，在 $V$ 的形式如上的分解式中，我们仅考虑满足 $T$ 将各 $V_k$ 都映射到本身的情形. 因此，现在我们给 $V$ 的被 $T$ 映射到本身的子空间取个名称.

**5.2 定义：不变子空间（invariant subspace）**

设 $T \in \mathcal{L}(V)$. 称 $V$ 的子空间 $U$ 是**不变的**，若对每个 $u \in U$ 均有 $Tu \in U$.

于是，如果 $T|_U$是 $U$ 上的算子，则 $U$ 在 $T$ 下是不变的.

**5.3 例：在微分算子下不变的子空间**

设 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 定义为 $Tp = p'$. 那么 $\mathcal{P}(\mathbf{R})$ 的子空间 $\mathcal{P}_4(\mathbf{R})$ 在 $T$ 下是不变的，因为若 $p \in \mathcal{P}(\mathbf{R})$ 的次数不高于 4，那么 $p'$ 的次数亦不高于 4.

**5.4 例：四个不变子空间（并不一定全都不同）**

若 $T \in \mathcal{L}(V)$，那么 $V$ 的以下子空间在 $T$ 下都是不变的.

| | |
|---|---|
| $\{0\}$ | 子空间 $\{0\}$ 在 $T$ 下是不变的，因为若 $u \in \{0\}$，那么 $u = 0$，因此 $Tu = 0 \in \{0\}$. |
| $V$ | 子空间 $V$ 在 $T$ 下是不变的，因为若 $u \in V$，那么 $Tu \in V$. |
| $\operatorname{null} T$ | 子空间 $\operatorname{null} T$ 在 $T$ 下是不变的，因为若 $u \in \operatorname{null} T$，那么 $Tu = 0$，因此 $Tu \in \operatorname{null} T$. |
| $\operatorname{range} T$ | 子空间 $\operatorname{range} T$ 在 $T$ 下是不变的，因为若 $u \in \operatorname{range} T$，那么 $Tu \in \operatorname{range} T$. |

一个算子 $T \in \mathcal{L}(V)$ 一定会有除了 $\{0\}$ 和 $V$ 以外的不变子空间吗？稍后我们会看到，如果 $V$ 是有限维的，且 $\dim V > 1$（$\mathbf{F} = \mathbf{C}$）或 $\dim V > 2$（$\mathbf{F} = \mathbf{R}$），那么这个问题的回答是肯定的. 可参看 5.19 与 5B 节的习题 29.

上例中提到，$\operatorname{null} T$ 和 $\operatorname{range} T$ 在 $T$ 下是不变的. 然而，不能仅凭这两个不变子空间存在，就轻率地回答上述有关除了 $\{0\}$ 和 $V$ 以外的不变子空间存在性的问题，因为有可能 $\operatorname{null} T$ 等于 $\{0\}$ 而 $\operatorname{range} T$ 等于 $V$（$T$ 可逆时就会出现这种情况）.

我们稍后再回过头来更深入地研究不变子空间. 现在我们转而研究最简单的非平凡不变子空间——维数为 1 的不变子空间.

任取 $v \in V$ 且 $v \neq 0$，并令 $U$ 等于 $v$ 的所有标量倍构成的集合：

$$U = \{\lambda v : \lambda \in \mathbf{F}\} = \operatorname{span}(v).$$

那么，$U$ 是 $V$ 的一维子空间（且只要适当选取 $v$，$V$ 的每个一维子空间都可写成这种形式）. 若 $U$ 在算子 $T \in \mathcal{L}(V)$ 下是不变的，那么 $Tv \in U$，因此存在标量 $\lambda \in \mathbf{F}$ 使得

$$Tv = \lambda v.$$

反之，若对某 $\lambda \in \mathbf{F}$ 有 $Tv = \lambda v$，那么 $\operatorname{span}(v)$ 是 $V$ 的在 $T$ 下不变的一维子空间.

我们刚才看到，等式 $Tv = \lambda v$ 与一维不变子空间紧密联系在一起. 这个等式很重要，因此满足该式的标量 $\lambda$ 和向量 $v$ 都有特别的名称.

**5.5 定义：特征值（eigenvalue）**

设 $T \in \mathcal{L}(V)$. 称数 $\lambda \in \mathbf{F}$ 为 $T$ 的**特征值**，若存在 $v \in V$ 使得 $v \neq 0$ 且 $Tv = \lambda v$.

在上面定义中，我们要求 $v \neq 0$，因为每个标量 $\lambda \in \mathbf{F}$ 都满足 $T0 = \lambda 0$.

上面的讨论表明，$V$ 有在 $T$ 下不变的一维子空间，当且仅当 $T$ 有特征值.

“**eigenvalue**”这个词，一半是德文，一半是英文. 德文前缀“**eigen**”意为“自身的”，是指表征固有的性质.

**5.6 例：特征值**

定义算子 $T \in \mathcal{L}(\mathbf{F}^3)$ 为：对 $(x, y, z) \in \mathbf{F}^3$，

$$T(x, y, z) = (7x + 3z, 3x + 6y + 9z, -6y).$$

那么 $T(3, 1, -1) = (18, 6, -6) = 6(3, 1, -1)$. 所以 6 是 $T$ 的特征值.

下面结论中的诸等价关系，乃至线性代数中许多深刻的结论，都只在有限维向量空间中成立.

**5.7 成为特征值的等价条件**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$.

那么下面几个命题等价.

(a) $\lambda$ 是 $T$ 的特征值.

(b) $T - \lambda I$ 不是单射.

(c) $T - \lambda I$ 不是满射.

(d) $T - \lambda I$ 不可逆.

提示：$I \in \mathcal{L}(V)$ 是恒等算子. 于是，对所有 $v \in V$ 有 $Iv = v$.

**证明** (a) 与 (b) 等价是因为等式 $Tv = \lambda v$ 等价于等式 $(T - \lambda I)v = 0$. (b)、(c) 和 (d) 等价是由于 3.65. ■

**5.8 定义：特征向量（eigenvector）**

设 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$ 是 $T$ 的特征值．若向量 $v \in V$ 满足 $v \neq 0$ 且 $Tv = \lambda v$，则称该向量是 $T$ 对应于 $\lambda$ 的**特征向量**.

换句话说，非零向量 $v \in V$ 是算子 $T \in \mathcal{L}(V)$ 的特征向量，当且仅当 $Tv$ 是 $v$ 的标量倍．因为 $Tv = \lambda v$ 当且仅当 $(T - \lambda I)v = 0$，所以，向量 $v \in V$（$v \neq 0$）是 $T$ 对应于 $\lambda$ 的特征向量，当且仅当 $v \in \operatorname{null}(T - \lambda I)$.

---

**5.9 例：特征值和特征向量**

设 $T \in \mathcal{L}(\mathbf{F}^2)$ 定义为 $T(w, z) = (-z, w)$.

(a) 首先考虑 $\mathbf{F} = \mathbf{R}$ 的情况．那么 $T$ 的作用就是将向量绕 $\mathbf{R}^2$ 的原点逆时针旋转 $90°$．一个算子有特征值，当且仅当其定义空间内存在可被其映射为自身标量倍的非零向量．将 $\mathbf{R}^2$ 中的非零向量逆时针旋转 $90°$ 不会得到原向量的标量倍．所以：若 $\mathbf{F} = \mathbf{R}$，那么 $T$ 没有特征值（于是也没有特征向量）.

(b) 现考虑 $\mathbf{F} = \mathbf{C}$ 的情况．为求出 $T$ 的特征值，我们必须求出标量 $\lambda$，使得 $T(w, z) = \lambda(w, z)$ 有除 $w = z = 0$ 之外的解．方程 $T(w, z) = \lambda(w, z)$ 等价于联立方程

$$-z = \lambda w, \quad w = \lambda z. \tag{5.10}$$

将第二个方程所给出的 $w$ 的表达式代入第一个方程，得

$$-z = \lambda^2 z.$$

由于 $z$ 不能等于 0【否则由式 (5.10) 得 $w = 0$，而我们要找的式 (5.10) 的解须满足 $(w, z)$ 不是向量 0】，所以上述方程就可化为

$$-1 = \lambda^2.$$

此方程的解是 $\lambda = \mathrm{i}$ 和 $\lambda = -\mathrm{i}$.

你可验证，$\mathrm{i}$ 和 $-\mathrm{i}$ 都是 $T$ 的特征值．的确，对应于特征值 $\mathrm{i}$ 的特征向量是形如 $(w, -w\mathrm{i})$ 的向量，其中 $w \in \mathbf{C}$ 且 $w \neq 0$；对应于特征值 $-\mathrm{i}$ 的特征向量是形如 $(w, w\mathrm{i})$ 的向量，其中 $w \in \mathbf{C}$ 且 $w \neq 0$.

---

在下面证明中，我们再次使用了如下等价关系：

$$Tv = \lambda v \iff (T - \lambda I)v = 0.$$

**5.11 线性无关的特征向量**

设 $T \in \mathcal{L}(V)$．那么分别对应于 $T$ 的不同特征值的特征向量构成的每个组都线性无关.

**证明** 假设欲证结论不成立．那么存在最小的正整数 $m$，使得 $T$ 对应于其互异特征值 $\lambda_1, \ldots, \lambda_m$ 的特征向量 $v_1, \ldots, v_m$ 构成线性相关向量组（注意 $m \geq 2$，因为根据定义，特征向量是非零的）.

于是，存在 $a_1,\ldots,a_m \in \mathbf{F}$，其中各数均非零（因为 $m$ 最小[1]），使得

$$a_1v_1+\cdots+a_mv_m=0.$$

将 $T-\lambda_m I$ 作用于上式两侧，得

$$a_1(\lambda_1-\lambda_m)v_1+\cdots+a_{m-1}(\lambda_{m-1}-\lambda_m)v_{m-1}=0.$$

因为特征值 $\lambda_1,\ldots,\lambda_m$ 互异，所以上式中各系数均不等于 0. 于是 $v_1,\ldots,v_{m-1}$ 就是由对应于 $T$ 的互异特征值的 $m-1$ 个特征向量构成的线性相关向量组，这和 $m$ 最小的假设相矛盾. 此矛盾就说明欲证命题成立. ■

上面结论可引出下面结论的一个简短证明. 下面结论给一个算子的互异特征值个数设置了上界.

**5.12 算子的特征值个数不多于向量空间的维数**

设 $V$ 是有限维的. 那么 $V$ 上的每个算子最多有 $\dim V$ 个互异特征值. ♡

**证明** 令 $T\in\mathcal{L}(V)$. 设 $\lambda_1,\ldots,\lambda_m$ 是 $T$ 的互异特征值. 令 $v_1,\ldots,v_m$ 是与之对应的特征向量. 那么 5.11 表明 $v_1,\ldots,v_m$ 线性无关. 于是 $m\le\dim V$（见 2.22），证毕. ■

## 将多项式作用于算子

算子（它们将向量空间映射到其自身）的理论比一般的线性映射的理论更丰富的主要原因，就在于算子可以自乘为幂. 在本小节中，我们给出算子的幂的定义，并提出将多项式作用于算子的概念. 这个概念，将是我们在下一节中证明“非零有限维复向量空间上的每个算子都有特征值”时，所使用的关键工具.

若 $T$ 是一个算子，那么 $TT$ 是有意义的（见 3.7），且它也是 $T$ 的定义空间上的算子. 我们通常把 $TT$ 写成 $T^2$. 更一般地，我们定义 $T^m$ 如下.

**5.13 记号：$T^m$**

设 $T\in\mathcal{L}(V)$，$m$ 是正整数.

- 定义 $T^m\in\mathcal{L}(V)$ 为 $T^m=\underbrace{T\cdots T}_{m\text{ 个 }T}$.
- 定义 $T^0$ 为 $V$ 上的恒等算子 $I$.
- 若 $T$ 是可逆的，且其逆为 $T^{-1}$，那么 $T^{-m}\in\mathcal{L}(V)$ 的定义是

$$T^{-m}=(T^{-1})^m.$$

♣

你应自行验证，若 $T$ 为算子，那么

$$T^mT^n=T^{m+n}\quad 且\quad (T^m)^n=T^{mn},$$

其中，若 $T$ 是可逆的，则 $m$ 和 $n$ 是任意整数；若 $T$ 不可逆，则 $m$ 和 $n$ 为非负整数.

定义了算子的幂，我们就可以定义什么是“将多项式作用于算子”了.

---

[1]原文：because of the minimality of $m$，直译为“因为 $m$ 的最小性”，意即 $m$ 是满足“$T$ 对应于其互异特征值 $\lambda_1,\ldots,\lambda_m$ 的特征向量 $v_1,\ldots,v_m$ 线性相关”的最小值. 后文通过得出更小的 $m-1$ 也满足此条件来推出矛盾.

**5.14 记号：$p(T)$**

设 $T \in \mathcal{L}(V)$ 且 $p \in \mathcal{P}(\mathbf{F})$ 是由下式给出的多项式：对所有 $z \in \mathbf{F}$，

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m.$$

那么 $p(T)$ 是 $V$ 上的算子，由下式定义：

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

♣

我们把 $p$ 应用在算子上而不仅是 $\mathbf{F}$ 中的数上，给符号 $p$ 增添了一种新用法．我们的想法是，要得到 $p(T)$，将 $p$ 的定义式中的 $z$ 替换成 $T$ 即可．注意，$p(z)$ 中的常数项 $a_0$ 变成了算子 $a_0 I$（这写法很合理：因为 $a_0 = a_0 z^0$，所以我们应将 $a_0$ 替换成 $a_0 T^0$，也就是 $a_0 I$）.

**5.15 例：将多项式作用于微分算子**

假设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是由 $Dq = q'$ 定义的微分算子，$p$ 是定义为 $p(x) = 7 - 3x + 5x^2$ 的多项式．那么 $p(D) = 7I - 3D + 5D^2$．于是对每个 $q \in \mathcal{P}(\mathbf{R})$，都有

$$\bigl(p(D)\bigr)q = 7q - 3q' + 5q''.$$

若取定一算子 $T \in \mathcal{L}(V)$，则从 $\mathcal{P}(\mathbf{F})$ 到 $\mathcal{L}(V)$ 的函数 $p \mapsto p(T)$ 是线性的．你应自行验证这一点.

**5.16 定义：多项式的乘积（product of polynomials）**

若 $p, q \in \mathcal{P}(\mathbf{F})$，那么 $pq \in \mathcal{P}(\mathbf{F})$ 是按下式定义的多项式：对所有 $z \in \mathbf{F}$，

$$(pq)(z) = p(z)q(z).$$

♣

如下面结论 (b) 所示，对单个算子的多项式取乘积时，顺序无关紧要.

**5.17 乘积性质**

设 $p, q \in \mathcal{P}(\mathbf{F})$ 且 $T \in \mathcal{L}(V)$．那么：

(a) $(pq)(T) = p(T)q(T)$；

(b) $p(T)q(T) = q(T)p(T)$.

不太正式的证明：利用分配性质展开多项式之乘积时，跟符号用 $z$ 还是 $T$ 没有关系.

♡

**证明**

(a) 设 $p(z) = \sum_{j=0}^{m} a_j z^j$ 且 $q(z) = \sum_{k=0}^{n} b_k z^k$ 对所有 $z \in \mathbf{F}$ 都成立．那么

$$(pq)(z) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k z^{j+k}.$$

于是

$$\begin{aligned}(pq)(T) &= \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k} \\ &= \Bigl(\sum_{j=0}^{m} a_j T^j\Bigr)\Bigl(\sum_{k=0}^{n} b_k T^k\Bigr) \\ &= p(T)q(T).\end{aligned}$$

(b) 运用 (a) 两次，我们可得 $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$. ■

之前我们已经看到，若 $T \in \mathcal{L}(V)$，那么子空间 $\operatorname{null} T$ 和 $\operatorname{range} T$ 在 $T$ 下是不变的（见例 5.4）. 现在我们证明，$T$ 的每个多项式的零空间和值域在 $T$ 下也都是不变的.

**5.18 $p(T)$ 的零空间和值域在 $T$ 下是不变的**

设 $T \in \mathcal{L}(V)$ 且 $p \in \mathcal{P}(\mathbf{F})$. 那么 $\operatorname{null} p(T)$ 和 $\operatorname{range} p(T)$ 在 $T$ 下不变.

**证明** 设 $u \in \operatorname{null} p(T)$. 那么 $p(T)u = 0$. 于是

$$\big(p(T)\big)(Tu) = \big(p(T)T\big)(u) = \big(Tp(T)\big)(u) = T\big(p(T)u\big) = T(0) = 0.$$

因此 $Tu \in \operatorname{null} p(T)$. 于是 $\operatorname{null} p(T)$ 在 $T$ 下是不变的.

设 $u \in \operatorname{range} p(T)$. 那么存在 $v \in V$ 使得 $u = p(T)v$. 于是

$$Tu = T\big(p(T)v\big) = p(T)(Tv).$$

因此 $Tu \in \operatorname{range} p(T)$. 于是 $\operatorname{range} p(T)$ 在 $T$ 下是不变的. ■

## 习题 5A

**1** 设 $T \in \mathcal{L}(V)$ 且 $U$ 是 $V$ 的子空间.

(a) 证明：如果 $U \subseteq \operatorname{null} T$，那么 $U$ 在 $T$ 下不变.

(b) 证明：如果 $\operatorname{range} T \subseteq U$，那么 $U$ 在 $T$ 下不变.

**2** 设 $T \in \mathcal{L}(V)$ 且 $V$ 的子空间 $V_1, \ldots, V_m$ 在 $T$ 下不变. 证明 $V_1 + \cdots + V_m$ 在 $T$ 下不变.

**3** 设 $T \in \mathcal{L}(V)$. 证明：$V$ 的任意一族在 $T$ 下不变的子空间的交集在 $T$ 下不变.

**4** 证明或给出一反例：如果 $V$ 是有限维的，而 $U$ 是 $V$ 的子空间且在 $V$ 上任一算子下都不变，那么 $U = \{0\}$ 或 $U = V$.

**5** 设 $T \in \mathcal{L}(\mathbf{R}^2)$ 定义为 $T(x, y) = (-3y, x)$. 求 $T$ 的特征值.

**6** 定义 $T \in \mathcal{L}(\mathbf{F}^2)$ 为 $T(w, z) = (z, w)$. 求出 $T$ 的所有特征值和特征向量.

**7** 定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 为 $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. 求出 $T$ 的所有特征值和特征向量.

**8** 设 $P \in \mathcal{L}(V)$，$P^2 = P$. 证明：如果 $\lambda$ 是 $P$ 的特征值，那么 $\lambda = 0$ 或 $\lambda = 1$.

**9** 定义 $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ 为 $Tp = p'$. 求出 $T$ 的所有特征值和特征向量.

**10** 定义 $T \in \mathcal{L}\big(\mathcal{P}_4(\mathbf{R})\big)$ 为 $(Tp)(x) = xp'(x)$ 对所有 $x \in \mathbf{R}$ 成立. 求出 $T$ 的所有特征值和特征向量.

**11** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$ 且 $\alpha \in \mathbf{F}$. 证明：存在 $\delta > 0$，使得对所有满足 $0 < |\alpha - \lambda| < \delta$ 的 $\lambda \in \mathbf{F}$ 都有 $T - \lambda I$ 可逆.

**12** 设 $V = U \oplus W$，其中 $U$ 和 $W$ 是 $V$ 的非零子空间. 定义 $P \in \mathcal{L}(V)$ 为，对任一 $u \in U$ 和任一 $w \in W$ 都有 $P(u + w) = u$. 求出 $P$ 的所有特征值和特征向量.

**13** 设 $T \in \mathcal{L}(V)$，并设 $S \in \mathcal{L}(V)$ 可逆.

(a) 证明 $T$ 和 $S^{-1}TS$ 特征值相同.

(b) $T$ 的特征向量和 $S^{-1}TS$ 的特征向量之间的关系是什么？

**14** 给出一例：$\mathbf{R}^4$ 上没有（实）特征值的算子.

**15** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，$\lambda \in \mathbf{F}$. 证明：$\lambda$ 是 $T$ 的特征值，当且仅当 $\lambda$ 是对偶算子 $T' \in \mathcal{L}(V')$ 的特征值.

**16** 设 $v_1, \ldots, v_n$ 是 $V$ 的基，$T \in \mathcal{L}(V)$. 证明：如果 $\lambda$ 是 $T$ 的特征值，那么

$$|\lambda| \leq n \max\left\{\left|\mathcal{M}(T)_{j,k}\right| : 1 \leq j, k \leq n\right\},$$

其中 $\mathcal{M}(T)_{j,k}$ 表示 $T$ 关于基 $v_1, \ldots, v_n$ 的矩阵第 $j$ 行第 $k$ 列的元素.

注 关于 $|\lambda|$ 的另一个界，见 6A 节习题 19.

**17** 设 $\mathbf{F} = \mathbf{R}$，$T \in \mathcal{L}(V)$，$\lambda \in \mathbf{R}$. 证明：$\lambda$ 是 $T$ 的特征值，当且仅当 $\lambda$ 是复化 $T_{\mathrm{C}}$ 的特征值.

注 $T_{\mathrm{C}}$ 的定义见 3B 节习题 33.

**18** 设 $\mathbf{F} = \mathbf{R}$，$T \in \mathcal{L}(V)$，$\lambda \in \mathbf{C}$. 证明：$\lambda$ 是复化 $T_{\mathrm{C}}$ 的特征值，当且仅当 $\overline{\lambda}$ 是 $T_{\mathrm{C}}$ 的特征值.

**19** 证明：定义为

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots)$$

的前向移位算子 $T \in \mathcal{L}(\mathbf{F}^\infty)$ 没有特征值.

**20** 定义后向移位算子 $S \in \mathcal{L}(\mathbf{F}^\infty)$ 为

$$S(z_1, z_2, z_3, \ldots) = (z_2, z_3, \ldots).$$

(a) 证明 $\mathbf{F}$ 的每个元素都是 $S$ 的特征值.

(b) 求出 $S$ 的所有特征向量.

**21** 设 $T \in \mathcal{L}(V)$ 可逆.

(a) 设 $\lambda \in \mathbf{F}$（$\lambda \neq 0$）. 证明：$\lambda$ 是 $T$ 的特征值，当且仅当 $\frac{1}{\lambda}$ 是 $T^{-1}$ 的特征值.

(b) 证明：$T$ 和 $T^{-1}$ 的特征向量相同.

**22** 设 $T \in \mathcal{L}(V)$ 且 $V$ 中存在非零向量 $u$ 和 $w$ 使得

$$Tu = 3w \quad 且 \quad Tw = 3u.$$

证明：3 或 $-3$ 是 $T$ 的特征值.

**23** 设 $V$ 是有限维的，$S, T \in \mathcal{L}(V)$. 证明：$ST$ 和 $TS$ 的特征值相同.

**24** 设 $A$ 是 $n \times n$ 矩阵，其元素属于 $\mathbf{F}$. 定义 $T \in \mathcal{L}(\mathbf{F}^n)$ 为 $Tx = Ax$，其中 $\mathbf{F}^n$ 的元素被视为 $n \times 1$ 列向量.

(a) 设 $A$ 的每一行元素之和都等于 1. 证明：1 是 $T$ 的特征值.

(b) 设 $A$ 的每一列元素之和都等于 1. 证明：1 是 $T$ 的特征值.

**25** 设 $T \in \mathcal{L}(V)$，$u, w$ 都是 $T$ 的特征向量，且 $u + w$ 也是 $T$ 的特征向量. 证明：$u$ 和 $w$ 是 $T$ 的对应于同一特征值的特征向量.

**26** 设 $T \in \mathcal{L}(V)$ 使得 $V$ 中每个非零向量都是 $T$ 的特征向量. 证明：$T$ 是恒等算子的标量倍.

**27** 设 $V$ 是有限维的，且 $k \in \{1, \ldots, \dim V - 1\}$. 设 $T \in \mathcal{L}(V)$ 使得 $V$ 的任一 $k$ 维子空间都在 $T$ 下不变. 证明：$T$ 是恒等算子的标量倍.

**28** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 证明：$T$ 最多有 $1 + \dim \operatorname{range} T$ 个不同的特征值.

**29** 设 $T \in \mathcal{L}(\mathbf{R}^3)$，且 $-4$、5、$\sqrt{7}$ 是 $T$ 的特征值. 证明：存在 $x \in \mathbf{R}^3$ 使得 $Tx - 9x = (-4, 5, \sqrt{7})$.

**30** 设 $T \in \mathcal{L}(V)$，且 $(T - 2I)(T - 3I)(T - 4I) = 0$. 设 $\lambda$ 是 $T$ 的特征值. 证明：$\lambda = 2$ 或者 $\lambda = 3$ 或者 $\lambda = 4$.

**31** 给出一例：$T \in \mathcal{L}(\mathbf{R}^2)$ 使得 $T^4 = -I$.

**32** 设 $T \in \mathcal{L}(V)$ 没有特征值且 $T^4 = I$. 证明 $T^2 = -I$.

**33** 设 $T \in \mathcal{L}(V)$，且 $m$ 是一正整数.

(a) 证明：$T$ 是单射，当且仅当 $T^m$ 是单射.

(b) 证明：$T$ 是满射，当且仅当 $T^m$ 是满射.

**34** 设 $V$ 是有限维的，$v_1, \ldots, v_m \in V$. 证明：组 $v_1, \ldots, v_m$ 线性无关，当且仅当存在 $T \in \mathcal{L}(V)$ 使得 $v_1, \ldots, v_m$ 是 $T$ 的对应于不同特征值的特征向量.

**35** 设 $\lambda_1, \ldots, \lambda_n$ 是一组相异实数. 证明：组 $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ 在由 $\mathbf{R}$ 上的实值函数构成的向量空间中线性无关.

提示 令 $V = \operatorname{span}(e^{\lambda_1 x}, \ldots, e^{\lambda_n x})$，并定义算子 $D \in \mathcal{L}(V)$ 为 $Df = f'$. 求 $D$ 的特征值和特征向量.

**36** 设 $\lambda_1, \ldots, \lambda_n$ 是一组相异正数. 证明：组 $\cos(\lambda_1 x), \ldots, \cos(\lambda_n x)$ 在由 $\mathbf{R}$ 上的实值函数构成的向量空间中线性无关.

**37** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 定义 $\mathcal{A} \in \mathcal{L}\big(\mathcal{L}(V)\big)$ 为

$$\mathcal{A}(S) = TS$$

对任一 $S \in \mathcal{L}(V)$ 都成立. 证明：$T$ 的特征值所成的集合与 $\mathcal{A}$ 的特征值所成的集合相同.

**38** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $U$ 是 $V$ 的在 $T$ 下不变的子空间. **商算子（quotient operator）** $T/U \in \mathcal{L}(V/U)$ 定义为

$$(T/U)(v + U) = Tv + U$$

对任一 $v \in V$ 都成立.

(a) 证明 $T/U$ 的定义是有意义的（这需要用到 $U$ 在 $T$ 下不变的条件），并证明 $T/U$ 是 $V/U$ 上的算子.

(b) 证明：$T/U$ 的每个特征值都是 $T$ 的特征值.

**39** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 证明：$T$ 有特征值，当且仅当存在 $V$ 的 $\dim V - 1$ 维子空间在 $T$ 下不变.

**40** 设 $S, T \in \mathcal{L}(V)$ 且 $S$ 可逆. 设 $p \in \mathcal{P}(\mathbf{F})$ 是一多项式. 证明：

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

**41** 设 $T \in \mathcal{L}(V)$ 且 $U$ 是 $V$ 的在 $T$ 下不变的子空间. 证明：对每个多项式 $p \in \mathcal{P}(\mathbf{F})$，均有 $U$ 在 $p(T)$ 下不变.

**42** 定义 $T \in \mathcal{L}(\mathbf{F}^n)$ 为 $T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n)$.

(a) 求出 $T$ 的所有特征值和特征向量.

(b) 求出 $\mathbf{F}^n$ 的所有在 $T$ 下不变的子空间.

**43** 设 $V$ 是有限维的，$\dim V > 1$，且 $T \in \mathcal{L}(V)$. 证明 $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$.

# 5B 最小多项式

## 复向量空间上特征值的存在性

现在我们给出关于有限维复向量空间上算子的一个核心结论.

**5.19 特征值的存在性**

非零有限维复向量空间上的每个算子都有特征值.

**证明** 设 $V$ 是有限维复向量空间，其维数 $n > 0$，$T \in \mathcal{L}(V)$. 取 $v \in V$ 且 $v \neq 0$. 那么

$$v, Tv, T^2v, \ldots, T^nv$$

不是线性无关的，因为 $V$ 的维数是 $n$，该组的长度却是 $n+1$. 因此上述向量的某个线性组合（其中各系数不全为 0）会等于向量 0. 于是，存在次数最小的非常值多项式 $p$ 使得

$$p(T)v = 0.$$

根据代数基本定理的版本一（见 4.12），存在 $\lambda \in \mathbf{C}$ 使得 $p(\lambda) = 0$. 所以，存在多项式 $q \in \mathcal{P}(\mathbf{C})$ 使得

$$p(z) = (z - \lambda)q(z)$$

对每个 $z \in \mathbf{C}$ 都成立（见 4.6）. 由此可得（利用 5.17）

$$0 = p(T)v = (T - \lambda I)\big(q(T)v\big).$$

因为 $q$ 的次数小于 $p$ 的次数，所以我们可知 $q(T)v \neq 0$. 于是上述等式就表明 $\lambda$ 是 $T$ 的一个特征值，$q(T)v$ 是对应于它的特征向量. ■

上述证明的关键是运用代数基本定理. 习题 16 后的注记有助于解释为何代数基本定理与上述结论结合得如此紧密.

上述结论中的 $\mathbf{F} = \mathbf{C}$ 这个假设不能替换为 $\mathbf{F} = \mathbf{R}$——从例 5.9 可看出. 下例则说明，上述结论中的“有限维”这个假设也不能去掉.

---

**5.20 例：复向量空间上的一个无特征值的算子**

设 $T \in \mathcal{L}(\mathcal{P}(\mathbf{C}))$ 定义为 $(Tp)(z) = zp(z)$. 若 $p \in \mathcal{P}(\mathbf{C})$ 是非零多项式，那么 $Tp$ 的次数比 $p$ 的次数多 1，从而 $Tp$ 不会等于 $p$ 的标量倍. 因此 $T$ 没有特征值.

因为 $\mathcal{P}(\mathbf{C})$ 是无限维的，所以本例并不与上述结论矛盾.

---

## 特征值与最小多项式

在本小节中，我们引入与每个算子都相关联的一个重要的多项式. 我们从下面定义出发.

**5.21 定义：首一多项式（monic polynomial）**

**首一多项式**是最高次项系数等于 1 的多项式.

例如，多项式 $2 + 9z^2 + z^7$ 是次数为 7 的首一多项式.

**5.22 最小多项式的存在性、唯一性和次数**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 那么存在唯一的次数最小的首一多项式 $p \in \mathcal{P}(\mathbf{F})$，使得 $p(T) = 0$. 此外，$\deg p \leq \dim V$.

**证明** 若 $\dim V = 0$，那么 $I$ 就是 $V$ 上的零算子，从而我们取 $p$ 为常值多项式 1 即可.

现在对 $\dim V$ 用归纳法，于是假设 $\dim V > 0$ 且欲证结论对于所有维数更小的向量空间上的所有算子都成立. 设 $u \in V$（$u \neq 0$）. 组 $u, Tu, \ldots, T^{\dim V}u$ 的长度为 $1 + \dim V$，因而是线性相关的. 根据线性相关性引理（2.19），存在最小正整数 $m \leq \dim V$ 使得 $T^m u$ 是 $u, Tu, \ldots, T^{m-1}u$ 的线性组合. 于是存在标量 $c_0, c_1, c_2, \ldots, c_{m-1} \in \mathbf{F}$ 使得

$$c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u + T^m u = 0. \tag{5.23}$$

定义首一多项式 $q \in \mathcal{P}_m(\mathbf{F})$ 为

$$q(z) = c_0 + c_1 z + \cdots + c_{m-1}z^{m-1} + z^m.$$

那么式 (5.23) 就表明 $q(T)u = 0$.

若 $k$ 为非负整数，那么

$$q(T)(T^k u) = T^k\big(q(T)u\big) = T^k(0) = 0.$$

线性相关性引理（2.19）表明，$u, Tu, \ldots, T^{m-1}u$ 线性无关，于是由上式可得 $\dim \operatorname{null} q(T) \geq m$. 因此

$$\dim \operatorname{range} q(T) = \dim V - \dim \operatorname{null} q(T) \leq \dim V - m.$$

因为 $\operatorname{range} q(T)$ 在 $T$ 下是不变的（由 5.18），所以我们可以将归纳假设应用于 $\operatorname{range} q(T)$ 上的算子 $T|_{\operatorname{range} q(T)}$，从而存在首一多项式 $s \in \mathcal{P}(\mathbf{F})$ 使得

$$\deg s \leq \dim V - m \quad 且 \quad s(T|_{\operatorname{range} q(T)}) = 0.$$

因此，对于所有 $v \in V$，我们有

$$\big((sq)(T)\big)(v) = s(T)\big(q(T)v\big) = 0,$$

这是因为 $q(T)v \in \operatorname{range} q(T)$ 且 $s(T)|_{\operatorname{range} q(T)} = s(T|_{\operatorname{range} q(T)}) = 0$. 于是 $sq$ 是满足 $\deg sq \leq \dim V$ 和 $(sq)(T) = 0$ 的首一多项式.

上段表明，存在次数不高于 $\dim V$ 的首一多项式，其作用于 $T$ 可得算子 0. 于是，作用于 $T$ 得到算子 0 的次数最小的首一多项式存在，这就完成了本结论存在性部分的证明.

令 $p \in \mathcal{P}(\mathbf{F})$ 是使 $p(T) = 0$ 成立的次数最小的首一多项式. 为证明本结论的唯一性部分，设 $r \in \mathcal{P}(\mathbf{F})$ 是次数与 $p$ 相同的首一多项式，且 $r(T) = 0$. 那么 $(p - r)(T) = 0$ 且 $\deg(p - r) < \deg p$. 若 $p - r$ 不等于 0，那么将 $p - r$ 除以其最高次项的系数，可得一个作用于 $T$ 会得到算子 0 的首一多项式（且其次数小于 $p$ 的次数）. 于是 $p - r = 0$，即唯一性得证. ■

上述结论为下面的定义提供了依据.

**5.24 定义：最小多项式（minimal polynomial）**

设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 那么 $T$ 的**最小多项式**是唯一使得 $p(T) = 0$ 成立的次数最小的首一多项式 $p \in \mathcal{P}(\mathbf{F})$.

为计算一个算子 $T \in \mathcal{L}(V)$ 的最小多项式，我们需要求出使得

$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m$$

有解 $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$ 的最小正整数 $m$. 如果我们选取 $V$ 的一个基，再将上述方程中的 $T$ 换成 $T$ 的矩阵，那么可以将上述方程看成有 $(\dim V)^2$ 个方程、$m$ 个未知数 $c_0, c_1, \ldots, c_{m-1} \in \mathbf{F}$ 的线性方程组. 要知道解是否存在，我们用高斯消元法或者线性方程组的其他快速解法，对 $m = 1, 2, \ldots$ 诸值依次作检验，直到发现有解即可. 由 5.22，对于某个最小的正整数 $m \leq \dim V$，上述方程组有解. 那么 $T$ 的最小多项式就是 $c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$.

（通常）更快的做法是，选取 $v \in V$（$v \neq 0$）并考虑方程

$$c_0 v + c_1 T v + \cdots + c_{\dim V - 1} T^{\dim V - 1} v = -T^{\dim V} v. \tag{5.25}$$

利用 $V$ 的一个基，将上述方程转化为有 $\dim V$ 个方程、$\dim V$ 个未知数 $c_0, c_1, \ldots, c_{\dim V-1}$ 的线性方程组. 若该方程组有唯一解 $c_0, c_1, \ldots, c_{\dim V-1}$（这是最常出现的情况），那么标量 $c_0, c_1, \ldots, c_{\dim V-1}, 1$ 就是 $T$ 的最小多项式的各系数（因为 5.22 指出最小多项式的次数最多是 $\dim V$）.

此处的百分比估计值基于对几百万个随机的矩阵的实测.

考虑 $\mathbf{R}^4$ 上的算子（可看成关于标准基的 $4\times4$ 矩阵），运用上段方法并取 $v = (1,0,0,0)$. 对于由区间 $[-10, 10]$ 上的整数构成的 $4\times4$ 矩阵，此法适用于其中超过 99.8% 的矩阵；而对于由区间 $[-100, 100]$ 上的整数构成的 $4\times4$ 矩阵，此法适用于其中超过 99.999% 的矩阵.

下例展示了上面讨论的快速方法.

**5.26 例：$\mathbf{F}^5$ 上一算子的最小多项式**

设 $T \in \mathcal{L}(\mathbf{F}^5)$，其关于标准基 $e_1, e_2, e_3, e_4, e_5$ 的矩阵为

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

在式 (5.25) 中取 $v = e_1$，有

$$\begin{aligned} Te_1 &= e_2, & T^4 e_1 &= T(T^3 e_1) = Te_4 = e_5, \\ T^2 e_1 &= T(Te_1) = Te_2 = e_3, & T^5 e_1 &= T(T^4 e_1) = Te_5 = -3e_1 + 6e_2, \\ T^3 e_1 &= T(T^2 e_1) = Te_3 = e_4, \end{aligned}$$

于是 $3e_1 - 6Te_1 = -T^5 e_1$. 组 $e_1, Te_1, T^2 e_1, T^3 e_1, T^4 e_1$ 就等于组 $e_1, e_2, e_3, e_4, e_5$，是线性无关的，所以该组的其他线性组合都不会等于 $-T^5 e_1$. 因此 $T$ 的最小多项式就是 $3 - 6z + z^5$.

回忆一下，根据定义，$V$ 上算子的特征值和 $\mathcal{P}(\mathbf{F})$ 中多项式的零点都必为 $\mathbf{F}$ 中的元素. 特别地，若 $\mathbf{F} = \mathbf{R}$，那么特征值和零点都必为实数.

**5.27 特征值即最小多项式的零点**

设 $V$ 是有限维的且 $T \in \mathcal{L}(V)$.

(a) $T$ 的最小多项式的零点即 $T$ 的特征值.

(b) 若 $V$ 是复向量空间，那么 $T$ 的最小多项式具有下述形式

$$(z-\lambda_1)\cdots(z-\lambda_m),$$

其中 $\lambda_1,\ldots,\lambda_m$ 是 $T$ 的所有特征值（可能有重复）.

**证明** 令 $p$ 是 $T$ 的最小多项式.

(a) 首先假设 $\lambda \in \mathbf{F}$ 是 $p$ 的零点之一. 那么 $p$ 可以写成如下形式

$$p(z) = (z-\lambda)q(z),$$

其中 $q$ 是系数在 $\mathbf{F}$ 中的首一多项式（见 4.6）. 因为 $p(T)=0$，所以对全体 $v \in V$ 我们都有

$$0 = (T-\lambda I)\big(q(T)v\big).$$

因为 $\deg q = (\deg p) - 1$ 且 $p$ 是 $T$ 的最小多项式，所以至少存在一个向量 $v \in V$ 使得 $q(T)v \neq 0$. 因而，上述等式表明 $\lambda$ 是 $T$ 的一个特征值，即 $p$ 的零点都为 $T$ 的特征值.

为证明 $T$ 的每个特征值都是 $p$ 的一个零点，现设 $\lambda \in \mathbf{F}$ 是 $T$ 的一个特征值. 于是，存在 $v \in V$（$v \neq 0$）使得 $Tv = \lambda v$. 将 $T$ 反复作用在此式两端，即得对每个非负整数 $k$ 均有 $T^k v = \lambda^k v$. 于是

$$p(T)v = p(\lambda)v.$$

因为 $p$ 是 $T$ 的最小多项式，所以我们有 $p(T)v = 0$. 所以上述等式就表明 $p(\lambda)=0$. 于是 $\lambda$ 是 $p$ 的一个零点，即 $T$ 的特征值都为 $p$ 的零点.

(b) 利用 (a) 和代数基本定理的版本二（见 4.13）即可证明. ■

非零多项式的互异零点个数最多不超过其次数（见 4.8）. 因此，将上个结论的 (a) 结合“$V$ 上的算子的最小多项式的次数不超过 $\dim V$”这个结论，就能以另一种方式证明 5.12（该定理内容为，$V$ 上的算子至多有 $\dim V$ 个互异的特征值）.

每个首一多项式都是某个算子的最小多项式（该命题在习题 16 中证明），这是对例 5.26 的推广. 因而，5.27 (a) 就表明，求出算子的特征值的确切表达式，就等价于求出多项式的零点的确切表达式（因而有些算子的特征值的确切值是求不出来的）.

---

**5.28 例：无法确切求出特征值的算子**

令 $T \in \mathcal{L}(\mathbf{C}^5)$ 为定义如下的算子

$$T(z_1,z_2,z_3,z_4,z_5) = (-3z_5,\, z_1+6z_5,\, z_2,\, z_3,\, z_4).$$

$T$ 关于 $\mathbf{C}^5$ 的标准基的矩阵如例 5.26 所示. 在那个例子中，我们求得 $T$ 的最小多项式为

$$3-6z+z^5.$$

转下页

上述多项式的零点都无法用有理数及其根式与通常的算数规则来表达（这个结论的证明远远超出了线性代数的范围）. 又由于上述多项式的零点也就是 $T$ 的特征值【由 5.27 (a)】，所以我们也就无法用我们熟悉的任何形式来确切表达 $T$ 的任何特征值.

用数值计算方法（此处我们不做讨论）可得上述多项式的零点，也即 $T$ 的特征值，近似为下面这五个复数：

$$-1.67, \quad 0.51, \quad 1.40, \quad -0.12+1.59i, \quad -0.12-1.59i.$$

注意，该多项式的两个非实数零点互成复共轭. 这个结果在我们意料之中，毕竟上述多项式是实系数多项式（见 4.14）.

---

下面结论完整地刻画了作用于算子时会得到算子 0 的多项式.

**5.29 $q(T)=0 \iff q$ 是最小多项式的多项式倍**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $q \in \mathcal{P}(\mathbf{F})$. 那么 $q(T)=0$ 当且仅当 $q$ 是 $T$ 的最小多项式的多项式倍[2].

**证明** 令 $p$ 表示 $T$ 的最小多项式.

首先假设 $q(T)=0$. 由多项式的带余除法（4.9），存在多项式 $s, r \in \mathcal{P}(\mathbf{F})$ 使得

$$q = ps + r \tag{5.30}$$

且 $\deg r < \deg p$. 我们有

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

上式表明 $r=0$（否则，将 $r$ 除以其最高次项的系数可得一个首一多项式，将该多项式作用于 $T$ 能得到 0，而该多项式的次数比最小多项式的次数还低，这就产生了矛盾）. 于是式 (5.30) 化为 $q=ps$. 因此，$q$ 是 $p$ 的多项式倍，此即待证命题的一个方向.

为了证明另一方向，现假设 $q$ 是 $p$ 的多项式倍. 那么存在多项式 $s \in \mathcal{P}(\mathbf{F})$ 使得 $q=ps$. 我们有

$$q(T) = p(T)s(T) = 0\, s(T) = 0,$$

则此方向得证. ■

下面结论是上述结论的一个很漂亮的推论.

**5.31 受限算子的最小多项式**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $U$ 是 $V$ 的在 $T$ 下不变的子空间. 那么 $T$ 的最小多项式是 $T|_U$ 的最小多项式的多项式倍.

**证明** 设 $p$ 是 $T$ 的最小多项式. 于是对于所有 $v \in V$，有 $p(T)v=0$. 特别地，对于所有 $u \in U$，$p(T)u=0$. 于是 $p(T|_U)=0$. 现在将 5.29 中的 $T$ 换成 $T|_U$（即将该结论作用于 $T|_U$ 而不是 $T$ 上），即可说明 $p$ 是 $T|_U$ 的最小多项式的多项式倍. ■

有关商算子的类似结论详见习题 25.

[2]某多项式 $p$ 的多项式倍，就是指形如 $pq$（其中 $q$ 也为多项式）的多项式.

下面结论表明，算子的最小多项式的常数项决定了该算子是否可逆.

**5.32 $T$ 不可逆 $\Longleftrightarrow$ $T$ 的最小多项式的常数项为 0**

设 $V$ 是有限维的且 $T \in \mathcal{L}(V)$. 那么 $T$ 不可逆，当且仅当 $T$ 的最小多项式的常数项为 0.

**证明** 设 $T \in \mathcal{L}(V)$ 且 $p$ 是 $T$ 的最小多项式. 那么

$$\begin{aligned} T \text{ 不可逆} &\Longleftrightarrow 0 \text{ 是 } T \text{ 的特征值} \\ &\Longleftrightarrow 0 \text{ 是 } p \text{ 的零点} \\ &\Longleftrightarrow p \text{ 的常数项是 } 0, \end{aligned}$$

其中，第一个等价关系成立是根据 5.7，第二个等价关系成立是根据 5.27 (a)，而最后一个等价关系成立是因为 $p$ 的常数项就等于 $p(0)$. ■

## 奇数维的实向量空间上的特征值

下面的结论，将是我们用于证明“奇数维的实向量空间上的算子都有特征值”的关键工具.

**5.33 偶数维的零空间**

设 $\mathbf{F} = \mathbf{R}$ 且 $V$ 是有限维的，并设 $T \in \mathcal{L}(V)$，$b, c \in \mathbf{R}$ 使得 $b^2 < 4c$. 那么 $\dim \operatorname{null}(T^2+bT+cI)$ 是偶数.

**证明** 回忆一下，由 5.18，$\operatorname{null}(T^2+bT+cI)$ 在 $T$ 下是不变的. 将 $V$ 替换成 $\operatorname{null}(T^2+bT+cI)$ 并将 $T$ 替换成限制在 $\operatorname{null}(T^2+bT+cI)$ 上的 $T$，我们便可认为 $T^2+bT+cI=0$，这样一来，我们需证明的就是 $\dim V$ 为偶数. [3]

设 $\lambda \in \mathbf{R}$ 及 $v \in V$ 满足 $Tv = \lambda v$. 那么

$$0 = (T^2+bT+cI)v = (\lambda^2+b\lambda+c)v = \left(\left(\lambda+\frac{b}{2}\right)^2 + c - \frac{b^2}{4}\right)v.$$

上式最外层括号中的数是个正数. 于是上式表明 $v=0$. 因此我们证明了 $T$ 没有特征向量.

令 $U$ 是 $V$ 的子空间，它在 $T$ 下是不变的，并且，所有在 $T$ 下不变且维数为偶的 $V$ 的子空间中，它具有最大的维数. 若 $U=V$，那么我们的证明完成了；否则，设存在某个 $w \in V$ 满足 $w \notin U$.

令 $W = \operatorname{span}(w, Tw)$. 那么 $W$ 在 $T$ 下不变，因为 $T(Tw) = -bTw - cw$. 另外，$\dim W = 2$，否则 $w$ 就会成为 $T$ 的特征向量. 这样一来，

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2,$$

其中 $U \cap W = \{0\}$，否则 $U \cap W$ 就会成为 $V$ 的一个 1 维子空间且在 $T$ 下不变（这是不可能的，因为 $T$ 没有特征向量）.

因为 $U+W$ 在 $T$ 下不变，所以上述等式表明，存在在 $T$ 下不变、维数为偶且大于 $\dim U$ 的 $V$ 的子空间. 于是，$U \neq V$ 的假设不正确. 因此，$V$ 的维数为偶数. ■

[3] 自此处至证明结束的 $V$，都是指 $\operatorname{null}(T^2+bT+cI)$，而非定理陈述中的 $V$.

接下来这条结果指出，奇数维的向量空间上的每个算子都有特征值. 我们已知此结论对于有限维复向量空间成立（且无需奇数维的假设）. 于是在下面的证明中，我们将假设 $\mathbf{F}=\mathbf{R}$.

**5.34 奇数维向量空间上的算子总有特征值**

奇数维向量空间上的每个算子都有特征值.

**证明** 设 $\mathbf{F}=\mathbf{R}$ 且 $V$ 是有限维的. 令 $n=\dim V$，并设 $n$ 为奇数. 令 $T\in\mathcal{L}(V)$. 为证明 $T$ 有特征值，我们对 $n$ 用步长为 2 的归纳法. 首先，注意到若 $\dim V=1$，那么欲证结论成立，因为这样的话 $V$ 中的每个非零向量都是 $T$ 的特征向量.

现在，假设 $n\geq 3$ 且欲证结论对于维数小于 $n$ 的所有奇数维向量空间上的所有算子都成立. 令 $p$ 表示 $T$ 的最小多项式. 若对于某 $\lambda\in\mathbf{R}$，$p$ 是 $x-\lambda$ 的多项式倍，那么 $\lambda$ 是 $T$ 的一个特征值【由 5.27 (a)】，证明就完成了. 于是我们可设存在 $b,c\in\mathbf{R}$ 使得 $b^2<4c$ 且 $p$ 是 $x^2+bx+c$ 的多项式倍（见 4.16）.

从而，存在首一多项式 $q\in\mathcal{P}(\mathbf{R})$ 使得 $p(x)=q(x)(x^2+bx+c)$ 对所有 $x\in\mathbf{R}$ 成立. 这样一来，就有

$$0=p(T)=\big(q(T)\big)(T^2+bT+cI),$$

由此，在 $\operatorname{range}(T^2+bT+cI)$ 上 $q(T)$ 等于算子 0. 因为 $\deg q<\deg p$ 且 $p$ 是 $T$ 的最小多项式，所以上式表明 $\operatorname{range}(T^2+bT+cI)\neq V$.

线性映射基本定理（3.21）告诉我们

$$\dim V=\dim\operatorname{null}(T^2+bT+cI)+\dim\operatorname{range}(T^2+bT+cI).$$

因为 $\dim V$ 是奇数（由命题假设）且 $\dim\operatorname{null}(T^2+bT+cI)$ 是偶数（由 5.33），所以上式就说明 $\dim\operatorname{range}(T^2+bT+cI)$ 是奇数.

所以，$\operatorname{range}(T^2+bT+cI)$ 是 $V$ 的在 $T$ 下不变的子空间（由 5.18），其维数为奇且小于 $\dim V$. 这样一来，由我们的归纳假设可得，限制于 $\operatorname{range}(T^2+bT+cI)$ 的算子 $T$ 有特征值，也即 $T$ 有特征值. ■

上述结果的其他证明方法见 8B 节的习题 23 和 9C 节的习题 10.

## 习题 5B

1 设 $T\in\mathcal{L}(V)$，证明：9 是 $T^2$ 的特征值，当且仅当 3 或 $-3$ 是 $T$ 的特征值.

2 设 $V$ 是复向量空间，且 $T\in\mathcal{L}(V)$ 没有特征值. 证明：$V$ 的每个在 $T$ 下不变的子空间，不是 $\{0\}$ 就是无限维的.

3 设 $n$ 为整数且 $n>1$，且 $T\in\mathcal{L}(\mathbf{F}^n)$ 定义为

$$T(x_1,\ldots,x_n)=(x_1+\cdots+x_n,\ldots,x_1+\cdots+x_n).$$

(a) 求出 $T$ 的所有特征值和特征向量.

(b) 求出 $T$ 的最小多项式.

**注** $T$ 关于 $\mathbf{F}^n$ 的标准基的矩阵的各元素全为 1.

4 设 $\mathbf{F}=\mathbf{C}$，$T\in\mathcal{L}(V)$，$p\in\mathcal{P}(\mathbf{C})$ 是不为常值的多项式，$\alpha\in\mathbf{C}$. 证明：$\alpha$ 是 $p(T)$ 的特征值，当且仅当 $\alpha=p(\lambda)$ 对 $T$ 的某个特征值 $\lambda$ 成立.

**5** 给出一例在 $\mathbf{R}^2$ 上的算子，用以表明：如果将 $\mathbf{C}$ 替换成 $\mathbf{R}$，那么第 4 题的结果不再成立.

**6** 设 $T \in \mathcal{L}(\mathbf{F}^2)$ 定义为 $T(w,z) = (-z,w)$. 求 $T$ 的最小多项式.

**7** (a) 给出一例：$S,T \in \mathcal{L}(\mathbf{F}^2)$，使得 $ST$ 的最小多项式不等于 $TS$ 的最小多项式.

(b) 设 $V$ 是有限维的，且 $S,T \in \mathcal{L}(V)$. 证明：如果 $S,T$ 当中至少有一个可逆，那么 $ST$ 的最小多项式等于 $TS$ 的最小多项式.

提示　证明，如果 $S$ 可逆且 $p \in \mathcal{P}(\mathbf{F})$ 就有 $p(TS) = S^{-1}p(ST)S$.

**8** 设 $T \in \mathcal{L}(\mathbf{R}^2)$ 是"逆时针旋转 1°"这一算子. 求 $T$ 的最小多项式.

注　因为 $\dim \mathbf{R}^2 = 2$，所以 $T$ 的最小多项式次数最高为 2. 从而，$T$ 的最小多项式并不是 $x^{180}+1$，尽管你可能因为 $T^{180} = -I$ 而不禁这么想.

**9** 设 $T \in \mathcal{L}(V)$，使得关于 $V$ 的某个基，$T$ 的矩阵所有元素都是有理数. 解释为什么 $T$ 的最小多项式所有系数都是有理数.

**10** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $v \in V$. 证明：

$$\operatorname{span}(v, Tv, \dots, T^m v) = \operatorname{span}(v, Tv, \dots, T^{\dim V - 1} v)$$

对所有整数 $m \geq \dim V - 1$ 成立.

**11** 设 $V$ 是二维向量空间，$T \in \mathcal{L}(V)$，且 $T$ 关于 $V$ 的某个基的矩阵是 $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

(a) 证明 $T^2 - (a+d)T + (ad-bc)I = 0$.

(b) 证明 $T$ 的最小多项式等于

$$\begin{cases} z - a, & \text{若 } b = c = 0 \text{ 且 } a = d, \\ z^2 - (a+d)z + (ad-bc), & \text{其他.} \end{cases}$$

**12** 定义 $T \in \mathcal{L}(\mathbf{F}^n)$ 为 $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$. 求 $T$ 的最小多项式.

**13** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，$p \in \mathcal{P}(\mathbf{F})$. 证明：存在唯一的 $r \in \mathcal{P}(\mathbf{F})$ 使得 $p(T) = r(T)$ 且 $\deg r$ 小于 $T$ 的最小多项式的次数.

**14** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$ 有最小多项式 $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. 求 $T^{-1}$ 的最小多项式.

**15** 设 $V$ 是有限维复向量空间（$\dim V > 0$），且 $T \in \mathcal{L}(V)$. 定义 $f : \mathbf{C} \to \mathbf{R}$ 为

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

证明 $f$ 不是连续函数.

**16** 设 $a_0, \dots, a_{n-1} \in \mathbf{F}$. 令 $T$ 为 $\mathbf{F}^n$ 上的算子，其矩阵（关于标准基）为

$$\begin{pmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & \ddots & & & -a_2 \\ & & \ddots & & & \vdots \\ & & & & 0 & -a_{n-2} \\ & & & & 1 & -a_{n-1} \end{pmatrix}.$$

此处的矩阵中，除对角线下方那条线（其中都是 1）以及最后一列（其中有些也可能是 0）

以外，所有元素都是 0. 证明：$T$ 的最小多项式是

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

注 以上矩阵被称为以上多项式的**友矩阵（companion matrix）**. 这道习题表明了每个首一多项式都是某个算子的最小多项式. 因而一个公式或算法，要是能够得出每个 $\mathbf{F}^n$ 上的每个算子确切的特征值，就能够得出每个多项式确切的零点【根据 5.27 (a)】. 所以是没有这样的公式或算法的. 不过，有一些高效的数值方法，能够很好地逼近算子的特征值.

**17** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $p$ 是 $T$ 的最小多项式. 设 $\lambda \in \mathbf{F}$，证明：$T - \lambda I$ 的最小多项式是定义为 $q(z) = p(z + \lambda)$ 的多项式 $q$.

**18** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $p$ 是 $T$ 的最小多项式. 设 $\lambda \in \mathbf{F} \backslash \{0\}$，证明：$\lambda T$ 的最小多项式是定义为 $q(z) = \lambda^{\deg p} p\left(\frac{z}{\lambda}\right)$ 的多项式 $q$.

**19** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 令 $\mathcal{E}$ 为 $\mathcal{L}(V)$ 的子空间，定义为：

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

证明：$\dim \mathcal{E}$ 等于 $T$ 的最小多项式的次数.

**20** 设 $T \in \mathcal{L}(\mathbf{F}^4)$，其特征值为 3、5 和 8. 证明 $(T - 3I)^2 (T - 5I)^2 (T - 8I)^2 = 0$.

**21** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明：$T$ 的最小多项式次数最高为 $1 + \dim \operatorname{range} T$.

注 如果 $\dim \operatorname{range} T < \dim V - 1$，那么关于 $T$ 的最小多项式的次数，这道习题比起 5.22 给出了更好的上界.

**22** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明：$T$ 可逆，当且仅当 $I \in \operatorname{span}(T, T^2, \ldots, T^{\dim V})$.

**23** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 令 $n = \dim V$. 证明：如果 $v \in V$，那么 $\operatorname{span}(v, Tv, \ldots, T^{n-1} v)$ 在 $T$ 下不变.

**24** 设 $V$ 是有限维复向量空间. 设 $T \in \mathcal{L}(V)$，使得 5 和 6 是 $T$ 的特征值且 $T$ 没有其他特征值. 证明 $(T - 5I)^{\dim V - 1} (T - 6I)^{\dim V - 1} = 0$.

**25** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $U$ 是 $V$ 的在 $T$ 下不变的子空间.

(a) 证明：$T$ 的最小多项式是商算子 $T/U$ 的最小多项式的多项式倍.

(b) 证明：

$$(T|_U \text{ 的最小多项式}) \times (T/U \text{ 的最小多项式})$$

是 $T$ 的最小多项式的多项式倍.

注 商算子 $T/U$ 已在 5A 节的习题 38 中定义.

**26** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $U$ 是 $V$ 的在 $T$ 下不变的子空间. 证明：$T$ 的特征值所成的集合，等于 $T|_U$ 的特征值所成集合与 $T/U$ 的特征值所成集合的并集.

**27** 设 $\mathbf{F} = \mathbf{R}$，$V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明：$T_{\mathbf{C}}$ 的最小多项式等于 $T$ 的最小多项式.

注 复化 $T_{\mathbf{C}}$ 已在 3B 节习题 33 中定义.

**28** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明：$T' \in \mathcal{L}(V')$ 的最小多项式等于 $T$ 的最小多项式.

注 对偶映射 $T'$ 已在 3F 节中定义.

**29** 证明：在维数至少为二的有限维向量空间上，每个算子都有二维的不变子空间.

注 5C 节的习题 6 将在 $\mathbf{F} = \mathbf{C}$ 条件下加强该结果.

# 5C 上三角矩阵

在第 3 章中，我们定义了从一个有限维向量空间到另一个有限维向量空间的线性映射的矩阵. 该矩阵依赖于这两个向量空间的基的选取. 我们现在研究将一向量空间映射至自身的算子，需要着重考虑的就是仅用一个基来描述它.

**5.35 定义：算子的矩阵（matrix of an operator）、$\mathcal{M}(T)$**

设 $T \in \mathcal{L}(V)$. $T$ 关于 $V$ 的基 $v_1, \dots, v_n$ 的**矩阵**是 $n \times n$ 矩阵

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix},$$

其中的元素 $A_{j,k}$ 定义为

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

若根据上下文无法明确看出选取哪个基，就用 $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ 这个记号.

算子的矩阵是方阵（其行数等于列数），与前面我们讨论的一般线性映射的长方形矩阵不同.

> 矩阵 $\mathcal{M}(T)$ 的第 $k$ 列，是由将 $Tv_k$ 写成基 $v_1, \dots, v_n$ 的线性组合时所用的系数构成的.

若 $T$ 是 $\mathbf{F}^n$ 上的算子，且未明确基的选取，那么就假定所取的基是标准基（其中第 $k$ 个基向量除第 $k$ 个坐标为 1 外，其余坐标均为 0）. 此时你可以认为 $\mathcal{M}(T)$ 的第 $k$ 列是将 $T$ 作用于第 $k$ 个基向量所得的结果（这里把 $\mathbf{F}^n$ 中的元素与 $n \times 1$ 列向量等同起来看）.

**5.36 例：一算子关于标准基的矩阵**

定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 为 $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. 那么 $T$ 关于 $\mathbf{F}^3$ 的标准基的矩阵为

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

你应自行验证.

线性代数的中心目标之一，就是证明对于有限维向量空间 $V$ 上的算子 $T$，存在 $V$ 的一个基，使得 $T$ 关于该基有个相当简单的矩阵. 这个说法比较含糊，说得再确切些，就是我们尝试选取 $V$ 的一个基，使得 $\mathcal{M}(T)$ 有很多 0.

若 $V$ 是有限维复向量空间，那么我们凭已学过的知识，足以证明存在 $V$ 的一个基，使得 $T$ 关于该基的矩阵的第一列中，除了第一个元素可能非零外，其余元素均是 0. 换言之，存在

$V$ 的一个基，使得 $T$ 关于该基的矩阵形如

$$\begin{pmatrix} \lambda & & & \\ 0 & & * & \\ \vdots & & & \\ 0 & & & \end{pmatrix},$$

此处 $*$ 代表除了第一列以外的所有元素. 为证明以上命题，令 $\lambda$ 是 $T$ 的一个特征值（由 5.19 可知其存在性）并令 $v$ 是对应于它的特征向量. 将 $v$ 扩充为 $V$ 的一个基，那么 $T$ 关于该基的矩阵就有上述形式. 很快我们就会看到，可以选取 $V$ 的一个基，使 $T$ 关于该基的矩阵有更多的 0.

**5.37 定义：矩阵的对角线（diagonal of a matrix）**

方阵的**对角线**由从它的左上角到右下角的直线上的元素所构成. ♣

例如，例 5.36 中的矩阵

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

的对角线是由 2、5 和 8 所构成，在上式中以红色标出.

**5.38 定义：上三角矩阵（upper-triangular matrix）**

称一个方阵为**上三角矩阵**，若其中所有在对角线之下的元素都是 0. ♣

例如，上面所示的 $3 \times 3$ 矩阵就是上三角矩阵.

我们一般将上三角矩阵表示成下面的形式

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

上述矩阵中的 0 表示在这个 $n \times n$ 矩阵中，在对角线之下的元素都等于 0. 可以认为上三角矩阵是相当简单的——若 $n$ 很大，那么在 $n \times n$ 上三角矩阵中至少有将近一半的元素都是 0.

我们常用 $*$ 来表示矩阵中那些我们未知的，或与正在讨论的问题无关的元素.

下面结论在上三角矩阵和不变子空间之间建立了有用的联系.

**5.39 上三角矩阵的条件**

设 $T \in \mathcal{L}(V)$ 且 $v_1, \dots, v_n$ 是 $V$ 的基. 那么下面几条结论等价.

(a) $T$ 关于 $v_1, \dots, v_n$ 的矩阵是上三角矩阵.

(b) 对每个 $k = 1, \dots, n$，均有 $\operatorname{span}(v_1, \dots, v_k)$ 在 $T$ 下不变.

(c) 对每个 $k = 1, \dots, n$，均有 $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$. ♡

**证明** 先设 (a) 成立. 为证明 (b) 成立，设 $k \in \{1, \dots, n\}$. 若 $j \in \{1, \dots, n\}$，那么

$$Tv_j \in \operatorname{span}(v_1, \dots, v_j),$$

这是因为 $T$ 关于 $v_1,\ldots,v_n$ 的矩阵是上三角矩阵. 因为 $\operatorname{span}(v_1,\ldots,v_j)\subseteq\operatorname{span}(v_1,\ldots,v_k)$ ( $j\leq k$ ), 所以我们可知, 对每个 $j\in\{1,\ldots,k\}$, 有

$$Tv_j\in\operatorname{span}(v_1,\ldots,v_k).$$

于是, $\operatorname{span}(v_1,\ldots,v_k)$ 在 $T$ 下不变, 这就证明了 (a) 蕴涵 (b).

现在设 (b) 成立, 因此对每个 $k=1,\ldots,n$, 均有 $\operatorname{span}(v_1,\ldots,v_k)$ 在 $T$ 下不变. 特别地, 对每个 $k=1,\ldots,n$, 均有 $Tv_k\in\operatorname{span}(v_1,\ldots,v_k)$. 于是 (b) 蕴涵 (c).

现在设 (c) 成立, 因此对每个 $k=1,\ldots,n$, 均有 $Tv_k\in\operatorname{span}(v_1,\ldots,v_k)$. 这就意味着, 当用基向量 $v_1,\ldots,v_n$ 的线性组合来表达 $Tv_k$ 时, 我们只需用到向量 $v_1,\ldots,v_k$. 因此, $\mathcal{M}(T)$ 对角线之下的所有元素都是 0. 于是, $\mathcal{M}(T)$ 是个上三角矩阵, 这就证明了 (c) 蕴涵 (a).

我们证明了 (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (a), 也就证明了 (a)、(b) 和 (c) 是等价的. ■

下面结论告诉我们, 若 $T\in\mathcal{L}(V)$, 且关于 $V$ 的某个基的矩阵是

$$\mathcal{M}(T)=\begin{pmatrix}\lambda_1 & & *\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$$

那么 $T$ 满足一个依赖于 $\lambda_1,\ldots,\lambda_n$ 的简单等式.

**5.40 具有上三角矩阵的算子满足的等式**

设 $T\in\mathcal{L}(V)$ 且存在 $V$ 的一个基, 使得 $T$ 关于该基有上三角矩阵, 且该矩阵的对角线元素是 $\lambda_1,\ldots,\lambda_n$, 那么

$$(T-\lambda_1 I)\cdots(T-\lambda_n I)=0.$$

**证明** 令 $v_1,\ldots,v_n$ 是 $V$ 的基, 且 $T$ 关于该基有对角线元素为 $\lambda_1,\ldots,\lambda_n$ 的上三角矩阵. 那么 $Tv_1=\lambda_1 v_1$, 这意味着 $(T-\lambda_1 I)v_1=0$, 由此可推出 $(T-\lambda_1 I)\cdots(T-\lambda_m I)v_1=0$ 对于 $m=1,\ldots,n$ 成立 ( 利用各 $T-\lambda_j I$ 和各 $T-\lambda_k I$ 的可交换性可得 ).

注意到 $(T-\lambda_2 I)v_2\in\operatorname{span}(v_1)$. 于是 $(T-\lambda_1 I)(T-\lambda_2 I)v_2=0$ ( 根据上段可得 ), 并可由此推出 $(T-\lambda_1 I)\cdots(T-\lambda_m I)v_2=0$ 对于 $m=2,\ldots,n$ 成立 ( 利用各 $T-\lambda_j I$ 和各 $T-\lambda_k I$ 的可交换性可得 ).

注意到 $(T-\lambda_3 I)v_3\in\operatorname{span}(v_1,v_2)$. 于是由上段可得, $(T-\lambda_1 I)(T-\lambda_2 I)(T-\lambda_3 I)v_3=0$, 并可由此推出 $(T-\lambda_1 I)\cdots(T-\lambda_m I)v_3=0$ 对于 $m=3,\ldots,n$ 成立 ( 利用各 $T-\lambda_j I$ 和各 $T-\lambda_k I$ 的可交换性可得 ).

按照这个方式继续下去, 我们就会发现对每个 $k=1,\ldots,n$ 都有 $(T-\lambda_1 I)\cdots(T-\lambda_n I)v_k=0$. 于是 $(T-\lambda_1 I)\cdots(T-\lambda_n I)$ 是 0 算子, 因为它把 $V$ 的一个基中所有向量都对应至 0. ■

由算子的矩阵确切计算出算子的特征值的方法并不存在, 这实为憾事. 然而, 如果我们有幸找到一个基, 使得算子关于该基的矩阵是上三角的, 那么特征值的计算问题就变得很平凡, 如下面结论所示.

**5.41 由上三角矩阵确定特征值**

设 $T\in\mathcal{L}(V)$ 关于 $V$ 的某个基有上三角矩阵. 那么 $T$ 的特征值恰为该上三角矩阵对角线上的各元素.

**证明** 设 $v_1, \ldots, v_n$ 是 $V$ 的基，且 $T$ 关于该基有上三角矩阵

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

因为 $Tv_1 = \lambda_1 v_1$，我们可见 $\lambda_1$ 是 $T$ 的一个特征值.

设 $k \in \{2, \ldots, n\}$. 那么 $(T - \lambda_k I)v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$. 于是 $T - \lambda_k I$ 将 $\operatorname{span}(v_1, \ldots, v_k)$ 映射至 $\operatorname{span}(v_1, \ldots, v_{k-1})$. 又因为

$$\dim \operatorname{span}(v_1, \ldots, v_k) = k \quad 且 \quad \dim \operatorname{span}(v_1, \ldots, v_{k-1}) = k - 1,$$

所以限制于 $\operatorname{span}(v_1, \ldots, v_k)$ 的算子 $T - \lambda_k I$ 不是单射（由 3.22）. 于是，存在 $v \in \operatorname{span}(v_1, \ldots, v_k)$ 满足 $v \neq 0$ 且 $(T - \lambda_k I)v = 0$. 于是 $\lambda_k$ 是 $T$ 的特征值. 因此我们证明了 $\mathcal{M}(T)$ 对角线上各元素均为 $T$ 的特征值.

为证明 $T$ 无其他特征值，令 $q$ 是定义为 $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$ 的多项式. 那么 $q(T) = 0$（由 5.40）. 因此 $q$ 是 $T$ 的最小多项式的多项式倍（由 5.29）. 于是 $T$ 的最小多项式的每个零点都是 $q$ 的零点. 因为 $T$ 的最小多项式的零点即 $T$ 的特征值（由 5.27），所以可推出 $T$ 的每个特征值都是 $q$ 的零点. 因此，$T$ 的特征值都包含在 $\lambda_1, \ldots, \lambda_n$ 这组数中. ■

---

**5.42 例：由上三角矩阵得特征值**

定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 为 $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. $T$ 关于标准基的矩阵为

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

由 5.41 可得 $T$ 的特征值为 2、5 和 8.

---

下面例子形象阐释了 5.44：一算子关于某个基有上三角矩阵，当且仅当该算子的最小多项式是一次多项式的乘积.

---

**5.43 例：$T$ 是否有上三角矩阵可能取决于 F**

定义 $T \in \mathcal{L}(\mathbf{F}^4)$ 为

$$T(z_1, z_2, z_3, z_4) = (-z_2, z_1, 2z_1 + 3z_3, z_3 + 3z_4).$$

于是 $T$ 关于 $\mathbf{F}^4$ 的标准基的矩阵是

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

$T$ 的最小多项式是定义如下的多项式 $p$：

$$p(z) = 9 - 6z + 10z^2 - 6z^3 + z^4.$$

转下页

你可以利用计算机来验证一下.

首先考虑 $\mathbf{F}=\mathbf{R}$ 的情况. 那么多项式 $p$ 可因式分解成

$$p(z)=(z^2+1)(z-3)(z-3),$$

其中 $z^2+1$ 无法被进一步分解成两个一次实系数多项式的乘积. 于是 5.44 表明, 不存在 $\mathbf{R}^4$ 的一个基, 使得 $T$ 关于该基有上三角矩阵.

现在考虑 $\mathbf{F}=\mathbf{C}$ 的情况. 那么多项式 $p$ 可因式分解成

$$p(z)=(z-\mathrm{i})(z+\mathrm{i})(z-3)(z-3),$$

其中各因子都具有 $z-\lambda_k$ 的形式. 于是 5.44 表明, 存在 $\mathbf{C}^4$ 的一个基, 使得 $T$ 关于该基有上三角矩阵. 的确如此, 你可以验证一下, 算子 $T$ 关于 $\mathbf{C}^4$ 的基 $(4-3\mathrm{i},-3-4\mathrm{i},-3+\mathrm{i},1),(4+3\mathrm{i},-3+4\mathrm{i},-3-\mathrm{i},1),(0,0,0,1),(0,0,1,0)$ 有上三角矩阵

$$\begin{pmatrix} \mathrm{i} & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

---

**5.44 存在上三角矩阵的充要条件**

设 $V$ 是有限维的, $T\in\mathcal{L}(V)$. 那么 $T$ 关于 $V$ 的某个基有上三角矩阵, 当且仅当 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$ (其中 $\lambda_1,\ldots,\lambda_m\in\mathbf{F}$).

**证明** 首先设 $T$ 关于 $V$ 的某个基具有上三角矩阵. 令 $\alpha_1,\ldots,\alpha_n$ 代表该矩阵对角线上各元素. 定义多项式 $q\in\mathcal{P}(\mathbf{F})$ 为

$$q(z)=(z-\alpha_1)\cdots(z-\alpha_n).$$

那么由 5.40 得 $q(T)=0$. 因此由 5.29 得, $q$ 是 $T$ 的最小多项式的多项式倍. 于是, $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$, 其中 $\lambda_1,\ldots,\lambda_m\in\mathbf{F}$ 且 $\{\lambda_1,\ldots,\lambda_m\}\subseteq\{\alpha_1,\ldots,\alpha_n\}$.

为了证明另一方向的蕴涵关系, 现在设 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$ (其中 $\lambda_1,\ldots,\lambda_m\in\mathbf{F}$). 我们对 $m$ 用归纳法. 首先, 若 $m=1$, 则 $z-\lambda_1$ 是 $T$ 的最小多项式, 由此可得 $T=\lambda_1 I$, 进而可推出 $T$ (关于 $V$ 的任何基) 的矩阵都是上三角的.

现设 $m>1$ 且欲证结论当 $m$ 为更小的正整数值时都成立. 令

$$U=\operatorname{range}(T-\lambda_m I).$$

那么 $U$ 在 $T$ 下是不变的【这是 5.18 在 $p(z)=z-\lambda_m$ 时的特殊情况】, 那么 $T|_U$ 就是 $U$ 上的算子.

若 $u\in U$, 那么对某 $v\in V$ 有 $u=(T-\lambda_m I)v$, 并且

$$(T-\lambda_1 I)\cdots(T-\lambda_{m-1}I)u=(T-\lambda_1 I)\cdots(T-\lambda_m I)v=0.$$

因此由 5.29, $(z-\lambda_1)\cdots(z-\lambda_{m-1})$ 就是 $T|_U$ 的最小多项式的多项式倍. 于是 $T|_U$ 的最小多项式就是至多 $m-1$ 个形如 $z-\lambda_k$ 的项的乘积.

由归纳假设，存在 $U$ 的基 $u_1,\ldots,u_M$，使得 $T|_U$ 关于该基有上三角矩阵．于是对于每个 $k\in\{1,\ldots,M\}$，利用 5.39，我们有

$$Tu_k=(T|_U)(u_k)\in\operatorname{span}(u_1,\ldots,u_k). \tag{5.45}$$

将 $u_1,\ldots,u_M$ 扩充为 $V$ 的基 $u_1,\ldots,u_M,v_1,\ldots,v_N$．对每个 $k\in\{1,\ldots,N\}$，我们有

$$Tv_k=(T-\lambda_m I)v_k+\lambda_m v_k.$$

由 $U$ 的定义可知，$(T-\lambda_m I)v_k\in U=\operatorname{span}(u_1,\ldots,u_M)$．于是上式表明

$$Tv_k\in\operatorname{span}(u_1,\ldots,u_M,v_1,\ldots,v_k). \tag{5.46}$$

由式 (5.45) 和式 (5.46)，并利用 5.39，我们可得出结论：$T$ 关于 $V$ 的基 $u_1,\ldots,u_M,v_1,\ldots,v_N$ 有上三角矩阵，命题得证． ■

上个结果中得到的数集 $\{\lambda_1,\ldots,\lambda_m\}$ 就等于 $T$ 的特征值所成的集合（因为，由 5.27 可知 $T$ 的最小多项式的零点构成的集合就等于 $T$ 的特征值构成的集合），尽管 $\lambda_1,\ldots,\lambda_m$ 这组数中可能有重复的数.

下面结论是个很好的结论，但即便如此，在第 8 章中，我们还可将其改进得更好，参看 8.37 和 8.46.

**5.47 若 $\mathbf{F}=\mathbf{C}$，那么 $V$ 上的每个算子都有上三角矩阵**

设 $V$ 是有限维复向量空间且 $T\in\mathcal{L}(V)$．那么 $T$ 关于 $V$ 的某个基具有上三角矩阵.

**证明** 由 5.44 和代数基本定理的版本二（4.13）即可证明． ■

对于满足 $ST=TS$ 的两个算子 $S$ 和 $T$，上述结论可进一步拓展，参见 5.80．另外，上述结论还可拓展到多于两个算子的情形，可参看 5E 节的习题 9 (b).

**注意：**若算子 $T\in\mathcal{L}(V)$ 关于 $V$ 的某个基 $v_1,\ldots,v_n$ 有上三角矩阵，那么如 5.41 所示，$T$ 的特征值恰为 $\mathcal{M}(T)$ 对角线上的元素，此外，$v_1$ 是 $T$ 的一个特征向量．然而，$v_2,\ldots,v_n$ 并不一定是 $T$ 的特征向量．事实上，基向量 $v_k$ 是 $T$ 的特征向量，当且仅当 $T$ 的矩阵的第 $k$ 列中除了第 $k$ 个元素可能非零外，其余各元素都是 0.

回忆一下，在以前的课程中，你可能学过，每个矩阵都可以被化成行阶梯形矩阵．如果是对方阵作变换，那么所得的行阶梯形矩阵是个上三角矩阵．但不要把这个上三角矩阵和算子关于某基的上三角矩阵（由 5.47，若 $\mathbf{F}=\mathbf{C}$ 则其一定存在）混淆起来．这两个上三角矩阵之间是没有联系的.

由算子的矩阵的行阶梯形并不能得到该算子的特征值．与之不同的是，算子关于某个基的上三角矩阵则可以给出该算子的所有特征值．然而，这样的上三角矩阵是无法确切计算得到的，即便 5.47 确保了 $\mathbf{F}=\mathbf{C}$ 时它一定存在.

## 习题 5C

**1** 证明或给出一反例：如果 $T\in\mathcal{L}(V)$ 且 $T^2$ 关于 $V$ 的某个基有上三角矩阵，那么 $T$ 关于 $V$ 的某个基有上三角矩阵.

**2** 设 $A$ 和 $B$ 是大小相同的上三角矩阵，$A$ 的对角线上是 $\alpha_1,\dots,\alpha_n$，$B$ 的对角线上是 $\beta_1,\dots,\beta_n$.

(a) 证明：$A+B$ 是上三角矩阵，其对角线上是 $\alpha_1+\beta_1,\dots,\alpha_n+\beta_n$.

(b) 证明：$AB$ 是上三角矩阵，其对角线上是 $\alpha_1\beta_1,\dots,\alpha_n\beta_n$.

注 这道习题的结果会在 5.81 的证明中用到.

**3** 设 $T\in\mathcal{L}(V)$ 可逆，且 $V$ 的基 $v_1,\dots,v_n$ 使得 $T$ 关于这个基的矩阵是上三角的，其对角线上是 $\lambda_1,\dots,\lambda_n$. 证明：$T^{-1}$ 关于这个基 $v_1,\dots,v_n$ 的矩阵也是上三角的，其对角线上是

$$\frac{1}{\lambda_1},\dots,\frac{1}{\lambda_n}.$$

**4** 给出一例：一个算子，关于某个基的矩阵对角线上只有 0，但是可逆.

注 本题和下一题表明，如果不假设所考虑的是上三角矩阵，5.41 就不成立了.

**5** 给出一例：一个算子，关于某个基的矩阵对角线上只有非零数，但是不可逆.

**6** 设 $\mathbf{F}=\mathbf{C}$，$V$ 是有限维的，且 $T\in\mathcal{L}(V)$. 证明：如果 $k\in\{1,\dots,\dim V\}$，那么 $V$ 有在 $T$ 下不变的 $k$ 维子空间.

**7** 设 $V$ 是有限维的，$T\in\mathcal{L}(V)$ 且 $v\in V$.

(a) 证明：存在唯一的最低次首一多项式 $p_v$ 使得 $p_v(T)v=0$.

(b) 证明：$T$ 的最小多项式是 $p_v$ 的多项式倍.

**8** 设 $V$ 是有限维的，$T\in\mathcal{L}(V)$，且存在非零向量 $v\in V$ 使得 $T^2v+2Tv=-2v$.

(a) 证明：如果 $\mathbf{F}=\mathbf{R}$，那么 $V$ 中不存在能使 $T$ 有上三角矩阵的基.

(b) 证明：如果 $\mathbf{F}=\mathbf{C}$ 且 $A$ 是 $T$ 关于 $V$ 的某个基得到的上三角矩阵，那么 $A$ 的对角线上会出现 $-1+\mathrm{i}$ 或 $-1-\mathrm{i}$.

**9** 设 $B$ 是方阵，其元素为复数. 证明：存在元素为复数的可逆方阵 $A$，使得 $A^{-1}BA$ 是上三角矩阵.

**10** 设 $T\in\mathcal{L}(V)$ 且 $v_1,\dots,v_n$ 是 $V$ 的基. 证明下列是等价的：

(a) $T$ 关于 $v_1,\dots,v_n$ 的矩阵是下三角的.

(b) 对每个 $k=1,\dots,n$，均有 $\operatorname{span}(v_k,\dots,v_n)$ 在 $T$ 下不变.

(c) 对每个 $k=1,\dots,n$，均有 $Tv_k\in\operatorname{span}(v_k,\dots,v_n)$.

注 一个方阵被称为**下三角的（lower-triangular）**，如果其对角线上方的元素都为 0.

**11** 设 $\mathbf{F}=\mathbf{C}$，且 $V$ 是有限维的. 证明：如果 $T\in\mathcal{L}(V)$，那么存在 $V$ 的一个基使得 $T$ 关于该基有下三角矩阵.

**12** 设 $V$ 是有限维的，$T\in\mathcal{L}(V)$ 关于 $V$ 的某个基有上三角矩阵，且 $U$ 是 $V$ 的在 $T$ 下不变的子空间.

(a) 证明：$T|_U$ 关于 $U$ 的某个基有上三角矩阵.

(b) 证明：商算子 $T/U$ 关于 $V/U$ 的某个基有上三角矩阵.

注 商算子 $T/U$ 已在 5A 节习题 38 中定义.

**13** 设 $V$ 是有限维的，且 $T\in\mathcal{L}(V)$. 设存在 $V$ 的在 $T$ 下不变的子空间 $U$，使得 $T|_U$ 关于 $U$ 的某个基有上三角矩阵，$T/U$ 关于 $V/U$ 的某个基也有上三角矩阵. 证明：$T$ 关于 $V$ 的某个基有上三角矩阵.

**14** 设 $V$ 是有限维的，且 $T\in\mathcal{L}(V)$. 证明：$T$ 关于 $V$ 的某个基有上三角矩阵，当且仅当对偶算子 $T'$ 关于对偶空间 $V'$ 的某个基有上三角矩阵.

# 5D 可对角化算子

## 对角矩阵

**5.48 定义：对角矩阵（diagonal matrix）**

**对角矩阵**是对角线之外元素均为 0 的方阵.

---

**5.49 例：对角矩阵**

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

是对角矩阵.

---

若一算子关于某基具有对角矩阵，那么该矩阵对角线上的元素就是该算子的特征值. 由 5.41 即可证明这个结论（也可专为对角矩阵构造个更简单直接的证法）.

> 每个对角矩阵都是上三角的. 一般而言，对角矩阵比大多数相同大小的上三角矩阵含有更多的 0.

**5.50 定义：可对角化（diagonalizable）**

若 $V$ 上的算子关于 $V$ 的某个基具有对角矩阵，则称该算子是**可对角化的**.

---

**5.51 例：对角化可能需要不同的基**

定义 $T \in \mathcal{L}(\mathbf{R}^2)$ 为

$$T(x, y) = (41x + 7y, -20x + 74y).$$

$T$ 关于 $\mathbf{R}^2$ 的标准基的矩阵是

$$\begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix},$$

这不是对角矩阵. 然而，$T$ 是可对角化的. 具体而言，$T$ 关于基 $(1,4), (7,5)$ 的矩阵是

$$\begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix},$$

因为 $T(1,4) = (69,276) = 69(1,4)$，$T(7,5) = (322,230) = 46(7,5)$.

---

下面，我们给被算子 $T$ 映射至其 $\lambda$（$\lambda \in \mathbf{F}$）倍的向量所构成的集合起个名字，并用个记号来表示它，这将便于我们进行讨论.

**5.52 定义：特征空间（eigenspace）、$E(\lambda, T)$**

设 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$. $T$ 对应于 $\lambda$ 的**特征空间**记作 $E(\lambda, T)$，是定义如下的 $V$ 的子空间

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

因此 $E(\lambda, T)$ 是 $T$ 对应于 $\lambda$ 的所有特征向量以及向量 0 所构成的集合.

对 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$，集合 $E(\lambda, T)$ 是 $V$ 的子空间，因为 $V$ 上任何线性映射的零空间都是 $V$ 的子空间. 由定义可得，$\lambda$ 是 $T$ 的特征值当且仅当 $E(\lambda, T) \neq \{0\}$.

**5.53 例：一算子的特征空间**

设算子 $T \in \mathcal{L}(V)$ 关于 $V$ 的基 $v_1, v_2, v_3$ 的矩阵是例 5.49 中的矩阵. 那么

$$E(8, T) = \operatorname{span}(v_1), \quad E(5, T) = \operatorname{span}(v_2, v_3).$$

若 $\lambda$ 是算子 $T \in \mathcal{L}(V)$ 的特征值，那么限制于 $E(\lambda, T)$ 的 $T$ 就是将向量乘以 $\lambda$ 的算子.

**5.54 特征空间之和是直和**

设 $T \in \mathcal{L}(V)$ 且 $\lambda_1, \dots, \lambda_m$ 是 $T$ 的互异特征值. 那么

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

是直和. 此外，若 $V$ 是有限维的，那么

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V.$$

**证明** 为证明 $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ 是直和，设

$$v_1 + \cdots + v_m = 0,$$

其中各 $v_k$ 分别属于 $E(\lambda_k, T)$. 因为对应于互异特征值的特征向量是线性无关的（由 5.11），故可得各 $v_k$ 都等于 0. 因此，$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ 是直和（由 1.45）.

现设 $V$ 是有限维的. 那么

$$\begin{aligned} \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim\bigl(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)\bigr) \\ &\leq \dim V, \end{aligned}$$

其中第一行源于 3.94 而第二行源于 2.37. ■

## 可对角化的条件

对于可对角化算子的如下刻画很有用处.

**5.55 可对角化的等价条件**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 令 $\lambda_1, \dots, \lambda_m$ 表示 $T$ 的所有互异特征值. 则下列命题等价.

(a) $T$ 是可对角化的.

(b) $V$ 有由 $T$ 的特征向量构成的基.

(c) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

(d) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

**证明** 算子 $T \in \mathcal{L}(V)$ 关于 $V$ 的某个基 $v_1, \ldots, v_n$ 具有对角矩阵

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

当且仅当对每个 $k$ 均有 $Tv_k = \lambda_k v_k$. 于是 (a) 与 (b) 等价.

设 (b) 成立，于是 $V$ 有由 $T$ 的特征向量构成的基. 因此 $V$ 中每个向量都是 $T$ 的特征向量的线性组合，而这意味着

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T).$$

将此式结合 5.54 可知 (c) 成立，也就证明了 (b) 蕴涵 (c).

由 3.94 立得 (c) 蕴涵 (d).

最后，设 (d) 成立，于是

$$\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T). \tag{5.56}$$

为各 $E(\lambda_k, T)$ 选取一个基，再将这些基合在一起，就形成由 $T$ 的特征向量构成的组 $v_1, \ldots, v_n$，其中 $n = \dim V$【由式 (5.56)】. 为证明这个组是线性无关的，设

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

其中 $a_1, \ldots, a_n \in \mathbf{F}$. 对每个 $k = 1, \ldots, m$，令 $u_k$ 表示所有 $a_j v_j$ 之和，其中各 $v_j \in E(\lambda_k, T)$. 于是各 $u_k$ 分别属于 $E(\lambda_k, T)$，且

$$u_1 + \cdots + u_m = 0.$$

因为对应于互异特征值的特征向量是线性无关的（见 5.11），故可得各 $u_k$ 都等于 0. 因为各 $u_k$ 是由若干 $a_j v_j$ 求和得到，其中这些 $v_j$ 构成 $E(\lambda_k, T)$ 的基，故可知各 $a_j$ 都等于 0. 于是 $v_1, \ldots, v_n$ 是线性无关的，因此是 $V$ 的基（由 2.38）. 所以 (d) 蕴涵 (b)，证明完成. ■

与可对角化等价的其他条件见于 5.62，本节习题 5 和习题 15，7B 节的习题 24，以及 8A 节的习题 15.

我们知道，非零有限维复向量空间上的每个算子都有特征值. 然而，非零有限维复向量空间上，不是每个算子都有足够多的特征向量使其可对角化，如下例所示.

**5.57 例：不可对角化的算子**

定义算子 $T \in \mathcal{L}(\mathbf{F}^3)$ 为 $T(a, b, c) = (b, c, 0)$. $T$ 关于 $\mathbf{F}^3$ 的标准基的矩阵为

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

它是上三角矩阵，但不是对角矩阵.

转下页

你应自行验证，0 是 $T$ 的唯一特征值，并且

$$E(0,T) = \{(a,0,0) \in \mathbf{F}^3 : a \in \mathbf{F}\}.$$

因此 5.55 中 (b)、(c) 和 (d) 这三条都不成立（当然，由于这几条是等价的，所以只要检验其中一条不成立就够了）. 于是 5.55 中的 (a) 这条也不成立. 因此不管 $\mathbf{F} = \mathbf{R}$ 还是 $\mathbf{F} = \mathbf{C}$，$T$ 都不可对角化.

下面结论表明，如果一个算子的互异特征值数目与其定义空间的维数一样，那么该算子就是可对角化的.

**5.58 特征值足够多意味着可对角化**

设 $V$ 是有限维的且 $T \in \mathcal{L}(V)$ 有 $\dim V$ 个互不相同的特征值. 那么 $T$ 是可对角化的.

**证明** 设 $T$ 有互不相同的特征值 $\lambda_1, \ldots, \lambda_{\dim V}$. 对每个 $k$，令 $v_k \in V$ 是对应于特征值 $\lambda_k$ 的一个特征向量. 因为对应于不同特征值的特征向量线性无关（见 5.11），所以 $v_1, \ldots, v_{\dim V}$ 是线性无关的.

$V$ 中由 $\dim V$ 个向量构成的线性无关组是 $V$ 的基（见 2.38）. 由此，$v_1, \ldots, v_{\dim V}$ 是 $V$ 的基. 关于这个由特征向量构成的基，$T$ 具有对角矩阵. ■

在后面的章节中我们还会得到更多能推出某个算子可对角化的条件，例如实谱定理（7.29）和复谱定理（7.31）.

上述结论给出了算子可对角化的充分条件. 然而，这个条件不是必要的. 例如，定义为 $T(x,y,z) = (6x, 6y, 7z)$ 的 $\mathbf{F}^3$ 上的算子 $T$ 只有两个特征值（6 和 7）而 $\dim \mathbf{F}^3 = 3$，但 $T$ 也是可对角化的（用 $\mathbf{F}^3$ 的标准基即可）.

下面例子展示了对角化的重要作用，即可以借助它来计算算子的高次幂——利用式 $T^k v = \lambda^k v$（$v$ 是 $T$ 对应于特征值 $\lambda$ 的特征向量）即可.

该技巧的一个更精彩的应用示于习题 21 中，该题展示了如何利用对角化来得出斐波那契数列的第 $n$ 项的确切表达式.

**5.59 例：利用对角化来计算 $T^{100}$**

定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 为 $T(x,y,z) = (2x+y, 5y+3z, 8z)$. 那么 $T$ 关于标准基的矩阵为

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

上面矩阵是上三角矩阵，但不是对角矩阵. 由 5.41 可知 $T$ 的特征值是 2、5 和 8. 因为 $T$ 是三维向量空间上的算子且 $T$ 有三个互不相同的特征值，5.58 就确保存在 $\mathbf{F}^3$ 的一个基使得 $T$ 关于该基具有对角矩阵.

为了求出这个基，我们只需求出对应于每个特征值的特征向量即可. 换言之，我们需要求出方程

$$T(x,y,z) = \lambda(x,y,z)$$

转下页

在 $\lambda=2$、$\lambda=5$ 和 $\lambda=8$ 这三种情形下的非零解. 求解这些简单的方程即得对应于 $\lambda=2$ 的特征向量 $(1,0,0)$、对应于 $\lambda=5$ 的特征向量 $(1,3,0)$ 和对应于 $\lambda=8$ 的特征向量 $(1,6,6)$.

于是, $(1,0,0),(1,3,0),(1,6,6)$ 是由 $T$ 的特征向量构成的 $\mathbf{F}^3$ 的基, 且 $T$ 关于这个基的矩阵即为对角矩阵

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

下面以计算 $T^{100}(0,0,1)$ 为例说明如何计算 $T^{100}$. 我们先把 $(0,0,1)$ 用特征向量构成的基的线性组合表示出来:

$$(0,0,1)=\frac{1}{6}(1,0,0)-\frac{1}{3}(1,3,0)+\frac{1}{6}(1,6,6).$$

再将 $T^{100}$ 作用于上式两端, 可得

$$\begin{aligned} T^{100}(0,0,1) &= \frac{1}{6}\big(T^{100}(1,0,0)\big)-\frac{1}{3}\big(T^{100}(1,3,0)\big)+\frac{1}{6}\big(T^{100}(1,6,6)\big) \\ &= \frac{1}{6}\big(2^{100}(1,0,0)-2\cdot 5^{100}(1,3,0)+8^{100}(1,6,6)\big) \\ &= \frac{1}{6}\big(2^{100}-2\cdot 5^{100}+8^{100}, 6\cdot 8^{100}-6\cdot 5^{100}, 6\cdot 8^{100}\big). \end{aligned}$$

之前我们已经看到, 有限维向量空间 $V$ 上的算子 $T$ 关于 $V$ 的某个基有上三角矩阵, 当且仅当 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$, 其中 $\lambda_1,\ldots,\lambda_m\in\mathbf{F}$ (见 5.44). 前面也提到过 (见 5.47), 这个结论当 $\mathbf{F}=\mathbf{C}$ 时总是成立的.

接下来的结果 5.62 的内容是, 算子 $T\in\mathcal{L}(V)$ 关于 $V$ 的某个基有对角矩阵, 当且仅当 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$, 其中 $\lambda_1,\ldots,\lambda_m\in\mathbf{F}$ 且**互不相同**. 在正式地陈述这个结论之前, 我们先给出两个应用它的实例.

**5.60 例: 可对角化, 但无法求出确切特征值**

定义 $T\in\mathcal{L}(\mathbf{C}^5)$ 为

$$T(z_1,z_2,z_3,z_4,z_5)=(-3z_5, z_1+6z_5, z_2, z_3, z_4).$$

$T$ 的矩阵如例 5.26 所示, 在此例中, 我们求得 $T$ 的最小多项式为 $3-6z+z^5$.

在例 5.28 中曾提及, 无法求出上述多项式的零点的确切表达式, 但是用数值计算方法可得该多项式的零点近似为 $-1.67, 0.51, 1.40, -0.12+1.59i, -0.12-1.59i$.

算出这些近似值的软件可以给出不止三位数字的精确值. 于是, 上述近似结果足以显示五个特征值是互异的. 而 $T$ 的最小多项式就是将以上各数作为零点的五次首一多项式. 则由 5.62 可知 $T$ 是可对角化的.

**5.61 例: 说明一算子不可对角化**

定义 $T\in\mathcal{L}(\mathbf{F}^3)$ 为

$$T(z_1,z_2,z_3)=(6z_1+3z_2+4z_3, 6z_2+2z_3, 7z_3).$$

转下页

$T$ 关于 $\mathbf{F}^3$ 的标准基的矩阵是

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

上述矩阵是上三角矩阵，但不是对角矩阵．$T$ 关于 $\mathbf{F}^3$ 的其他基是否可能具有对角矩阵呢？

为了回答这个问题，我们需要求出 $T$ 的最小多项式．首先，注意到 $T$ 的特征值是上述矩阵对角线上的项（由 5.41）．于是 $T$ 的最小多项式的零点是 6 和 7【由 5.27 (a)】．由上述矩阵的对角线又可得 $(T-6I)^2(T-7I)=0$（由 5.40）．$T$ 的最小多项式的次数最高为 3（由 5.22）．综合所有这些结论，我们即得 $T$ 的最小多项式要么是 $(z-6)(z-7)$，要么是 $(z-6)^2(z-7)$.

简单计算即知 $(T-6I)(T-7I)\neq 0$．因而 $T$ 的最小多项式是 $(z-6)^2(z-7)$.

则由 5.62 可知 $T$ 不可对角化.

**5.62　可对角化的充要条件**

设 $V$ 是有限维的，且 $T\in\mathcal{L}(V)$．那么 $T$ 是可对角化的，当且仅当 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$，其中 $\lambda_1,\dots,\lambda_m\in\mathbf{F}$ 为互不相同的数.

**证明**　首先设 $T$ 是可对角化的. 那么存在由 $T$ 的特征向量构成的 $V$ 的基 $v_1,\dots,v_n$. 令 $\lambda_1,\dots,\lambda_m$ 是 $T$ 的所有互异特征值．那么对每个 $v_j$，存在 $\lambda_k$ 使得 $(T-\lambda_k I)v_j=0$．于是

$$(T-\lambda_1 I)\cdots(T-\lambda_m I)v_j=0,$$

这意味着 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$.

为了证明另一方向的蕴涵关系，现设 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$，其中 $\lambda_1,\dots,\lambda_m\in\mathbf{F}$ 为互不相同的数．于是

$$(T-\lambda_1 I)\cdots(T-\lambda_m I)=0. \tag{5.63}$$

我们将通过对 $m$ 用归纳法来证明 $T$ 是可对角化的．首先设 $m=1$．那么 $T-\lambda_1 I=0$，这意味着 $T$ 是恒等算子的标量倍，也就说明了 $T$ 是可对角化的.

现在设 $m>1$ 且欲证结论在 $m$ 为更小的值时都成立．$\operatorname{range}(T-\lambda_m I)$ 这个子空间在 $T$ 下是不变的【这是 5.18 在 $p(z)=z-\lambda_m$ 时的特殊情况】．那么限制在 $\operatorname{range}(T-\lambda_m I)$ 上的 $T$ 就是 $\operatorname{range}(T-\lambda_m I)$ 上的算子.

若 $u\in\operatorname{range}(T-\lambda_m I)$，那么对某个 $v\in V$ 有 $u=(T-\lambda_m I)v$，结合式 (5.63) 可得

$$(T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u=(T-\lambda_1 I)\cdots(T-\lambda_m I)v=0. \tag{5.64}$$

因此，$(z-\lambda_1)\cdots(z-\lambda_{m-1})$ 是限制于 $\operatorname{range}(T-\lambda_m I)$ 上的 $T$ 的最小多项式的多项式倍（由 5.29）．于是，由归纳假设，$\operatorname{range}(T-\lambda_m I)$ 中存在由 $T$ 的特征向量构成的基.

设 $u\in\operatorname{range}(T-\lambda_m I)\cap\operatorname{null}(T-\lambda_m I)$．那么 $Tu=\lambda_m u$．结合式 (5.64) 可得

$$\begin{aligned} 0 &= (T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u \\ &= (\lambda_m-\lambda_1)\cdots(\lambda_m-\lambda_{m-1})u. \end{aligned}$$

因为 $\lambda_1,\dots,\lambda_m$ 是互异的，所以上式表明 $u=0$．因此 $\operatorname{range}(T-\lambda_m I)\cap\operatorname{null}(T-\lambda_m I)=\{0\}$.

于是 $\operatorname{range}(T-\lambda_m I)+\operatorname{null}(T-\lambda_m I)$ 是直和（由 1.46），且其维数是 $\dim V$（由 3.94 和 3.21）. 因此 $\operatorname{range}(T-\lambda_m I)\oplus\operatorname{null}(T-\lambda_m I)=V$. $\operatorname{null}(T-\lambda_m I)$ 中的每个非零向量都是 $T$ 对应于特征值 $\lambda_m$ 的特征向量. 在该证明的靠前部分，我们已证明 $\operatorname{range}(T-\lambda_m I)$ 存在由 $T$ 的特征向量构成的基. 将该基与 $\operatorname{null}(T-\lambda_m I)$ 的基合并，就得到了由 $T$ 的特征向量构成的 $V$ 的基. $T$ 关于这个基的矩阵就是对角矩阵，命题得证. ■

5 次以上多项式没有求根公式. 然而，利用上述结论，连最小多项式零点的近似值都不用求，就能判断复向量空间上一算子是否可对角化，详见习题 15.

下面这条结论将是我们证明两算子的同时对角化时所运用的关键工具，详见 5.76. 由 5.62，可从最小多项式角度刻画可对角化算子，请留意我们如何利用这一点来引出下面这条结论的简短证明.

**5.65 将可对角化算子限制于不变子空间**

设 $T\in\mathcal{L}(V)$ 是可对角化的，$U$ 是 $V$ 的子空间且在 $T$ 下不变. 那么 $T|_U$ 是 $U$ 上的可对角化算子.

**证明** 因为算子 $T$ 是可对角化的，所以 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$，其中 $\lambda_1,\dots,\lambda_m\in\mathbf{F}$ 为互不相同的数（由 5.62）. $T$ 的最小多项式是 $T|_U$ 的最小多项式的多项式倍（由 5.31）. 因此 $T|_U$ 的最小多项式具有 5.62 所需的形式，由此可说明 $T|_U$ 是可对角化的. ■

## 格什戈林圆盘定理

**5.66 定义：格什戈林圆盘（Gershgorin disks）**

设 $T\in\mathcal{L}(V)$ 且 $v_1,\dots,v_n$ 是 $V$ 的基. 令 $A$ 表示 $T$ 关于该基的矩阵. $T$ 关于基 $v_1,\dots,v_n$ 的**格什戈林圆盘**是形如

$$\left\{z\in\mathbf{F}:|z-A_{j,j}|\le\sum_{\substack{k=1\\k\neq j}}^{n}|A_{j,k}|\right\}$$

的集合，其中 $j\in\{1,\dots,n\}$.

因为上面定义中的 $j$ 有 $n$ 种取值，所以 $T$ 有 $n$ 个格什戈林圆盘. 若 $\mathbf{F}=\mathbf{C}$，那么对每个 $j\in\{1,\dots,n\}$，与之对应的格什戈林圆盘就是 $\mathbf{C}$ 中以 $A_{j,j}$ 为圆心的闭圆盘，其中 $A_{j,j}$ 是 $A$ 的对角线上的第 $j$ 个元素. 这个闭圆盘的半径，等于 $A$ 的第 $j$ 行除对角线上的元素外各元素绝对值之和. 若 $\mathbf{F}=\mathbf{R}$，那么格什戈林圆盘是 $\mathbf{R}$ 中的闭区间.

在上述方阵 $A$ 为对角阵的特殊情况下，每个格什戈林圆盘都仅包含一个点，即 $A$ 的对角线上的元素（进而 $T$ 的每个特征值都是这样的点，这也是下面结论所保证的）. 下面的结论有个推论：如果 $A$ 的非对角线元素很小，那么 $T$ 的每个特征值都与 $A$ 的对角线上的一个元素很接近.

**5.67 格什戈林圆盘定理（Gershgorin disk theorem）**

设 $T\in\mathcal{L}(V)$ 且 $v_1,\dots,v_n$ 是 $V$ 的基. 那么 $T$ 的每个特征值都被包含在 $T$ 关于基 $v_1,\dots,v_n$ 的某个格什戈林圆盘中.

**证明** 设 $\lambda \in \mathbf{F}$ 是 $T$ 的特征值. 令 $w \in V$ 是与之对应的特征向量. 存在 $c_1, \ldots, c_n \in \mathbf{F}$, 使得

$$w = c_1 v_1 + \cdots + c_n v_n. \tag{5.68}$$

令 $A$ 表示 $T$ 关于基 $v_1, \ldots, v_n$ 的矩阵. 将 $T$ 作用于上式两侧, 可得

$$\begin{aligned}\lambda w &= \sum_{k=1}^{n} c_k T v_k &(5.69)\\ &= \sum_{k=1}^{n} c_k \sum_{j=1}^{n} A_{j,k} v_j\\ &= \sum_{j=1}^{n} \left( \sum_{k=1}^{n} A_{j,k} c_k \right) v_j. &(5.70)\end{aligned}$$

令 $j \in \{1, \ldots, n\}$ 使得

$$|c_j| = \max\{|c_1|, \ldots, |c_n|\}.$$

利用式 (5.68), 我们发现, 在式 (5.69) 左侧, 各 $v_j$ 的系数等于 $\lambda c_j$, 而这些系数必须与式 (5.70) 右侧中各 $v_j$ 的系数相等. 换言之,

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k} c_k.$$

在此式两侧同时减去 $A_{j,j} c_j$, 再同除以 $c_j$, 可得

$$\begin{aligned}|\lambda - A_{j,j}| &= \left| \sum_{\substack{k=1\\k\neq j}}^{n} A_{j,k} \frac{c_k}{c_j} \right|\\ &\leq \sum_{\substack{k=1\\k\neq j}}^{n} |A_{j,k}|.\end{aligned}$$

于是 $\lambda$ 处在关于基 $v_1, \ldots, v_n$ 的第 $j$ 个格什戈林圆盘中. ■

格什戈林圆盘定理是用谢苗・阿罗诺维奇・格什戈林（Semyon Aronovich Gershgorin）的名字命名的, 他于 1931 年发表了该结论.

习题 22 给出了格什戈林圆盘定理的一个很漂亮的应用实例.

习题 23 则提出, 可将每个格什戈林圆盘的半径改为等于其对应列（而不是行）中除对角线元素外各元素绝对值之和, 而格什戈林圆盘定理仍然成立.

## 习题 5D

**1** 设 $V$ 是有限维复向量空间, 且 $T \in \mathcal{L}(V)$.

(a) 证明: 如果 $T^4 = I$, 那么 $T$ 可对角化.

(b) 证明: 如果 $T^4 = T$, 那么 $T$ 可对角化.

(c) 给出一例: 算子 $T \in \mathcal{L}(\mathbf{C}^2)$, 使得 $T^4 = T^2$ 且 $T$ 不可对角化.

**2** 设 $T \in \mathcal{L}(V)$ 关于 $V$ 的一个基有对角矩阵 $A$. 证明: 如果 $\lambda \in \mathbf{F}$, 那么 $\lambda$ 在 $A$ 的对角线上恰好出现 $\dim E(\lambda, T)$ 次.

**3** 设 $V$ 是有限维的, 且 $T \in \mathcal{L}(V)$. 证明: 如果算子 $T$ 可对角化, 那么 $V = \operatorname{null} T \oplus \operatorname{range} T$.

**4** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 证明下列是等价的：

(a) $V = \operatorname{null} T \oplus \operatorname{range} T$.

(b) $V = \operatorname{null} T + \operatorname{range} T$.

(c) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

**5** 设 $V$ 是有限维复向量空间，且 $T \in \mathcal{L}(V)$. 证明：$T$ 可对角化，当且仅当

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

对任一 $\lambda \in \mathbf{C}$ 都成立.

**6** 设 $T \in \mathcal{L}(\mathbf{F}^5)$ 且 $\dim E(8, T) = 4$. 证明：$T - 2I$ 或 $T - 6I$ 可逆.

**7** 设 $T \in \mathcal{L}(V)$ 可逆. 证明：

$$E(\lambda, T) = E\left(\tfrac{1}{\lambda}, T^{-1}\right)$$

对任一 $\lambda \in \mathbf{F}$（$\lambda \neq 0$）都成立.

**8** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V)$. 令 $\lambda_1, \ldots, \lambda_m$ 表示 $T$ 的相异非零特征值. 证明：

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

**9** 设 $R, T \in \mathcal{L}(\mathbf{F}^3)$ 都有特征值 2、6 和 7. 证明：存在一可逆算子 $S \in \mathcal{L}(\mathbf{F}^3)$ 使得 $R = S^{-1}TS$.

**10** 求 $R, T \in \mathcal{L}(\mathbf{F}^4)$. 其满足：$R$ 和 $T$ 都有特征值 2、6 和 7 且没有其他特征值，同时不存在一可逆算子 $S \in \mathcal{L}(\mathbf{F}^4)$ 使得 $R = S^{-1}TS$.

**11** 求 $T \in \mathcal{L}(\mathbf{C}^3)$. 其满足：6 和 7 是 $T$ 的特征值，且 $T$ 关于 $\mathbf{C}^3$ 的任何基都没有对角矩阵.

**12** 设 $T \in \mathcal{L}(\mathbf{C}^3)$ 使得 6 和 7 是 $T$ 的特征值. 又假设 $T$ 关于 $\mathbf{C}^3$ 的任何基都没有对角矩阵. 证明：存在 $(z_1, z_2, z_3) \in \mathbf{C}^3$ 使得

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

**13** 设 $A$ 是对角线上元素各异的对角矩阵，$B$ 是和 $A$ 大小相同的矩阵. 证明：$AB = BA$ 当且仅当 $B$ 是对角矩阵.

**14** (a) 给出一例：一有限维复向量空间，以及该空间上的算子 $T$，使得 $T^2$ 可对角化但 $T$ 不可对角化.

(b) 设 $\mathbf{F} = \mathbf{C}$，$k$ 是一正整数，且 $T \in \mathcal{L}(V)$ 可逆. 证明：$T$ 可对角化当且仅当 $T^k$ 可对角化.

**15** 设 $V$ 是有限维复向量空间，$T \in \mathcal{L}(V)$，且 $p$ 是 $T$ 的最小多项式. 证明下列是等价的：

(a) $T$ 可对角化.

(b) 不存在 $\lambda \in \mathbf{C}$ 使得 $p$ 是 $(z - \lambda)^2$ 的多项式倍.

(c) $p$ 和其导数 $p'$ 没有公共零点.

(d) $p$ 和 $p'$ 的最大公因式是常多项式 1.

注 $p$ 和 $p'$ 的**最大公因式**（**greatest common divisor**）是使得 $p$ 和 $p'$ 都为其多项式倍的最高次首一多项式 $q$. 多项式的欧几里得算法（自己查一下）可以快速确定两多项式的最大公因式，而不需要关于多项式零点的任何信息. 因此以上 (a) 和 (d) 的等价关系表明，我们可在对 $p$ 的零点一无所知的情况下确定 $T$ 是否可对角化.

**16** 设 $T \in \mathcal{L}(V)$ 可对角化. 令 $\lambda_1, \ldots, \lambda_m$ 表示 $T$ 的所有互异特征值. 证明：$V$ 的子空间 $U$ 在 $T$ 下不变，当且仅当存在 $V$ 的子空间 $U_1, \ldots, U_m$ 使得 $U_k \subseteq E(\lambda_k, T)$ 对每个 $k$ 成立且 $U = U_1 \oplus \cdots \oplus U_m$.

**17** 设 $V$ 是有限维的. 证明: $\mathcal{L}(V)$ 中存在由可对角化算子构成的基.

**18** 设 $T \in \mathcal{L}(V)$ 可对角化, 且 $U$ 是 $V$ 的在 $T$ 下不变的子空间. 证明: 商算子 $T/U$ 是 $V/U$ 上的可对角化算子.

注 商算子 $T/U$ 已在 5A 节习题 38 中定义.

**19** 证明或给出一反例: 如果 $T \in \mathcal{L}(V)$ 且存在 $V$ 的在 $T$ 下不变的子空间 $U$, 使得 $T|_U$ 和 $T/U$ 都可对角化, 那么 $T$ 可对角化.

注 关于上三角矩阵的类似表述, 见 5C 节习题 13.

**20** 设 $V$ 是有限维的, 且 $T \in \mathcal{L}(V)$. 证明: $T$ 可对角化, 当且仅当对偶算子 $T'$ 可对角化.

**21** **斐波那契数列 (Fibonacci sequence)** $F_0, F_1, F_2, \ldots$ 定义为

$$F_0 = 0, F_1 = 1, \text{ 且 } F_n = F_{n-2} + F_{n-1} \text{ 对 } n \geq 2 \text{ 成立.}$$

定义 $T \in \mathcal{L}(\mathbf{R}^2)$ 为 $T(x, y) = (y, x + y)$.

(a) 证明: $T^n(0, 1) = (F_n, F_{n+1})$ 对任一非负整数 $n$ 都成立.

(b) 求 $T$ 的特征值.

(c) 求 $\mathbf{R}^2$ 的一个由 $T$ 的特征向量构成的基.

(d) 用 (c) 的求解结果计算 $T^n(0, 1)$. 得出结论:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

对任一非负整数 $n$ 都成立.

(e) 利用 (d) 得出结论: 如果 $n$ 是非负整数, 那么斐波那契数 $F_n$ 是最接近

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

的整数.

注 每个 $F_n$ 都是非负整数, 尽管 (d) 中公式右侧看起来不像整数. 其中的

$$\frac{1+\sqrt{5}}{2}$$

被称为**黄金比例 (golden ratio)**.

**22** 设 $T \in \mathcal{L}(V)$ 且 $n \times n$ 矩阵 $A$ 是 $T$ 关于 $V$ 的某个基的矩阵. 证明: 如果

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|$$

对每个 $j \in \{1, \ldots, n\}$ 成立, 那么 $T$ 可逆.

注 这道习题说明, 如果 $T$ 的矩阵对角线元素很大 (相较于非对角线元素) [4], 那么 $T$ 可逆.

**23** 假设格什戈林圆盘的定义被替换为: 第 $k$ 个圆盘的半径为 $A$ 的第 $k$ 列 (而不是行) 元素 (不含对角线上的元素) 绝对值之和. 证明格什戈林圆盘定理 (5.67) 在替换定义后仍然成立.

[4] 这种矩阵也被称为严格对角占优矩阵 (strictly diagonally dominant matrix).

## 5E 可交换算子

**5.71 定义：可交换（commute）**

- 对于同一向量空间上的两个算子 $S$ 和 $T$，若 $ST = TS$，则它们**可交换**.
- 对于两个大小相同的方阵 $A$ 和 $B$，若 $AB = BA$，则它们**可交换**.

♣

例如，若 $I$ 是 $V$ 上的恒等算子且 $\lambda \in \mathbf{F}$，那么 $\lambda I$ 与 $V$ 上每个算子都可交换.

又例如，若 $T$ 是算子，那么 $T^2$ 和 $T^3$ 可交换. 更一般地，若 $p, q \in \mathcal{P}(\mathbf{F})$，那么 $p(T)$ 和 $q(T)$ 可交换【见 5.17 (b)】.

---

**5.72 例：偏微分算子可交换**

设 $m$ 是非负整数. 令 $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ 表示具有两个自变量且次数最高为 $m$ 的复系数（系数属于 $\mathbf{C}$ 的）多项式构成的复向量空间，其带有复值（值属于 $\mathbf{C}$ 的）函数的一般加法和标量乘法运算. 于是 $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ 中的元素，是从 $\mathbf{C}^2$ 到 $\mathbf{C}$ 的形式如下的函数 $p$：

$$p(w, z) = \sum_{j+k \le m} a_{j,k} w^j z^k, \tag{5.73}$$

其中下标 $j$ 和 $k$ 可取遍每个满足 $j + k \le m$ 的非负整数值，每个 $a_{j,k}$ 都属于 $\mathbf{C}$，$w^j z^k$ 表示定义为 $(w, z) \mapsto w^j z^k$ 的 $\mathbf{C}^2$ 上的函数.

定义算子 $D_w, D_z \in \mathcal{L}\big(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})\big)$ 为

$$D_w p = \frac{\partial p}{\partial w} = \sum_{j+k \le m} j a_{j,k} w^{j-1} z^k \quad 与 \quad D_z p = \frac{\partial p}{\partial z} = \sum_{j+k \le m} k a_{j,k} w^j z^{k-1},$$

其中 $p$ 定义如式 (5.73). 因为算子 $D_w$ 和 $D_z$ 都只对两个变量之一求微分，而将另一个变量当作常数，所以它们称为偏微分算子（partial differentiation operator）.

算子 $D_w$ 和 $D_z$ 可交换，因为若 $p$ 定义如式 (5.73)，则

$$(D_w D_z)p = \sum_{j+k \le m} jk a_{j,k} w^{j-1} z^{k-1} = (D_z D_w)p.$$

等式 $D_w D_z = D_z D_w$ 在 $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ 上成立，还说明了另一个更为一般的结论，那就是对于性质良好的函数，偏微分运算的顺序是无关紧要的.

---

可交换矩阵并不很常见. 例如，各元素均为区间 $[-5, 5]$ 内整数的 $2 \times 2$ 矩阵，两两共可凑出 214,358,881 对（考虑顺序），但如此多对矩阵中仅有约 0.3% 是可交换的.

> 用计算机检验得，讨论范围内的 214,358,881（等于 $11^8$）对 $2\times2$ 矩阵中，仅有 674,609 对矩阵是可交换的.

下面结论说明了，两个算子可交换当且仅当它们（关于同一个基）的矩阵是可交换的.

**5.74 可交换算子对应可交换矩阵**

设 $S, T \in \mathcal{L}(V)$ 且 $v_1, \dots, v_n$ 是 $V$ 的基. 那么 $S$ 和 $T$ 可交换，当且仅当 $\mathcal{M}\big(S, (v_1, \dots, v_n)\big)$ 和 $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ 可交换.

♡

**证明** 我们有

$$\begin{aligned} S \text{ 和 } T \text{ 可交换} &\Longleftrightarrow ST = TS \\ &\Longleftrightarrow \mathcal{M}(ST) = \mathcal{M}(TS) \\ &\Longleftrightarrow \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\ &\Longleftrightarrow \mathcal{M}(S) \text{ 和 } \mathcal{M}(T) \text{ 可交换,} \end{aligned}$$

命题得证. ■

接下来这条结果表明，若两个算子可交换，那么其中一个算子的每个特征空间必在另一算子下不变. 一对可交换算子比起一对不可交换的算子具有更好的性质，其主要原因之一就在于此，我们也会多次用到这一结果.

**5.75 特征空间在可交换算子下不变**

设 $S, T \in \mathcal{L}(V)$ 可交换且 $\lambda \in \mathbf{F}$. 那么 $E(\lambda, S)$ 在 $T$ 下不变.

**证明** 设 $v \in E(\lambda, S)$. 那么

$$S(Tv) = (ST)v = (TS)v = T(Sv) = T(\lambda v) = \lambda Tv.$$

上式即表明 $Tv \in E(\lambda, S)$. 因此 $E(\lambda, S)$ 在 $T$ 下不变. ■

设我们有两个算子并且它们都可对角化. 如果我们想进行的计算同时涉及它们两者（例如，涉及它们的和），那么我们就希望这两个算子可关于相同的基作对角化. 根据下面的结论，这种操作在这两个算子可交换时是可行的.

**5.76 可同时对角化 $\Longleftrightarrow$ 可交换性**

同一向量空间上的两个可对角化算子关于相同的基都有对角矩阵，当且仅当这两个算子可交换.

**证明** 首先设 $S, T \in \mathcal{L}(V)$ 关于同一个基有对角矩阵. 两个大小相同的对角矩阵的乘积，等于将这两个矩阵对角线上的元素对应相乘所得的对角矩阵. 因此任意两个大小相同的对角矩阵都可交换. 于是 $S$ 和 $T$ 可交换（由 5.74）.

为了证明另一方向的蕴涵关系，现设 $S, T \in \mathcal{L}(V)$ 是可对角化算子且可交换. 令 $\lambda_1, \ldots, \lambda_m$ 代表 $S$ 的所有互异特征值. 因为 $S$ 可对角化，所以 5.55 (c) 就表明

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S). \tag{5.77}$$

子空间 $E(\lambda_k, S)$（$k = 1, \ldots, m$）在 $T$ 下不变（由 5.75）. 因为 $T$ 是可对角化的，所以 5.65 表明，对于每个 $k$，$T|_{E(\lambda_k, S)}$ 均可对角化. 所以对每个 $k = 1, \ldots, m$，都存在由 $T$ 的特征向量组成的 $E(\lambda_k, S)$ 的基. 将这些基合并起来就得到了 $V$ 的基【由式 (5.77)】，且该基中每个向量都既是 $S$ 的特征向量，又是 $T$ 的特征向量. 于是 $S$ 和 $T$ 关于这个基均具有对角矩阵，命题得证. ■

习题 2 将上述结论推广到多于两个算子的情况.

设 $V$ 是非零有限维复向量空间．那么 $V$ 上的每个算子都有特征向量（见 5.19）．下面结论说明，如果 $V$ 上的两个算子可交换，那么 $V$ 中存在这两个算子共有的特征向量（但这两个可交换算子并不一定有共同的特征值）．习题 9 (a) 将该结论推广到多于两个算子的情况．

**5.78 可交换算子的公共特征向量**

非零有限维复向量空间上的每对可交换算子都有公共的特征向量．

**证明** 设 $V$ 是非零有限维复向量空间且 $S, T \in \mathcal{L}(V)$ 可交换．令 $\lambda$ 是 $S$ 的特征值（5.19 告诉我们 $S$ 肯定有特征值）．于是 $E(\lambda, S) \neq \{0\}$．并且，$E(\lambda, S)$ 在 $T$ 下不变（由 5.75）．

于是，再次利用 5.19 得，$T|_{E(\lambda,S)}$ 具有特征向量，且该向量既是 $S$ 的特征向量又是 $T$ 的特征向量，证毕．■

---

**5.79 例：偏微分算子的公共特征向量**

令 $\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})$ 定义如例 5.72 所示，并令 $D_w, D_z \in \mathcal{L}\big(\mathcal{P}_m(\mathbf{C}^2, \mathbf{C})\big)$ 是该例中的可交换偏微分算子．你可验证，这两个算子的唯一特征值是 0．并且，

$$E(0, D_w) = \left\{\sum_{k=0}^{m} a_k z^k : a_0, \ldots, a_m \in \mathbf{C}\right\},$$

$$E(0, D_z) = \left\{\sum_{j=0}^{m} c_j w^j : c_0, \ldots, c_m \in \mathbf{C}\right\}.$$

这两个特征空间的交集是由这两个算子共有的特征向量所构成的集合．因为 $E(0, D_w) \cap E(0, D_z)$ 是由常值函数构成的集合，所以我们可看出 $D_w$ 和 $D_z$ 确实有公共的特征向量，正如 5.78 所言．

---

下面结论将 5.47（使算子具有上三角矩阵的基的存在性）拓展至两个可交换算子的情形中．

**5.80 可交换算子可同时上三角化**

设 $V$ 是有限维复向量空间，$S, T$ 是 $V$ 上的可交换算子．那么存在 $V$ 的一个基，使得 $S$ 和 $T$ 关于该基均有上三角矩阵．

**证明** 令 $n = \dim V$．我们对 $n$ 用归纳法．欲证结论对 $n = 1$ 是成立的，因为所有的 $1 \times 1$ 矩阵都是上三角矩阵．现在假设 $n > 1$，且欲证结论对于所有维数是 $n - 1$ 的复向量空间都成立．

令 $v_1$ 为 $S$ 和 $T$ 共有的特征向量（利用 5.78）．因此 $Sv_1 \in \operatorname{span}(v_1)$ 且 $Tv_1 \in \operatorname{span}(v_1)$．令 $W$ 为 $V$ 的子空间且满足

$$V = \operatorname{span}(v_1) \oplus W.$$

关于 $W$ 这一子空间的存在性，请看 2.33．定义线性映射 $P: V \to W$ 为：对各 $a \in \mathbf{C}$ 和各 $w \in W$ 有

$$P(av_1 + w) = w.$$

定义 $\hat{S},\hat{T} \in \mathcal{L}(W)$ 为：对每个 $w \in W$ 有

$$\hat{S}w = P(Sw) \quad 及 \quad \hat{T}w = P(Tw).$$

为了将我们的归纳假设应用于 $\hat{S}$ 和 $\hat{T}$，我们首先必须说明 $W$ 上的这两个算子可交换. 为此，设 $w \in W$. 那么存在 $a \in \mathbf{C}$，使得

$$(\hat{S}\hat{T})w = \hat{S}\big(P(Tw)\big) = \hat{S}(Tw - av_1) = P\big(S(Tw - av_1)\big) = P\big((ST)w\big),$$

其中最后一个等号成立是因为 $v_1$ 是 $S$ 的特征向量且 $Pv_1 = 0$. 类似有

$$(\hat{T}\hat{S})w = P\big((TS)w\big).$$

因为算子 $S$ 和 $T$ 可交换，所以上述两式就可说明 $(\hat{S}\hat{T})w = (\hat{T}\hat{S})w$. 因此 $\hat{S}$ 和 $\hat{T}$ 可交换.

于是我们可以利用归纳假设得出，存在 $W$ 的一个基 $v_2,\dots,v_n$ 使得 $\hat{S}$ 和 $\hat{T}$ 关于该基都有上三角矩阵. 组 $v_1,\dots,v_n$ 就是 $V$ 的基.

若 $k \in \{2,\dots,n\}$，那么存在 $a_k,b_k \in \mathbf{C}$，使得

$$Sv_k = a_kv_1 + \hat{S}v_k \quad 及 \quad Tv_k = b_kv_1 + \hat{T}v_k.$$

因为 $\hat{S}$ 和 $\hat{T}$ 关于 $v_2,\dots,v_n$ 有上三角矩阵，所以我们可知，$\hat{S}v_k \in \operatorname{span}(v_2,\dots,v_k)$ 且 $\hat{T}v_k \in \operatorname{span}(v_2,\dots,v_k)$. 因此由上述等式得

$$Sv_k \in \operatorname{span}(v_1,\dots,v_k) \quad 及 \quad Tv_k \in \operatorname{span}(v_1,\dots,v_k).$$

于是，$S$ 和 $T$ 关于基 $v_1,\dots,v_n$ 有上三角矩阵，命题得证. ■

习题 9 (b) 将上述结论推广到多于两个算子的情形.

一般来说，仅凭两个算子的特征值，无法确定这两个算子的和或积的特征值. 然而，当这两个算子可交换的时候，我们就有下面这个很棒的结论.

**5.81 可交换算子的和与积的特征值**

设 $V$ 是有限维复向量空间，$S,T$ 是 $V$ 上的可交换算子. 那么

- $S+T$ 的每个特征值都等于 $S$ 的特征值加上 $T$ 的特征值.
- $ST$ 的每个特征值都等于 $S$ 的特征值乘以 $T$ 的特征值.

**证明** 存在 $V$ 的一个基，使得 $S$ 和 $T$ 关于该基都有上三角矩阵（由 5.80）. 由 3.35 和 3.43，关于该基的矩阵满足

$$\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T) \quad 及 \quad \mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T).$$

矩阵加法的定义表明，$\mathcal{M}(S+T)$ 对角线上的每个元素都等于 $\mathcal{M}(S)$ 对角线与 $\mathcal{M}(T)$ 对角线上对应元素之和. 类似地，由于 $\mathcal{M}(S)$ 和 $\mathcal{M}(T)$ 都是上三角矩阵，所以矩阵乘法的定义表明，$\mathcal{M}(ST)$ 对角线上的每个元素都等于 $\mathcal{M}(S)$ 对角线与 $\mathcal{M}(T)$ 对角线上对应元素之积. 此外，$\mathcal{M}(S+T)$ 和 $\mathcal{M}(ST)$ 都是上三角矩阵（见 5C 节的习题 2）.

$\mathcal{M}(S)$ 对角线上的每个元素都是 $S$ 的特征值，$\mathcal{M}(T)$ 对角线上的每个元素都是 $T$ 的特征值（由 5.41）. $S+T$ 的每个特征值都在 $\mathcal{M}(S+T)$ 对角线上，$ST$ 的每个特征值都在 $\mathcal{M}(ST)$ 对角线上（这两条结论同样源于 5.41）. 综上所述，我们即可得出结论，$S+T$ 的每个特征值都等于 $S$ 的特征值加上 $T$ 的特征值，$ST$ 的每个特征值都等于 $S$ 的特征值乘以 $T$ 的特征值. ■

## 习题 5E

**1** 给出一例：$\mathbf{F}^4$ 上的两个可交换算子 $S,T$，使得 $\mathbf{F}^4$ 中有在 $S$ 下不变但不在 $T$ 下不变的子空间，以及在 $T$ 下不变但不在 $S$ 下不变的子空间.

**2** 设 $\mathcal{E}$ 是 $\mathcal{L}(V)$ 的子集，且 $\mathcal{E}$ 中每个元素都可对角化. 证明：存在 $V$ 的一个基使得 $\mathcal{E}$ 的每个元素关于它都有对角矩阵，当且仅当 $\mathcal{E}$ 中每对元素都可交换.

注 本题推广了 5.76，在那里考虑的是 $\mathcal{E}$ 只含两个元素的情形. 对本题而言，$\mathcal{E}$ 可以包含任意数量的元素，甚至可以是无限集.

**3** 设 $S,T \in \mathcal{L}(V)$ 使得 $ST = TS$. 设 $p \in \mathcal{P}(\mathbf{F})$.

(a) 证明：$\operatorname{null} p(S)$ 在 $T$ 下不变.

(b) 证明：$\operatorname{range} p(S)$ 在 $T$ 下不变.

注 $S = T$ 的特殊情形见 5.18.

**4** 证明或给出一反例：如果 $A$ 是对角矩阵且 $B$ 是和 $A$ 大小相同的上三角矩阵，那么 $A$ 和 $B$ 可交换.

**5** 证明：有限维向量空间上的一对算子可交换，当且仅当其对偶算子可交换.

注 算子的对偶的定义见 3.118.

**6** 设 $V$ 是非零有限维复向量空间，且 $S,T \in \mathcal{L}(V)$ 可交换. 证明：存在 $\alpha,\lambda \in \mathbf{C}$ 使得

$$\operatorname{range}(S-\alpha I)+\operatorname{range}(T-\lambda I) \neq V.$$

**7** 设 $V$ 是复向量空间，$S \in \mathcal{L}(V)$ 可对角化，且 $T \in \mathcal{L}(V)$ 和 $S$ 可交换. 证明：存在 $V$ 的一个基使得 $S$ 关于该基有对角矩阵而 $T$ 关于该基有上三角矩阵.

**8** 设例 5.72 中 $m = 3$，根据该例可设 $D_x, D_y$ 是 $\mathcal{P}_3(\mathbf{R}^2)$ 上可交换的偏微分算子. 求 $\mathcal{P}_3(\mathbf{R}^2)$ 的一个基，使得 $D_x, D_y$ 关于该基都有上三角矩阵.

**9** 设 $V$ 是非零有限维复向量空间. 设 $\mathcal{E} \subseteq \mathcal{L}(V)$，使得对所有 $S,T \in \mathcal{E}$ 都有 $S$ 和 $T$ 可交换.

(a) 证明：$V$ 中存在一向量，该向量是 $\mathcal{E}$ 每个元素的特征向量.

(b) 证明：存在 $V$ 的一个基，使得 $\mathcal{E}$ 每个元素关于它都有上三角矩阵.

注 本题推广了 5.78 和 5.80. 在那里考虑的是 $\mathcal{E}$ 只含两个元素的情形. 对本题而言，$\mathcal{E}$ 可以包含任意数量的元素，甚至可以是无限集.

**10** 给出一例：在一有限维实向量空间上的两个可交换算子 $S,T$，使得 $S+T$ 有特征值不等于 $S$ 的特征值加上 $T$ 的特征值，且 $ST$ 有特征值不等于 $S$ 的特征值乘以 $T$ 的特征值.

注 本题表明 5.81 在实向量空间上不再成立.

# 第 6 章 内积空间

在定义向量空间时，我们推广了 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 的线性结构（加法和标量乘法），而未顾及几何上的特性，例如长度和角度等概念．内积的概念就蕴含了这些几何层面的想法，它是我们本章研究的主题.

每种内积都可诱导出一种范数（你可以把范数看成长度）．范数满足一些重要的性质，例如毕达哥拉斯定理、三角不等式、平行四边形等式和柯西-施瓦兹不等式.

在讨论内积空间时，我们将欧几里得几何中的垂直向量这一概念重命名为正交向量．我们将看到，规范正交基在内积空间中非常有用．格拉姆-施密特过程可构造出这样的基．本章结尾处，我们将综合运用上述工具来解决最小化问题.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 和 $W$ 代表 $\mathbf{F}$ 上的向量空间.

Matthew Petroff CC BY-SA

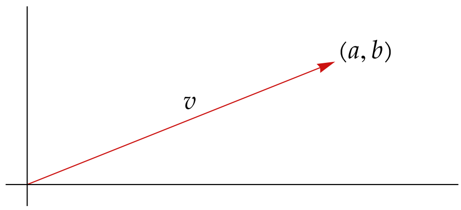
图为乔治·皮博迪图书馆（The George Peabody Library），现在是约翰·霍普金斯大学（Johns Hopkins University）的一部分．该图书馆开馆时，詹姆斯·西尔维斯特（James Sylvester，1814-1897）正担任这所大学的首位数学教授．西尔维斯特在他发表的著作中首次将**"矩阵"（matrix）**一词引入数学领域.

# 6A 内积和范数

## 内积

首先说明定义内积这个概念的动机. 将 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 中的向量看成始于原点的箭头. $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中向量 $v$ 的长度被称为 $v$ 的**范数**（**norm**）并记作 $\|v\|$. 于是对 $v=(a,b)\in\mathbf{R}^2$，我们有

$$\|v\|=\sqrt{a^2+b^2}.$$




图中的向量 $v$ 的范数为 $\sqrt{a^2+b^2}$.

类似地，若 $v=(a,b,c)\in\mathbf{R}^3$，那么 $\|v\|=\sqrt{a^2+b^2+c^2}$.

尽管我们画不出高维空间中的图形，但将上面概念推广到 $\mathbf{R}^n$ 并不难：我们定义 $x=(x_1,\dots,x_n)\in\mathbf{R}^n$ 的范数为

$$\|x\|=\sqrt{x_1^2+\cdots+x_n^2}.$$

范数在 $\mathbf{R}^n$ 上不是线性的. 为了把线性引入讨论中，我们引入点积.

**6.1 定义：点积（dot product）**

对 $x,y\in\mathbf{R}^n$，$x$ 和 $y$ 的**点积**记作 $x\cdot y$，由下式定义：

$$x\cdot y=x_1y_1+\cdots+x_ny_n,$$

其中 $x=(x_1,\dots,x_n)$，$y=(y_1,\dots,y_n)$. ♣

$\mathbf{R}^n$ 中两个向量的点积是数而不是向量. 注意，对所有 $x\in\mathbf{R}^n$，有 $x\cdot x=\|x\|^2$. 此外，$\mathbf{R}^n$ 上的点积还满足下列性质.

如果我们把向量视为点而不是箭头，那么应将 $\|x\|$ 解释为从原点到点 $x$ 的距离.

- 对所有 $x\in\mathbf{R}^n$，均有 $x\cdot x\geq 0$.
- $x\cdot x=0$ 当且仅当 $x=0$.
- 对于固定的 $y\in\mathbf{R}^n$，$\mathbf{R}^n$ 到 $\mathbf{R}$ 的将 $x\in\mathbf{R}^n$ 对应到 $x\cdot y$ 的映射是线性的.
- 对所有 $x,y\in\mathbf{R}^n$，均有 $x\cdot y=y\cdot x$.

内积是对点积的推广. 现在你可能会猜测，内积的定义就是通过将上一段讨论的点积性质抽象化得出的. 对于实向量空间，这个猜想没错. 然而，为了让我们所作的定义同时适用于实向量空间和复向量空间，我们需要在下定义之前考虑复数的情况.

回忆一下，如果 $\lambda=a+bi$，其中 $a,b\in\mathbf{R}$，那么

- $\lambda$ 的绝对值记作 $|\lambda|$，定义为 $|\lambda|=\sqrt{a^2+b^2}$；
- $\lambda$ 的复共轭记作 $\bar{\lambda}$，定义为 $\bar{\lambda}=a-bi$；
- $|\lambda|^2=\lambda\bar{\lambda}$.

绝对值和复共轭的定义及基本性质见于第 4 章.

对于 $z=(z_1,\dots,z_n)\in\mathbf{C}^n$，我们定义 $z$ 的范数为

$$\|z\|=\sqrt{|z_1|^2+\cdots+|z_n|^2}.$$

上式中需要用到绝对值，因为我们想让 $\|z\|$ 是个非负实数. 注意到

$$\|z\|^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

我们想将 $\|z\|^2$ 视为 $z$ 与自身的内积，就如在 $\mathbf{R}^n$ 中那般. 于是由上式可知，向量 $w = (w_1, \ldots, w_n) \in \mathbf{C}^n$ 和 $z$ 的内积应等于

$$w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

如果将 $w$ 和 $z$ 的角色互换，上述表达式就会被其复共轭所替代. 于是，我们期望内积满足：$w$ 和 $z$ 的内积等于 $z$ 和 $w$ 的内积的复共轭. 说明这点动机后，我们现在可以定义任意实向量空间或复向量空间 $V$ 上的内积了.

关于下面定义中所用的记号，有一点需要注意：

- 对于 $\lambda \in \mathbf{C}$，记号 $\lambda \geq 0$ 意为 $\lambda$ 是非负实数.

**6.2 定义：内积（inner product）**

$V$ 上的**内积**是一个函数，它将由 $V$ 中元素构成的每个有序对 $(u, v)$ 对应至一个数 $\langle u, v\rangle \in \mathbf{F}$，并满足如下性质.

**正性（positivity）**

对于所有 $v \in V$，均有 $\langle v, v\rangle \geq 0$.

**定性（definiteness）**

$\langle v, v\rangle = 0$ 当且仅当 $v = 0$.

**第一个位置上的可加性（additivity in first slot）**

对于所有 $u, v, w \in V$，均有 $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

**第一个位置上的齐次性（homogeneity in first slot）**

对于所有 $\lambda \in \mathbf{F}$ 和所有 $u, v \in V$，均有 $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$.

**共轭对称性（conjugate symmetry）**

对于所有 $u, v \in V$，均有 $\langle u, v\rangle = \overline{\langle v, u\rangle}$.

♣

每个实数都等于其复共轭. 因此如果我们讨论的是实向量空间，那么我们可以从上面最后一个条件中省去复共轭，并直接把它表述为：对于所有 $u, v \in V$，均有 $\langle u, v\rangle = \langle v, u\rangle$.

大部分数学家都把内积定义如上. 不过，许多物理学家使用的定义中，要求齐次性在第二个位置上成立，而不是第一个位置.

**6.3 例：内积**

(a) $\mathbf{F}^n$ 上的**欧几里得内积（Euclidean inner product）**定义为：对所有 $(w_1, \ldots, w_n), (z_1, \ldots, z_n) \in \mathbf{F}^n$，

$$\big\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\big\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

(b) 若 $c_1, \ldots, c_n$ 是正数，那么可在 $\mathbf{F}^n$ 上定义内积如下：对所有 $(w_1, \ldots, w_n), (z_1, \ldots, z_n) \in \mathbf{F}^n$，

$$\big\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\big\rangle = c_1 w_1\overline{z_1} + \cdots + c_n w_n\overline{z_n}.$$

转下页

(c) 在由定义在区间 $[-1,1]$ 上的全体连续实值函数构成的向量空间上，可定义内积如下：对所有定义在区间 $[-1,1]$ 上的连续实值函数 $f,g$，

$$\langle f,g\rangle=\int_{-1}^{1}fg.$$

(d) 在 $\mathcal{P}(\mathbf{R})$ 上，可定义内积如下：对所有 $p,q\in\mathcal{P}(\mathbf{R})$，

$$\langle p,q\rangle=p(0)q(0)+\int_{-1}^{1}p'q'.$$

(e) 在 $\mathcal{P}(\mathbf{R})$ 上，还可定义内积如下：对所有 $p,q\in\mathcal{P}(\mathbf{R})$，

$$\langle p,q\rangle=\int_{0}^{\infty}p(x)q(x)\mathrm{e}^{-x}\mathrm{d}x.$$

**6.4 定义：内积空间（inner product space）**

带有内积的向量空间称为**内积空间**. ♣

内积空间的最重要的例子，就是带有如上例 (a) 所示欧几里得内积的 $\mathbf{F}^n$. 当称 $\mathbf{F}^n$ 是内积空间时，除非另有说明，你都应假设其上定义的内积是欧几里得内积.

为了让我们不用反复重申 $V$ 和 $W$ 是内积空间这个前提条件，我们作出如下假设.

**6.5 记号：$V$、$W$**

在本章的剩余部分和下章中，$V$ 和 $W$ 都指代 $\mathbf{F}$ 上的内积空间. ♣

注意，这里稍微有些滥用术语. 内积空间是带有内积的向量空间. 当我们称一向量空间 $V$ 为内积空间时，我们或是将 $V$ 上的内积隐含于其中，或是由上下文可明确 $V$ 上的内积如何定义（又或者，如果这个向量空间是 $\mathbf{F}^n$，那么所用内积就是欧几里得内积）.

**6.6 内积的基本性质**

(a) 对每个固定的 $v\in V$，将 $u\in V$ 对应到 $\langle u,v\rangle$ 的函数都是 $V$ 到 $\mathbf{F}$ 的线性映射.

(b) 对每个 $v\in V$，均有 $\langle 0,v\rangle=0$.

(c) 对每个 $v\in V$，均有 $\langle v,0\rangle=0$.

(d) 对所有 $u,v,w\in V$，均有 $\langle u,v+w\rangle=\langle u,v\rangle+\langle u,w\rangle$.

(e) 对所有 $\lambda\in\mathbf{F}$ 和 $u,v\in V$，均有 $\langle u,\lambda v\rangle=\overline{\lambda}\langle u,v\rangle$. ♡

**证明**

(a) 对 $v\in V$，由内积定义中第一个位置上的可加性和齐次性，即可得 $u\mapsto\langle u,v\rangle$ 的线性.

(b) 每个线性映射都将 0 对应到 0. 所以由 (a) 即知 (b) 成立.

(c) 若 $v\in V$，那么由内积定义中的共轭对称性质和 (b) 即得 $\langle v,0\rangle=\overline{\langle 0,v\rangle}=\overline{0}=0$.

(d) 设 $u, v, w \in V$. 则

$$\begin{aligned}\langle u, v+w\rangle &= \overline{\langle v+w, u\rangle} \\ &= \overline{\langle v, u\rangle + \langle w, u\rangle} \\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} \\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

(e) 设 $\lambda \in \mathbf{F}$ 及 $u, v \in V$. 则

$$\begin{aligned}\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle} \\ &= \overline{\lambda \langle v, u\rangle} \\ &= \bar{\lambda}\,\overline{\langle v, u\rangle} \\ &= \bar{\lambda}\langle u, v\rangle.\end{aligned}$$

■

## 范数

我们定义内积的动机最初来源于 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 上向量的范数. 现在我们就会看到，每种内积都能确定一种范数.

**6.7 定义：范数（norm）、$\|v\|$**

对 $v \in V$，$v$ 的**范数**记作 $\|v\|$，定义为

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

♣

**6.8 例：范数**

(a) 若 $(z_1, \ldots, z_n) \in \mathbf{F}^n$（其上定义了欧几里得内积），那么

$$\|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

(b) 定义在区间 $[-1, 1]$ 上的连续实值函数所构成的向量空间【其上内积定义如 6.3 (c)】中的元素 $f$ 满足

$$\|f\| = \sqrt{\int_{-1}^{1} f^2}.$$

**6.9 范数的基本性质**

设 $v \in V$.

(a) $\|v\| = 0$ 当且仅当 $v = 0$.

(b) 对所有 $\lambda \in \mathbf{F}$，均有 $\|\lambda v\| = |\lambda|\,\|v\|$.

♡

### 证明

(a) 因为 $\langle v, v\rangle = 0$ 当且仅当 $v = 0$，所以原命题成立.

(b) 设 $\lambda \in \mathbf{F}$. 那么

$$\begin{aligned}\|\lambda v\|^2 &= \langle \lambda v, \lambda v\rangle \\ &= \lambda\langle v, \lambda v\rangle \\ &= \lambda\bar{\lambda}\langle v, v\rangle \\ &= |\lambda|^2\|v\|^2.\end{aligned}$$

两边开平方根即得我们欲证的等式. ■

上述结论中 (b) 的证明形象说明了一个普遍适用的道理：处理范数的平方通常比直接处理范数来得容易.

现在我们给出一个关键的定义.

**6.10 定义：正交（orthogonal）**

称两个向量 $u, v \in V$ 是**正交的**，若 $\langle u, v\rangle = 0$. ♣

在上述定义中，两个向量的顺序是无关紧要的，因为 $\langle u, v\rangle = 0$ 当且仅当 $\langle v, u\rangle = 0$. 除了说“$u$ 和 $v$ 是正交的”，我们有时也说“$u$ 正交于 $v$”.

**“orthogonal”** 这个词来自希腊语中的 **“orthogonios”**，后者意为“直角的”.

习题 15 要求你证明若 $u, v$ 是 $\mathbf{R}^2$ 中的非零向量，那么

$$\langle u, v\rangle = \|u\|\|v\| \cos\theta,$$

其中 $\theta$ 是 $u$ 和 $v$ 间的夹角（将 $u$ 和 $v$ 看成始于原点的箭头）. 于是 $\mathbf{R}^2$ 中两个非零向量（在欧几里得内积下）是正交的，当且仅当它们之间夹角的余弦值是 0，这又等价于通常讨论平面几何时所说的两向量垂直. 于是，你可将**正交**这个词看成表达**垂直（perpendicular）**之意的一个更炫的说法.

我们从一个简单的结论开始研究正交性.

**6.11 正交性和 0**

(a) 0 与 $V$ 中每个向量都正交.

(b) 0 是 $V$ 中唯一与自身正交的向量. ♡

**证明**

(a) 回忆一下，6.6 (b) 曾提到，对每个 $v \in V$ 都有 $\langle 0, v\rangle = 0$.

(b) 若 $v \in V$ 且 $\langle v, v\rangle = 0$，那么 $v = 0$（由内积的定义可得）. ■

下面的定理在 $V = \mathbf{R}^2$ 时的特殊情形，早在 3500 多年前就已经为古巴比伦人所知，在 2500 余年前又被希腊人重新发现并证明. 当然，下面的证明用的不是原始的证法.

**6.12 毕达哥拉斯定理[1]（Pythagorean theorem）**

设 $u, v \in V$. 若 $u$ 和 $v$ 是正交的，那么

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$ ♡

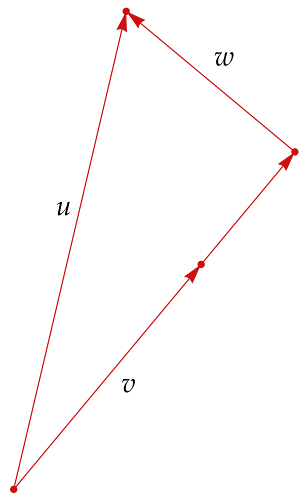
[1]该定理又称勾股定理.

**证明** 设 $\langle u,v\rangle=0$. 那么

$$\begin{aligned}\|u+v\|^2&=\langle u+v,u+v\rangle\\&=\langle u,u\rangle+\langle u,v\rangle+\langle v,u\rangle+\langle v,v\rangle\\&=\|u\|^2+\|v\|^2.\end{aligned}$$

■

设 $u,v\in V$ 且 $v\neq 0$. 我们想将 $u$ 写成 $v$ 的标量倍加上正交于 $v$ 的向量 $w$ 的形式，如下图所示.




正交分解：将 $u$ 表示成 $v$ 的标量倍加上正交于 $v$ 的向量的形式.

为探究如何将 $u$ 写成 $v$ 的标量倍加上正交于 $v$ 的向量的形式，令 $c\in\mathbf{F}$ 表示一标量. 则

$$u=cv+(u-cv).$$

于是我们需选取 $c$ 使得 $v$ 与 $u-cv$ 正交. 因此我们希望有

$$0=\langle u-cv,v\rangle=\langle u,v\rangle-c\|v\|^2.$$

上式表明，我们应将 $c$ 取为 $\langle u,v\rangle/\|v\|^2$. 取定该 $c$ 后，我们就可写出

$$u=\frac{\langle u,v\rangle}{\|v\|^2}v+\left(u-\frac{\langle u,v\rangle}{\|v\|^2}v\right).$$

你可验证得，上述等式将 $u$ 显式地写成了 $v$ 的标量倍加上正交于 $v$ 的向量的形式. 于是我们就证明了下面这个关键的结论.

**6.13 一种正交分解**

设 $u,v\in V$，且 $v\neq 0$. 取 $c=\dfrac{\langle u,v\rangle}{\|v\|^2}$ 及 $w=u-\dfrac{\langle u,v\rangle}{\|v\|^2}v$. 那么

$$u=cv+w \quad 且 \quad \langle w,v\rangle=0.$$

我们将利用正交分解 6.13 来证明下面的柯西-施瓦兹不等式，它是数学中最重要的不等式之一.

**6.14 柯西-施瓦兹不等式（Cauchy-Schwarz inequality）**

设 $u, v \in V$. 那么

$$|\langle u, v\rangle| \leq \|u\|\|v\|.$$

当且仅当 $u, v$ 成标量倍数关系时，上述不等式取得等号.

**证明** 若 $v = 0$，那么欲证不等式的两端都等于 0，即得证. 于是我们可假设 $v \neq 0$. 考察 6.13 给出的正交分解式

$$u = \frac{\langle u, v\rangle}{\|v\|^2}v + w$$

其中 $w$ 正交于 $v$. 由毕达哥拉斯定理，

$$\begin{aligned}\|u\|^2 &= \left\|\frac{\langle u, v\rangle}{\|v\|^2}v\right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned} \tag{6.15}$$

将上式两边都乘以 $\|v\|^2$ 并同时开平方根即得欲证的不等式.

1821 年，奥古斯汀-路易·柯西（Augustin-Louis Cauchy, 1789-1857）证明了 6.16 (a). 1859 年，柯西的学生维克多·布尼亚科夫斯基（Viktor Bunyakovsky，1804-1889）证明了类似于 6.16 (b) 的积分不等式. 几十年后，赫尔曼·施瓦兹（Hermann Schwarz，1843-1921）发现了类似的结论并引发了更多的关注，柯西-施瓦兹不等式由此得名.

由上段的证明可得柯西-施瓦兹不等式取得等号当且仅当式 (6.15) 取得等号，这又等价于 $w = 0$. 而 $w = 0$ 当且仅当 $u$ 是 $v$ 的标量倍（见 6.13）. 于是柯西-施瓦兹不等式取得等号，当且仅当 $u$ 是 $v$ 的标量倍或 $v$ 是 $u$ 的标量倍（或互成标量倍数关系）. 这样说可以涵盖 $u$ 或 $v$ 等于 0 这个特殊情形. ■

**6.16 例：柯西-施瓦兹不等式**

(a) 若 $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbf{R}$，那么

$$(x_1y_1 + \cdots + x_ny_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2),$$

这就是将柯西-施瓦兹不等式应用于向量 $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbf{R}^n$ 的结果（所用内积是通常的欧几里得内积）.

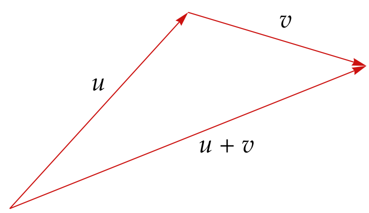
(b) 若 $f, g$ 是定义在 $[-1, 1]$ 上的连续实值函数，那么

$$\left|\int_{-1}^{1} fg\right|^2 \leq \left(\int_{-1}^{1} f^2\right)\left(\int_{-1}^{1} g^2\right),$$

这就是将柯西-施瓦兹不等式应用于例 6.3 (c) 的结果.

接下来的结论被称为三角不等式，从几何角度解释，就是三角形任意一边的长度都小于另外两边的长度之和.

注意，三角不等式表明，连接两点的折线路径中，最短的是一条线段（可视作由共线的线段构成的折线）.




在这个三角形里，$u + v$ 的长度小于 $u$ 的长度加上 $v$ 的长度.

**6.17 三角不等式（triangle inequality）**

设 $u, v \in V$. 那么

$$\|u+v\| \leq \|u\|+\|v\|.$$

该不等式取得等号，当且仅当 $u, v$ 中任意一者是另一者的非负实数倍.

**证明** 我们有

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle+\langle v,v\rangle+\langle u,v\rangle+\langle v,u\rangle \\
&= \langle u,u\rangle+\langle v,v\rangle+\langle u,v\rangle+\overline{\langle u,v\rangle} \\
&= \|u\|^2+\|v\|^2+2\operatorname{Re}\langle u,v\rangle \\
&\leq \|u\|^2+\|v\|^2+2|\langle u,v\rangle| \qquad (6.18)\\
&\leq \|u\|^2+\|v\|^2+2\|u\|\|v\| \qquad (6.19)\\
&= (\|u\|+\|v\|)^2,
\end{aligned}$$

其中式 (6.19) 源自柯西-施瓦兹不等式（6.14）. 将上式两侧同时开平方根，即得欲证的不等式.

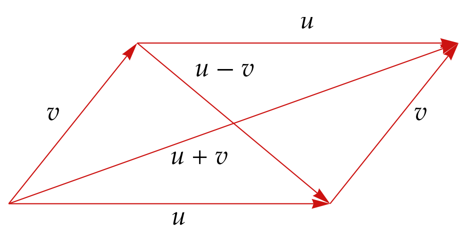
上述证明表明，三角不等式取得等号，当且仅当式 (6.18) 和 (6.19) 取得等号. 于是三角不等式取得等号等价于

$$\langle u, v\rangle = \|u\|\|v\|. \tag{6.20}$$

如果 $u, v$ 之一是另一者的非负实数倍，那么式 (6.20) 成立. 反之，设式 (6.20) 成立. 那么柯西-施瓦兹不等式（6.14）取等号的条件就表明，$u, v$ 必成标量倍数关系. 由式 (6.20) 又可知该标量必须是非负实数，这就完成了证明. ■

反向三角不等式见于本节习题 20.

下面结果因其几何解释而被称作平行四边形等式：在平行四边形中，对角线的长度平方之和等于四条边的长度平方之和. 注意，比起常见的欧几里得几何中的证明，此处的证明更直截了当.




图示平行四边形的对角线是 $u+v$ 和 $u-v$.

**6.21 平行四边形等式（parallelogram equality）**

设 $u, v \in V$. 那么

$$\|u+v\|^2+\|u-v\|^2 = 2(\|u\|^2+\|v\|^2).$$

**证明** 我们有

$$\begin{aligned}\|u+v\|^2+\|u-v\|^2 &= \langle u+v,u+v\rangle+\langle u-v,u-v\rangle\\ &= \|u\|^2+\|v\|^2+\langle u,v\rangle+\langle v,u\rangle\\ &\quad +\|u\|^2+\|v\|^2-\langle u,v\rangle-\langle v,u\rangle\\ &= 2(\|u\|^2+\|v\|^2),\end{aligned}$$

命题得证. ■

## 习题 6A

**1** 证明或给出一反例：如果 $v_1,\ldots,v_m\in V$，那么

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j,v_k\rangle\geq 0.$$

**2** 设 $S\in\mathcal{L}(V)$. 定义 $\langle\cdot,\cdot\rangle_1$ 为

$$\langle u,v\rangle_1=\langle Su,Sv\rangle$$

对所有 $u,v\in V$ 成立. 证明：$\langle\cdot,\cdot\rangle_1$ 是 $V$ 上的内积，当且仅当 $S$ 是单射.

**3** (a) 证明：将 $\mathbf{R}^2$ 中元素的有序对 $\big((x_1,x_2),(y_1,y_2)\big)$ 映到 $|x_1y_1|+|x_2y_2|$ 的函数，不是 $\mathbf{R}^2$ 上的内积.

(b) 证明：将 $\mathbf{R}^3$ 中元素的有序对 $\big((x_1,x_2,x_3),(y_1,y_2,y_3)\big)$ 映到 $x_1y_1+x_3y_3$ 的函数，不是 $\mathbf{R}^3$ 上的内积.

**4** 设 $T\in\mathcal{L}(V)$ 使得 $\|Tv\|\leq\|v\|$ 对每个 $v\in V$ 成立. 证明：$T-\sqrt{2}I$ 是单射.

**5** 设 $V$ 是实内积空间.

(a) 证明：$\langle u+v,u-v\rangle=\|u\|^2-\|v\|^2$ 对任意 $u,v\in V$ 都成立.

(b) 证明：如果 $u,v\in V$ 范数相同，那么 $u+v$ 正交于 $u-v$.

(c) 利用 (b) 证明：菱形的对角线相互垂直.

**6** 设 $u,v\in V$. 证明：$\langle u,v\rangle=0\iff$ 对所有 $a\in\mathbf{F}$ 都有 $\|u\|\leq\|u+av\|$.

**7** 设 $u,v\in V$. 证明：$\|au+bv\|=\|bu+av\|$ 对所有 $a,b\in\mathbf{R}$ 成立，当且仅当 $\|u\|=\|v\|$.

**8** 设 $a,b,c,x,y\in\mathbf{R}$ 且 $a^2+b^2+c^2+x^2+y^2\leq 1$. 证明：$a+b+c+4x+9y\leq 10$.

**9** 设 $u,v\in V$，$\|u\|=\|v\|=1$ 且 $\langle u,v\rangle=1$. 证明：$u=v$.

**10** 设 $u,v\in V$，$\|u\|\leq 1$ 且 $\|v\|\leq 1$. 证明：

$$\sqrt{1-\|u\|^2}\sqrt{1-\|v\|^2}\leq 1-|\langle u,v\rangle|.$$

**11** 求向量 $u,v\in\mathbf{R}^2$，使得 $u$ 是 $(1,3)$ 的标量倍，$v$ 正交于 $(1,3)$，且 $(1,2)=u+v$.

**12** 设 $a,b,c,d$ 是正数.

(a) 证明：$(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)\geq 16$.

(b) $a,b,c,d$ 取多少时，以上不等式取等?

**13** 证明：均值的平方小于或等于平方的均值. 更准确地说，是证明：如果 $a_1,\ldots,a_n\in\mathbf{R}$，那么 $a_1,\ldots,a_n$ 的均值的平方小于或等于 $a_1^2,\ldots,a_n^2$ 的均值.

**14** 设 $v \in V$（$v \neq 0$）. 证明：$v/\|v\|$ 是 $V$ 的单位球面上唯一最接近 $v$ 的元素. 更准确地说，是证明：如果 $u \in V$ 且 $\|u\| = 1$，那么

$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\|,$$

且仅当 $u = v/\|v\|$ 时取等.

**15** 设 $u, v$ 是 $\mathbf{R}^2$ 中的非零向量. 证明：

$$\langle u, v \rangle = \|u\| \, \|v\| \cos\theta,$$

其中 $\theta$ 是 $u$ 和 $v$ 的夹角（将 $u$ 和 $v$ 视为从原点出发的箭头）.

提示 在由 $u, v$ 和 $u - v$ 形成的三角形上应用余弦定理.

**16** $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中两向量（视为从原点出发的箭头）的夹角可以用几何方法定义. 然而，当 $n > 3$ 时，$\mathbf{R}^n$ 中的几何就不那样清晰了. 因此，定义两非零向量 $x, y \in \mathbf{R}^n$ 的夹角为

$$\arccos \frac{\langle x, y \rangle}{\|x\| \, \|y\|}.$$

这一定义的动机来自第 15 题. 请解释，为什么证明这一定义有意义需要用到柯西-施瓦兹不等式.

**17** 证明：

$$\left( \sum_{k=1}^{n} a_k b_k \right)^2 \leq \left( \sum_{k=1}^{n} k a_k^2 \right) \left( \sum_{k=1}^{n} \frac{b_k^2}{k} \right)$$

对所有实数 $a_1, \ldots, a_n$ 和 $b_1, \ldots, b_n$ 成立.

**18** (a) 设 $f : [1, \infty) \to [0, \infty)$ 连续. 证明：

$$\left( \int_1^\infty f \right)^2 \leq \int_1^\infty x^2 \bigl( f(x) \bigr)^2 \mathrm{d}x.$$

(b) 当连续函数 $f : [1, \infty) \to [0, \infty)$ 是什么函数时，以上不等式取等且两边均有限？

**19** 设 $v_1, \ldots, v_n$ 是 $V$ 的基，且 $T \in \mathcal{L}(V)$. 证明：如果 $\lambda$ 是 $T$ 的特征值，那么

$$|\lambda|^2 \leq \sum_{j=1}^{n} \sum_{k=1}^{n} \left| \mathcal{M}(T)_{j,k} \right|^2,$$

其中 $\mathcal{M}(T)_{j,k}$ 表示 $T$ 关于基 $v_1, \ldots, v_n$ 的矩阵第 $j$ 行第 $k$ 列的元素.

**20** 证明：如果 $u, v \in V$，那么 $\bigl| \|u\| - \|v\| \bigr| \leq \|u - v\|$.

注 以上不等式称为**反向三角不等式**. 当 $V = \mathbf{C}$ 时的反向三角不等式，见第 4 章习题 2.

**21** 设 $u, v \in V$ 使得

$$\|u\| = 3, \quad \|u + v\| = 4, \quad \|u - v\| = 6.$$

请问 $\|v\|$ 等于多少？

**22** 证明：如果 $u, v \in V$，那么

$$\|u + v\| \, \|u - v\| \leq \|u\|^2 + \|v\|^2.$$

**23** 设 $v_1, \ldots, v_m \in V$ 使得 $\|v_k\| \leq 1$ 对任一 $k = 1, \ldots, m$ 都成立. 证明：存在 $a_1, \ldots, a_m \in \{1, -1\}$ 使得

$$\|a_1 v_1 + \cdots + a_m v_m\| \leq \sqrt{m}.$$

**24** 证明或给出一反例：如果 $\|\cdot\|$ 是与 $\mathbf{R}^2$ 上的一内积关联的范数，那么存在 $(x, y) \in \mathbf{R}^2$ 使得 $\|(x, y)\| \neq \max\{|x|, |y|\}$.

**25** 设 $p > 0$. 证明：$\mathbf{R}^2$ 上存在一内积，使得与其关联的范数由下式给出：对所有 $(x, y) \in \mathbf{R}^2$，

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p},$$

当且仅当 $p = 2$.

**26** 设 $V$ 是实内积空间. 证明：

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

对所有 $u, v \in V$ 成立.

**27** 设 $V$ 是复内积空间，证明：

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+\mathrm{i}v\|^2\mathrm{i} - \|u-\mathrm{i}v\|^2\mathrm{i}}{4}$$

对所有 $u, v \in V$ 成立.

**28** 对于向量空间 $U$，其上的范数是这样一个函数

$$\|\cdot\| : U \to [0, \infty),$$

其满足下列性质：$\|u\| = 0$ 当且仅当 $u = 0$；$\|\alpha u\| = |\alpha|\,\|u\|$ 对所有 $\alpha \in \mathbf{F}$ 和所有 $u \in U$ 成立；$\|u+v\| \leq \|u\| + \|v\|$ 对所有 $u, v \in U$ 成立. 证明：满足平行四边形等式的范数来自内积——换句话说，是要证明，如果 $\|\cdot\|$ 是 $U$ 上满足平行四边形等式的范数，那么 $U$ 上就存在一内积 $\langle\cdot,\cdot\rangle$ 使得对所有 $u \in U$ 有 $\|u\| = \langle u, u\rangle^{1/2}$.

**29** 设 $V_1, \ldots, V_m$ 是内积空间. 证明等式

$$\big\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \big\rangle = \langle u_1, v_1\rangle + \cdots + \langle u_m, v_m\rangle$$

定义了在 $V_1 \times \cdots \times V_m$ 上的一内积.

注 上式右侧的表达式中，对于每个 $k = 1, \ldots, m$，$\langle u_k, v_k\rangle$ 表示 $V_k$ 上的内积. 尽管这里用了同样的记号，但是 $V_1, \ldots, V_m$ 这些空间上可能有不同的内积.

**30** 设 $V$ 是实内积空间. 对于 $u, v, w, x \in V$，定义

$$\langle u + \mathrm{i}v, w + \mathrm{i}x\rangle_{\mathbf{C}} = \langle u, w\rangle + \langle v, x\rangle + (\langle v, w\rangle - \langle u, x\rangle)\mathrm{i}.$$

(a) 证明：$\langle\cdot,\cdot\rangle_{\mathbf{C}}$ 使 $V_{\mathbf{C}}$ 成为复内积空间.

(b) 证明：如果 $u, v \in V$，那么

$$\langle u, v\rangle_{\mathbf{C}} = \langle u, v\rangle \quad 且 \quad \|u + \mathrm{i}v\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

注 复化 $V_{\mathbf{C}}$ 的定义见 1B 节的习题 8.

**31** 设 $u, v, w \in V$. 证明：

$$\left\|w - \tfrac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

**32** 设 $E$ 是 $V$ 的子集，具有这一性质：$u, v \in E$ 蕴涵 $\frac{1}{2}(u+v) \in E$. 令 $w \in V$. 证明：$E$ 中至多有一个最接近 $w$ 的点. 换句话说，是要证明，至多存在一个 $u \in E$ 使得

$$\|w - u\| \leq \|w - x\|$$

对所有 $x \in E$ 成立.

**33** 设 $f, g$ 是从 $\mathbf{R}$ 到 $\mathbf{R}^n$ 的可微函数.

(a) 证明:

$$\langle f(t), g(t)\rangle' = \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle.$$

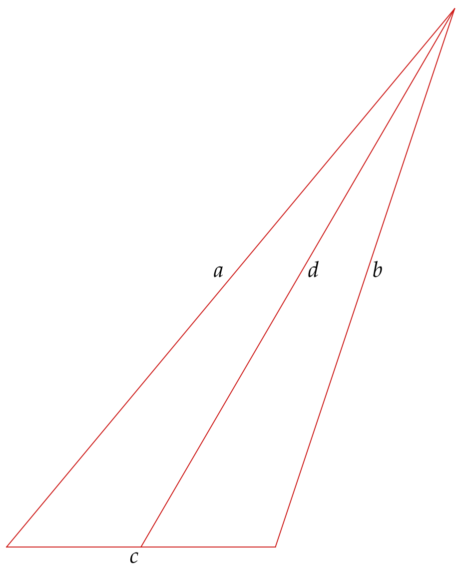
(b) 设 $c$ 是一正数，且 $\|f(t)\| = c$ 对任意 $t \in \mathbf{R}$ 成立. 证明：$\langle f'(t), f(t)\rangle = 0$ 对任意 $t \in \mathbf{R}$ 成立.

(c) 用几何的方式，从 $\mathbf{R}^n$ 中以原点为心的球面上一曲线的切向量这一角度，解释 (b) 中的结果.

注 称函数 $f : \mathbf{R} \to \mathbf{R}^n$ 是可微的，如果存在从 $\mathbf{R}$ 到 $\mathbf{R}$ 的可微函数 $f_1, \ldots, f_n$ 使得 $f(t) = (f_1(t), \ldots, f_n(t))$ 对任意 $t \in \mathbf{R}$ 成立. 进一步，对任意 $t \in \mathbf{R}$，导数 $f'(t) \in \mathbf{R}^n$ 定义为 $f'(t) = (f_1'(t), \ldots, f_n'(t))$.

**34** 利用内积证明阿波罗尼斯恒等式（Apollonius's identity）：在三边长为 $a, b, c$ 的三角形中，令 $d$ 为从长为 $c$ 的边中点到其对顶点的线段长度，则有

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$




**35** 取定一正整数 $n$. $\mathbf{R}^n$ 上的二阶可微实值函数 $p$ 的**拉普拉斯算子（Laplacian）**$\Delta p$ 是 $\mathbf{R}^n$ 上的函数，定义为

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

如果 $\Delta p = 0$，则称函数 $p$ 为**调和的（harmonic）**.

$\mathbf{R}^n$ 上的**多项式**是形如 $x_1^{m_1} \cdots x_n^{m_n}$ 的函数的线性组合（其系数在 $\mathbf{R}$ 中），其中 $m_1, \ldots, m_n$ 是非负整数.

设 $q$ 是 $\mathbf{R}^n$ 上的多项式. 证明：$\mathbf{R}^n$ 上存在一调和多项式 $p$，使得 $p(x) = q(x)$ 对每个满足 $\|x\| = 1$ 的 $x \in \mathbf{R}^n$ 成立.

注 就本题而言，关于调和函数，你只需要知道，如果 $p$ 是 $\mathbf{R}^n$ 上的调和函数且 $p(x) = 0$ 对所有满足 $\|x\| = 1$ 的 $x \in \mathbf{R}^n$ 成立，那么 $p = 0$.

**提示** 一个合理的猜测是，待求的调和多项式 $p$ 形如 $q+(1-\|x\|^2)r$，其中 $r$ 是某个多项式. 通过在合适的向量空间上定义算子 $T$ 为

$$Tr=\Delta\big((1-\|x\|^2)r\big)$$

并证明 $T$ 是单射从而是满射，以证明存在 $\mathbf{R}^n$ 上的多项式 $r$ 使得 $q+(1-\|x\|^2)r$ 是调和的.

In realms of numbers, where the secrets lie,
A noble truth emerges from the deep,
Cauchy and Schwarz, their wisdom they apply,
An inequality for all to keep.

Two vectors, by this bond, are intertwined,
As inner products weave a gilded thread,
Their magnitude, by providence, confined,
A bound to which their destiny is wed.

Though shadows fall, and twilight dims the day,
This inequality will stand the test,
To guide us in our quest, to light the way,
And in its truth, our understanding rest.

So sing, ye muses, of this noble feat,
Cauchy-Schwarz, the bound that none can beat.

柯西施瓦显真章，
数学领域绽光芒.
向量序列皆适用，
点积模长不等长.
解析几何显神通，
优化问题倚为强.
统计分析多应用，
数学之美尽飘扬.

左：由 ChatGPT 所作，输入提示词为“Shakespearean sonnet on Cauchy–Schwarz inequality”（主题为柯西-施瓦兹不等式的莎士比亚风格十四行诗）

右：由文心一言所作，输入提示词为“就‘柯西-施瓦兹不等式’写一首七言律诗”[2]

[2]经过讨论，译者决定保持诗歌的原貌，并仿照作者，利用 AI 大模型生成一首诗，但采用中文的诗体，读者可比照阅读.

# 6B 规范正交基

## 规范正交组和格拉姆-施密特过程

**6.22 定义：规范正交（orthonormal）**

- 如果一个向量组中所有向量的范数都是 1，且每个向量与其他向量都正交，则称该向量组是**规范正交的**.
- 换言之，$V$ 中向量组 $e_1, \ldots, e_m$ 是规范正交的，若对所有 $j, k \in \{1, \ldots, m\}$，

$$\langle e_j, e_k \rangle = \begin{cases} 1, & \text{若 } j = k, \\ 0, & \text{若 } j \neq k. \end{cases}$$

**6.23 例：规范正交组**

(a) $\mathbf{F}^n$ 的标准基是一个规范正交组.

(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ 是 $\mathbf{F}^3$ 中的规范正交组.

(c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ 是 $\mathbf{F}^3$ 中的规范正交组.

(d) 设 $n$ 是正整数. 那么，正如习题 4 要求你验证的那样，

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}$$

是 $C[-\pi, \pi]$ 中的规范正交向量组. $C[-\pi, \pi]$ 是由定义在区间 $[-\pi, \pi]$ 上的连续实值函数所构成的向量空间，且其上内积定义为

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

上述规范正交组经常用于建立潮汐等周期现象的数学模型.

(e) 假设我们在 $\mathcal{P}_2(\mathbf{R})$ 上定义内积为：对于所有 $p, q \in \mathcal{P}_2(\mathbf{R})$，

$$\langle p, q \rangle = \int_{-1}^{1} pq.$$

这样 $\mathcal{P}_2(\mathbf{R})$ 便成为内积空间. $\mathcal{P}_2(\mathbf{R})$ 的标准基 $1, x, x^2$ 不是规范正交组，因为该组中向量的范数不是 1. 将其中每个向量都除以各自的范数，可得组 $1/\sqrt{2}, \sqrt{3/2}x, \sqrt{5/2}x^2$，其中每个向量的范数都是 1，并且第二个向量与第一个、第三个向量正交. 然而，第一个向量和第三个向量不正交，从而该组不是规范正交组. 很快我们就会知道如何由标准基 $1, x, x^2$ 构造一个规范正交组（见例 6.34）.

处理规范正交组很容易，如下面结论所示.

**6.24 规范正交组线性组合的范数**

设 $e_1, \ldots, e_m$ 是 $V$ 中的规范正交向量组. 那么对所有 $a_1, \ldots, a_m \in \mathbf{F}$，有

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2.$$

**证明** 因为各 $e_k$ 的范数都是 1，所以只要反复利用毕达哥拉斯定理（6.12）即可得欲证等式. ■

下面结论是上述结论的一个重要推论.

**6.25　规范正交组是线性无关的**

每个规范正交向量组都是线性无关的.

**证明**　设 $e_1, \ldots, e_m$ 是 $V$ 中的规范正交向量组，且 $a_1, \ldots, a_m \in \mathbf{F}$ 满足

$$a_1 e_1 + \cdots + a_m e_m = 0.$$

那么 $|a_1|^2 + \cdots + |a_m|^2 = 0$（由 6.24），这表明各 $a_k$ 都等于 0. 于是 $e_1, \ldots, e_m$ 是线性无关的. ■

现在我们给出一个重要的不等式.

**6.26　贝塞尔不等式（Bessel's inequality）**

设 $e_1, \ldots, e_m$ 是 $V$ 中的规范正交向量组. 若 $v \in V$，那么

$$|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \leq \|v\|^2.$$

**证明**　设 $v \in V$. 那么

$$v = \underbrace{\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m}_{u} + \underbrace{v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m}_{w}.$$

令 $u$ 和 $w$ 按上式定义. 若 $k \in \{1, \ldots, m\}$，那么 $\langle w, e_k \rangle = \langle v, e_k \rangle - \langle v, e_k \rangle \langle e_k, e_k \rangle = 0$. 这意味着 $\langle w, u \rangle = 0$. 现由毕达哥拉斯定理可得

$$\begin{aligned}\|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\geq \|u\|^2 \\ &= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2,\end{aligned}$$

其中最后一行来自 6.24. ■

我们在下面定义中引入研究内积空间时最有用的概念之一.

**6.27　定义：规范正交基（orthonormal basis）**

$V$ 的**规范正交基**，就是 $V$ 中既是规范正交组又是基的向量组.

例如，标准基是 $\mathbf{F}^n$ 的一个规范正交基.

**6.28　长度恰好的规范正交组是规范正交基**

设 $V$ 是有限维的. 那么 $V$ 中每个长度为 $\dim V$ 的规范正交向量组都是 $V$ 的规范正交基.

**证明**　由 6.25，$V$ 中每个规范正交向量组都是线性无关的. 于是每个这样的组，只要长度恰好，就是一个基——见 2.38. ■

---

**6.29　例：$\mathbf{F}^4$ 的一个规范正交基**

上面提到，标准基是 $\mathbf{F}^4$ 的规范正交基. 我们现在说明

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

转下页

同样是 $\mathbf{F}^4$ 的规范正交基.

我们有

$$\left\|\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)\right\| = \sqrt{\frac{1}{2^2}+\frac{1}{2^2}+\frac{1}{2^2}+\frac{1}{2^2}} = 1.$$

类似地，上述组中其余三个向量的范数也为 1.

注意到

$$\left\langle\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)\right\rangle = \frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\left(-\frac{1}{2}\right)+\frac{1}{2}\cdot\left(-\frac{1}{2}\right) = 0.$$

类似地，上述组中任意两个不同向量的内积都等于 0.

于是上述组是规范正交的. 因为该组是 4 维向量空间 $\mathbf{F}^4$ 中长度为 4 的规范正交组，所以由 6.28 可得该组是 $\mathbf{F}^4$ 的规范正交基.

---

一般来说，给定 $V$ 的基 $e_1,\dots,e_n$ 与向量 $v\in V$，我们知道存在一组标量 $a_1,\dots,a_n\in\mathbf{F}$ 满足

$$v = a_1e_1+\cdots+a_ne_n.$$

对于 $V$ 的一个任取的基来说，计算出满足上述等式的数 $a_1,\dots,a_n$ 可能很麻烦. 然而，下面结论表明，若取的是规范正交基，各系数的计算就很容易了——只需取 $a_k=\langle v,e_k\rangle$.

下面结论中，$\langle v,e_1\rangle,\dots,\langle v,e_n\rangle$ 充当 $\mathbf{F}^n$ 中向量坐标的角色，从而使得维数为 $n$ 的内积空间看起来就像 $\mathbf{F}^n$ 一样. 请注意观察下面结论是如何体现这点的.

下面关于 $\|v\|$ 的公式被称为帕塞瓦尔恒等式（Parseval's identity），它于 1799 年发表在研究傅里叶级数（Fourier series）的文章中.

**6.30 将向量写成规范正交基的线性组合**

设 $e_1,\dots,e_n$ 是 $V$ 的规范正交基且 $u,v\in V$. 那么

(a) $v=\langle v,e_1\rangle e_1+\cdots+\langle v,e_n\rangle e_n$,

(b) $\|v\|^2=|\langle v,e_1\rangle|^2+\cdots+|\langle v,e_n\rangle|^2$,

(c) $\langle u,v\rangle=\langle u,e_1\rangle\overline{\langle v,e_1\rangle}+\cdots+\langle u,e_n\rangle\overline{\langle v,e_n\rangle}$.

**证明** 因为 $e_1,\dots,e_n$ 是 $V$ 的基，所以存在标量 $a_1,\dots,a_n$ 满足

$$v=a_1e_1+\cdots+a_ne_n.$$

因为 $e_1,\dots,e_n$ 是规范正交的，所以将上式两端同时与 $e_k$ 作内积，可得 $\langle v,e_k\rangle=a_k$. 于是 (a) 成立.

由 (a) 和 6.24 立刻可知 (b) 成立.

将 $u$ 同时与式 (a) 两侧作内积，并利用内积在第二个位置上的共轭线性【6.6 (d) 和 6.6 (e)】即可得 (c). ■

**6.31 例：求出一个线性组合的系数**

设我们想将向量 $(1,2,4,7)\in\mathbf{F}^4$ 写成 $\mathbf{F}^4$ 中规范正交基（见例 6.29）

$$\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right),\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right),\left(-\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)$$

的线性组合．如果我们用的是非规范正交基，我们一般需要解含有四个方程、四个未知数的线性方程组来得到系数，而此处我们只需求四个内积并利用 6.30 (a)，即可得 $(1,2,4,7)$ 等于

$$7\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)-4\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)+\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)+2\left(-\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right).$$

既然我们懂得了规范正交基很有用，那么我们怎么求出它们呢？例如，带有例 6.3 (c) 中内积的 $\mathcal{P}_m(\mathbf{R})$ 具有规范正交基吗？接下来这个结论会将我们引向这些问题的答案．

下面证明中所用的算法称作**格拉姆-施密特过程**．它给出了将一个线性无关组转化成与之张成空间相同的规范正交组的方法．

约根·格拉姆（Jørgen Gram，1850-1916）和埃哈德·施密特（Erhard Schmidt，1876-1959）推广了这个构造规范正交组的算法．

**6.32 格拉姆-施密特过程（Gram-Schmidt procedure）**

设 $v_1,\ldots,v_m$ 是 $V$ 中的线性无关向量组．令 $f_1=v_1$．对 $k=2,\ldots,m$，依次定义 $f_k$ 为

$$f_k=v_k-\frac{\langle v_k,f_1\rangle}{\|f_1\|^2}f_1-\cdots-\frac{\langle v_k,f_{k-1}\rangle}{\|f_{k-1}\|^2}f_{k-1}.$$

对每个 $k=1,\ldots,m$，令 $e_k=\dfrac{f_k}{\|f_k\|}$．那么 $e_1,\ldots,e_m$ 是 $V$ 中的规范正交向量组，且对每个 $k=1,\ldots,m$ 满足

$$\operatorname{span}(v_1,\ldots,v_k)=\operatorname{span}(e_1,\ldots,e_k).$$

**证明** 我们通过对 $k$ 用归纳法来证明欲证结论成立．首先从 $k=1$ 开始，注意因为 $e_1=f_1/\|f_1\|$ 所以我们有 $\|e_1\|=1$；另外，$\operatorname{span}(v_1)=\operatorname{span}(e_1)$ 因为 $e_1$ 是 $v_1$ 的非零倍．

设 $1<k\leq m$，并设由 6.32 构造出的组 $e_1,\ldots,e_{k-1}$ 是规范正交组且满足下式：

$$\operatorname{span}(v_1,\ldots,v_{k-1})=\operatorname{span}(e_1,\ldots,e_{k-1}).\tag{6.33}$$

因为 $v_1,\ldots,v_m$ 是线性无关的，所以我们有 $v_k\notin\operatorname{span}(v_1,\ldots,v_{k-1})$．于是 $v_k\notin\operatorname{span}(e_1,\ldots,e_{k-1})=\operatorname{span}(f_1,\ldots,f_{k-1})$，这意味着 $f_k\neq 0$．于是 6.32 中定义 $e_k$ 时就不会出现除以零的问题．将一个向量除以其范数会得到范数为 1 的向量，因而 $\|e_k\|=1$．

令 $j\in\{1,\ldots,k-1\}$．那么

$$\begin{aligned}\langle e_k,e_j\rangle&=\frac{1}{\|f_k\|\|f_j\|}\langle f_k,f_j\rangle\\&=\frac{1}{\|f_k\|\|f_j\|}\left\langle v_k-\frac{\langle v_k,f_1\rangle}{\|f_1\|^2}f_1-\cdots-\frac{\langle v_k,f_{k-1}\rangle}{\|f_{k-1}\|^2}f_{k-1},f_j\right\rangle\\&=\frac{1}{\|f_k\|\,\|f_j\|}\big(\langle v_k,f_j\rangle-\langle v_k,f_j\rangle\big)\\&=0.\end{aligned}$$

于是 $e_1,\ldots,e_k$ 是规范正交组．

由 6.32 给出的 $e_k$ 定义，可见 $v_k \in \operatorname{span}(e_1,\ldots,e_k)$. 将其与式 (6.33) 相结合可得

$$\operatorname{span}(v_1,\ldots,v_k) \subseteq \operatorname{span}(e_1,\ldots,e_k).$$

上式中的两个组都是线性无关的（各 $v$ 线性无关是由于前提条件，各 $e$ 线性无关是由于它们规范正交和 6.25）. 于是这两个子空间的维数都为 $k$，因此它们相等，这就完成了归纳步骤的证明，由此完成了整个证明. ■

---

**6.34 例：$\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基**

假设我们在 $\mathcal{P}_2(\mathbf{R})$ 上定义内积为：对于所有 $p,q \in \mathcal{P}_2(\mathbf{R})$，

$$\langle p,q\rangle = \int_{-1}^{1} pq.$$

这样 $\mathcal{P}_2(\mathbf{R})$ 便成为内积空间. 我们知道 $1,x,x^2$ 是 $\mathcal{P}_2(\mathbf{R})$ 的基，但不是规范正交基. 我们将通过把格拉姆-施密特过程应用在 $v_1=1, v_2=x, v_3=x^2$ 上来求出 $\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基.

首先，取 $f_1=v_1=1$. 于是 $\|f_1\|^2=\int_{-1}^{1}1=2$. 则由 6.32 中的公式得

$$f_2 = v_2 - \frac{\langle v_2,f_1\rangle}{\|f_1\|^2}f_1 = x - \frac{\langle x,1\rangle}{\|f_1\|^2} = x,$$

式中最后一个等号成立是由于 $\langle x,1\rangle=\int_{-1}^{1}t\,\mathrm{d}t=0$.

由上述 $f_2$ 的表达式得 $\|f_2\|^2=\int_{-1}^{1}t^2\mathrm{d}t=\frac{2}{3}$. 现在由 6.32 中的公式得

$$f_3 = v_3 - \frac{\langle v_3,f_1\rangle}{\|f_1\|^2}f_1 - \frac{\langle v_3,f_2\rangle}{\|f_2\|^2}f_2 = x^2 - \frac{1}{2}\langle x^2,1\rangle - \frac{3}{2}\langle x^2,x\rangle x = x^2-\frac{1}{3}.$$

由上述 $f_3$ 的表达式得

$$\|f_3\|^2 = \int_{-1}^{1}\left(t^2-\frac{1}{3}\right)^2\mathrm{d}t = \int_{-1}^{1}\left(t^4-\frac{2}{3}t^2+\frac{1}{9}\right)\mathrm{d}t = \frac{8}{45}.$$

现在将 $f_1,f_2,f_3$ 都除以其范数，可得规范正交组

$$\sqrt{\frac{1}{2}},\ \sqrt{\frac{3}{2}}x,\ \sqrt{\frac{45}{8}}\left(x^2-\frac{1}{3}\right).$$

这个规范正交组的长度是 3，恰等于 $\mathcal{P}_2(\mathbf{R})$ 的维数. 因此该规范正交组就是 $\mathcal{P}_2(\mathbf{R})$ 的规范正交基（由 6.28）.

---

现在我们可以回答有关规范正交基存在性的问题了.

**6.35 规范正交基的存在性**

每个有限维内积空间都有规范正交基.

**证明** 设 $V$ 是有限维的. 选取 $V$ 的一个基，对其应用格拉姆-施密特过程（6.32），可得长为 $\dim V$ 的规范正交组. 由 6.28，这个规范正交组就是 $V$ 的一个规范正交基. ■

有时我们不仅需要知道规范正交基存在，还要知道每个规范正交组都能被扩充成一个规范正交基. 在下面的推论中，格拉姆-施密特过程表明，这样的扩充总是可以完成的.

**6.36 每个规范正交组都可被扩充为规范正交基**

设 $V$ 是有限维的. 那么 $V$ 中每个规范正交向量组都能被扩充为 $V$ 的一个规范正交基.

**证明** 设 $e_1, \ldots, e_m$ 是 $V$ 中的一个规范正交向量组. 那么 $e_1, \ldots, e_m$ 是线性无关的（由 6.25）. 因此该组可被扩充为 $V$ 的一个基 $e_1, \ldots, e_m, v_1, \ldots, v_n$（见 2.32）. 现在对 $e_1, \ldots, e_m, v_1, \ldots, v_n$ 应用格拉姆-施密特过程（6.32），可得规范正交组

$$e_1, \ldots, e_m, f_1, \ldots, f_n.$$

格拉姆-施密特过程所用的公式保持组中前 $m$ 个向量不变，因为这些向量已经是规范正交的了. 由 6.28 可知上述组是 $V$ 的规范正交基. ■

回忆一下，称一个矩阵为上三角矩阵，如果它形如

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix},$$

其中的 0 指的是对角线之下的所有元素均为 0，星号代表对角线及其上方的各元素.

在上章中，我们给出了一个算子关于某个基具有上三角矩阵的充要条件（见 5.44）. 既然我们现在研究内积空间，我们就想知道是否存在一个**规范正交**基，使得算子关于它存在上三角矩阵. 下面结论说明，算子关于某规范正交基具有上三角矩阵的条件，和关于一个任取的基具有上三角矩阵的条件是一样的.

**6.37 关于某个规范正交基有上三角矩阵**

设 $V$ 是有限维的，$T \in \mathcal{L}(V)$. 那么 $T$ 关于 $V$ 的某个规范正交基有上三角矩阵，当且仅当 $T$ 的最小多项式等于 $(z - \lambda_1) \cdots (z - \lambda_m)$，其中 $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

**证明** 设 $T$ 关于 $V$ 的某个基 $v_1, \ldots, v_n$ 具有上三角矩阵. 于是对各 $k = 1, \ldots, n$，$\operatorname{span}(v_1, \ldots, v_k)$ 在 $T$ 下是不变的（见 5.39）.

对 $v_1, \ldots, v_n$ 应用格拉姆-施密特过程，可得 $V$ 的一个规范正交基 $e_1, \ldots, e_n$. 因为

$$\operatorname{span}(e_1, \ldots, e_k) = \operatorname{span}(v_1, \ldots, v_k)$$

对每个 $k$ 都成立（见 6.32），所以我们可推断出，对各 $k = 1, \ldots, n$，$\operatorname{span}(e_1, \ldots, e_k)$ 在 $T$ 下是不变的. 从而由 5.39，$T$ 关于规范正交基 $e_1, \ldots, e_n$ 具有上三角矩阵. 现在利用 5.44 即可完成证明. ■

对于复向量空间，下面结论是上述结论的重要应用. 舒尔定理适用于多个算子的版本见于习题 20.

伊赛·舒尔（Issai Schur, 1875-1941）于 1909 年发表了下述结论的证明.

**6.38 舒尔定理（Schur's theorem）**

有限维复内积空间上的每个算子都关于某个规范正交基有上三角矩阵.

**证明** 由代数基本定理的版本二（4.13）和 6.37 即可证明. ■

## 内积空间上的线性泛函

因为映射至标量域 $\mathbf{F}$ 的线性映射扮演特殊的角色，所以我们在 3F 节中，为它们以及它们构成的向量空间起了特殊的名称. 因为你可能跳过 3F 节，所以下面重述一下这两个定义.

**6.39 定义：线性泛函（linear functional），对偶空间（dual space）、$V'$**

- $V$ 上的一个**线性泛函**是一个从 $V$ 到 $\mathbf{F}$ 的线性映射.
- $V$ 的**对偶空间**记作 $V'$，是 $V$ 上全体线性泛函所构成的向量空间. 换言之，$V' = \mathcal{L}(V, \mathbf{F})$. ♣

**6.40 例：$\mathbf{F}^3$ 上的线性泛函**

定义为

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

的函数 $\varphi : \mathbf{F}^3 \to \mathbf{F}$ 是 $\mathbf{F}^3$ 上的线性泛函. 我们可将该泛函写成如下形式：对每个 $z \in \mathbf{F}^3$，

$$\varphi(z) = \langle z, w \rangle$$

其中 $w = (2, -5, 1)$.

**6.41 例：$\mathcal{P}_5(\mathbf{R})$ 上的线性泛函**

定义为

$$\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\mathrm{d}t$$

的函数 $\varphi : \mathcal{P}_5(\mathbf{R}) \to \mathbf{R}$ 是 $\mathcal{P}_5(\mathbf{R})$ 上的线性泛函.

如果 $v \in V$，那么将 $u$ 对应至 $\langle u, v \rangle$ 的映射是 $V$ 上的线性泛函. 下面结论说明，$V$ 上每个线性泛函都可写成该形式. 例如，对于例 6.40，我们可取 $v = (2, -5, 1)$.

为纪念弗里杰什·里斯（Frigyes Riesz，1880-1956），下面结论用他的名字命名. 里斯在 20 世纪早期证明了几个类似于下述结论的定理.

设我们通过定义 $\langle p, q \rangle = \int_{-1}^{1} pq$ 使向量空间 $\mathcal{P}_5(\mathbf{R})$ 成为内积空间. 令 $\varphi$ 定义如例 6.41. 我们没法一眼看出存在 $q \in \mathcal{P}_5(\mathbf{R})$ 使得

$$\int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\mathrm{d}t = \langle p, q \rangle$$

对每个 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立【我们不能取 $q(t) = \cos(\pi t)$，因为该 $q$ 不属于 $\mathcal{P}_5(\mathbf{R})$】. 下面结论却能给出一个惊人的结果，那就是的确存在多项式 $q \in \mathcal{P}_5(\mathbf{R})$ 使得上式对于所有 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立.

**6.42 里斯表示定理（Riesz representation theorem）**

设 $V$ 是有限维的，$\varphi$ 是 $V$ 上的线性泛函. 那么存在唯一的向量 $v \in V$，使得对每个 $u \in V$ 都有

$$\varphi(u) = \langle u, v \rangle.$$

♡

**证明** 首先我们证明存在向量 $v \in V$ 使得对每个 $u \in V$ 有 $\varphi(u) = \langle u, v \rangle$. 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基. 那么对每个 $u \in V$ 有

$$\begin{aligned}\varphi(u) &= \varphi(\langle u, e_1 \rangle e_1 + \cdots + \langle u, e_n \rangle e_n) \\ &= \langle u, e_1 \rangle \varphi(e_1) + \cdots + \langle u, e_n \rangle \varphi(e_n) \\ &= \left\langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n \right\rangle,\end{aligned}$$

其中第一个等号源于 6.30 (a). 从而，设

$$v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n, \tag{6.43}$$

我们就可得对每个 $u \in V$，都有 $\varphi(u) = \langle u, v \rangle$，存在性得证.

下面我们证明上述表示式中的向量 $v \in V$ 是唯一的. 设 $v_1, v_2 \in V$ 使得对于每个 $u \in V$，有

$$\varphi(u) = \langle u, v_1 \rangle = \langle u, v_2 \rangle.$$

那么对每个 $u \in V$，有

$$0 = \langle u, v_1 \rangle - \langle u, v_2 \rangle = \langle u, v_1 - v_2 \rangle.$$

取 $u = v_1 - v_2$ 就可得 $v_1 - v_2 = 0$. 于是 $v_1 = v_2$，这就完成了原命题唯一性部分的证明. ■

---

**6.44 例：里斯表示定理的计算演示**

假设我们想求出多项式 $q \in \mathcal{P}_2(\mathbf{R})$ 使得

$$\int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\mathrm{d}t = \int_{-1}^{1} pq \tag{6.45}$$

对所有多项式 $p \in \mathcal{P}_2(\mathbf{R})$ 都成立. 为此，我们定义对于 $p, q \in \mathcal{P}_2(\mathbf{R})$，$\langle p, q \rangle$ 等于上式右侧，从而使 $\mathcal{P}_2(\mathbf{R})$ 成为内积空间. 注意到上式左侧并不等于 $p$ 与函数 $t \mapsto \cos(\pi t)$ 在 $\mathcal{P}_2(\mathbf{R})$ 上的内积，因为后者并不是多项式.

定义 $\mathcal{P}_2(\mathbf{R})$ 上的线性泛函 $\varphi$ 使得对于每个 $p \in \mathcal{P}_2(\mathbf{R})$ 都有

$$\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\mathrm{d}t.$$

现在，选用例 6.34 中的规范正交基并运用里斯表示定理证明过程中的式 (6.43)，我们就能得出，若 $p \in \mathcal{P}_2(\mathbf{R})$，那么 $\varphi(p) = \langle p, q \rangle$，其中

$$\begin{aligned}q(x) = {}&\left(\int_{-1}^{1} \sqrt{\frac{1}{2}} \cos(\pi t)\mathrm{d}t\right)\sqrt{\frac{1}{2}} + \left(\int_{-1}^{1} \sqrt{\frac{3}{2}}\, t \cos(\pi t)\mathrm{d}t\right)\sqrt{\frac{3}{2}}\, x \\ &+ \left(\int_{-1}^{1} \sqrt{\frac{45}{8}}\Big(t^2 - \frac{1}{3}\Big)\cos(\pi t)\mathrm{d}t\right)\sqrt{\frac{45}{8}}\Big(x^2 - \frac{1}{3}\Big).\end{aligned}$$

运用一些微积分的知识，即可解出

$$q(x) = \frac{15}{2\pi^2}(1 - 3x^2).$$

运用相同的步骤，即可说明如果想让 $q \in \mathcal{P}_5(\mathbf{R})$ 使得式 (6.45) 对所有 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立，那么应取

$$q(x) = \frac{105}{8\pi^4}\Big((27 - 2\pi^2) + (24\pi^2 - 270)x^2 + (315 - 30\pi^2)x^4\Big).$$

设 $V$ 是有限维的且 $\varphi$ 是 $V$ 上的线性泛函. 那么式 (6.43) 给出了使得

$$\varphi(u)=\langle u,v\rangle$$

对所有 $u\in V$ 都成立的向量 $v$ 的公式. 具体而言, 我们有

$$v=\overline{\varphi(e_1)}e_1+\cdots+\overline{\varphi(e_n)}e_n.$$

上式右侧似乎同时依赖于规范正交基 $e_1,\ldots,e_n$ 与 $\varphi$. 然而, 6.42 告诉我们, $v$ 由 $\varphi$ 唯一确定. 于是上式右端同样与 $V$ 的规范正交基 $e_1,\ldots,e_n$ 的选取无关.

里斯表示定理的另外两种不同证法见于 6.58 和 6C 节习题 13.

## 习题 6B

**1** 设 $e_1,\ldots,e_m$ 是 $V$ 中的一向量组, 使得

$$\|a_1e_1+\cdots+a_me_m\|^2=|a_1|^2+\cdots+|a_m|^2$$

对所有 $a_1,\ldots,a_m\in\mathbf{F}$ 成立. 证明 $e_1,\ldots,e_m$ 是规范正交组.

注 本题提出了 6.24 的逆命题.

**2** (a) 设 $\theta\in\mathbf{R}$. 证明:

$$(\cos\theta,\sin\theta),(-\sin\theta,\cos\theta)\quad\text{和}\quad(\cos\theta,\sin\theta),(\sin\theta,-\cos\theta)$$

都是 $\mathbf{R}^2$ 的规范正交基.

(b) 证明: $\mathbf{R}^2$ 的每个规范正交基都具有 (a) 中两者之一的形式.

**3** 设 $e_1,\ldots,e_m$ 是 $V$ 中的一规范正交组, 且 $v\in V$. 证明:

$$\|v\|^2=|\langle v,e_1\rangle|^2+\cdots+|\langle v,e_m\rangle|^2\iff v\in\operatorname{span}(e_1,\ldots,e_m).$$

**4** 设 $n$ 为正整数. 证明:

$$\frac{1}{\sqrt{2\pi}},\frac{\cos x}{\sqrt{\pi}},\frac{\cos 2x}{\sqrt{\pi}},\ldots,\frac{\cos nx}{\sqrt{\pi}},\frac{\sin x}{\sqrt{\pi}},\frac{\sin 2x}{\sqrt{\pi}},\ldots,\frac{\sin nx}{\sqrt{\pi}}$$

是 $C[-\pi,\pi]$ 中的一组规范正交向量. $C[-\pi,\pi]$ 是 $[-\pi,\pi]$ 上的连续实值函数构成的向量空间, 具有这样的内积——

$$\langle f,g\rangle=\int_{-\pi}^{\pi}fg.$$

提示 下列公式应该会有用.

$$(\sin x)(\cos y)=\frac{\sin(x-y)+\sin(x+y)}{2}$$
$$(\sin x)(\sin y)=\frac{\cos(x-y)-\cos(x+y)}{2}$$
$$(\cos x)(\cos y)=\frac{\cos(x-y)+\cos(x+y)}{2}$$

**5** 设 $f:[-\pi,\pi]\to\mathbf{R}$ 是连续的. 对任一非负整数 $k$, 定义

$$a_k=\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi}f(x)\cos(kx)\mathrm{d}x\quad\text{和}\quad b_k=\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi}f(x)\sin(kx)\mathrm{d}x.$$

证明:

$$\frac{a_0^2}{2}+\sum_{k=1}^{\infty}(a_k^2+b_k^2)\leq\int_{-\pi}^{\pi}f^2.$$

注 以上不等式实际上是对所有连续函数 $f:[-\pi,\pi]\to\mathbf{R}$ 成立的等式. 然而，证明这个不等式取等，涉及傅里叶级数方法，这超出了本书的范围.

**6** 设 $e_1,\ldots,e_n$ 是 $V$ 的规范正交基.

(a) 证明：如果 $v_1,\ldots,v_n$ 是 $V$ 中的向量，使得

$$\|e_k-v_k\|<\frac{1}{\sqrt{n}}$$

对任一 $k$ 成立，那么 $v_1,\ldots,v_n$ 是 $V$ 的基.

(b) 证明：存在 $v_1,\ldots,v_n\in V$，使得

$$\|e_k-v_k\|\leq\frac{1}{\sqrt{n}}$$

对任一 $k$ 成立，但 $v_1,\ldots,v_n$ 不是线性无关的.

注 本题的 (a) 指出，规范正交基带有适当小的扰动仍是基. 然后 (b) 说明，(a) 中不等式右侧的 $1/\sqrt{n}$ 不能更大了.

**7** 设 $T\in\mathcal{L}(\mathbf{R}^3)$ 关于基 $(1,0,0),(1,1,1),(1,1,2)$ 有上三角矩阵. 求 $\mathbf{R}^3$ 的一规范正交基，使得 $T$ 关于它有上三角矩阵.

**8** 定义内积——对所有 $p,q\in\mathcal{P}_2(\mathbf{R})$ 都有 $\langle p,q\rangle=\int_0^1 pq$，使 $\mathcal{P}_2(\mathbf{R})$ 成为内积空间.

(a) 对基 $1,x,x^2$ 应用格拉姆-施密特过程，得到 $\mathcal{P}_2(\mathbf{R})$ 的一规范正交基.

(b) $\mathcal{P}_2(\mathbf{R})$ 上的微分算子（该算子由 $p$ 得 $p'$）关于基 $1,x,x^2$（不是规范正交基）有上三角矩阵. 求 $\mathcal{P}_2(\mathbf{R})$ 上的微分算子关于 (a) 中所求规范正交基的矩阵，并验证其为上三角矩阵，这一点可以从 6.37 的证明中预料到.

**9** 设 $e_1,\ldots,e_m$ 是对 $V$ 中的线性无关组 $v_1,\ldots,v_m$ 应用格拉姆-施密特过程得到的结果. 证明：$\langle v_k,e_k\rangle>0$ 对任一 $k=1,\ldots,m$ 成立.

**10** 设 $v_1,\ldots,v_m$ 是 $V$ 中的线性无关组. 解释为什么：通过格拉姆-施密特过程的公式，得到的规范正交组 $e_1,\ldots,e_m$，是仅有的对任一 $k=1,\ldots,m$ 都有 $\langle v_k,e_k\rangle>0$ 且 $\operatorname{span}(v_1,\ldots,v_k)=\operatorname{span}(e_1,\ldots,e_k)$ 的规范正交组.

注 本题的结果将用于 7.58 的证明.

**11** 求一多项式 $q\in\mathcal{P}_2(\mathbf{R})$，使得 $p(\frac{1}{2})=\int_0^1 pq$ 对任一 $p\in\mathcal{P}_2(\mathbf{R})$ 都成立.

**12** 求一多项式 $q\in\mathcal{P}_2(\mathbf{R})$，使得

$$\int_0^1 p(x)\cos(\pi x)\mathrm{d}x=\int_0^1 pq$$

对任一 $p\in\mathcal{P}_2(\mathbf{R})$ 都成立.

**13** 证明：$V$ 中一组向量 $v_1,\ldots,v_m$ 线性相关，当且仅当 6.32 中的格拉姆-施密特公式对某个 $k\in\{1,\ldots,m\}$ 得到 $f_k=0$.

注 关于确定内积空间中一组向量是否线性相关，本题提供了高斯消元法的一种替代方法.

**14** 设 $V$ 是实内积空间，且 $v_1,\ldots,v_m$ 是 $V$ 中的一个线性无关向量组. 证明：$V$ 中存在恰好 $2^m$ 个规范正交向量组 $e_1,\ldots,e_m$ 使得

$$\operatorname{span}(v_1,\ldots,v_k)=\operatorname{span}(e_1,\ldots,e_k)$$

对所有 $k\in\{1,\ldots,m\}$ 成立.

**15** 设 $\langle\cdot,\cdot\rangle_1$ 和 $\langle\cdot,\cdot\rangle_2$ 是 $V$ 上的内积，使得 $\langle u,v\rangle_1=0$ 当且仅当 $\langle u,v\rangle_2=0$. 证明：存在正数 $c$ 使得 $\langle u,v\rangle_1=c\langle u,v\rangle_2$ 对任意 $u,v\in V$ 成立.

注 本题表明，如果两个内积下的正交向量对相同，那么其中一个内积是另一个的标量倍.

**16** 设 $V$ 是有限维的. 设 $\langle\cdot,\cdot\rangle_1$ 和 $\langle\cdot,\cdot\rangle_2$ 是 $V$ 上的内积，与之关联的范数是 $\|\cdot\|_1$ 和 $\|\cdot\|_2$. 证明存在正数 $c$ 使得 $\|v\|_1\leq c\|v\|_2$ 对任一 $v\in V$ 成立.

**17** 设 $\mathbf{F}=\mathbf{C}$ 且 $V$ 是有限维的. 证明：如果 $T$ 是 $V$ 上一算子，使得其唯一特征值是 1 且对所有 $v\in V$ 有 $\|Tv\|\leq\|v\|$，那么 $T$ 是恒等算子.

**18** 设 $u_1,\ldots,u_m$ 是 $V$ 中一线性无关组. 证明：存在 $v\in V$ 使得 $\langle u_k,v\rangle=1$ 对所有 $k\in\{1,\ldots,m\}$ 成立.

**19** 设 $v_1,\ldots,v_n$ 是 $V$ 的基. 证明：存在 $V$ 的一个基 $u_1,\ldots,u_n$ 使得

$$\langle v_j,u_k\rangle=\begin{cases}0 & \text{若 } j\neq k,\\ 1 & \text{若 } j=k.\end{cases}$$

**20** 设 $\mathbf{F}=\mathbf{C}$，$V$ 是有限维的，且 $\mathcal{E}\subseteq\mathcal{L}(V)$ 使得

$$ST=TS$$

对所有 $S,T\in\mathcal{E}$ 成立. 证明：存在 $V$ 的一规范正交基使得 $\mathcal{E}$ 的每个元素关于它都有上三角矩阵.

注 本题证明了 5E 节的习题 9 (b) 中的基可以选定为规范正交基，从而（在内积空间的背景下）加强了那道题的结论.

**21** 设 $\mathbf{F}=\mathbf{C}$，$V$ 是有限维的，$T\in\mathcal{L}(V)$，且 $T$ 的所有特征值绝对值都小于 1. 令 $\epsilon>0$，证明：存在一正整数 $m$ 使得 $\|T^m v\|\leq\epsilon\|v\|$ 对任一 $v\in V$ 成立.

**22** 设 $C[-1,1]$ 是区间 $[-1,1]$ 上的连续实值函数构成的向量空间，具有内积——对所有 $f,g\in C[-1,1]$ 成立的

$$\langle f,g\rangle=\int_{-1}^{1}fg.$$

令 $\varphi$ 为 $C[-1,1]$ 上的线性泛函，定义为 $\varphi(f)=f(0)$. 证明：不存在 $g\in C[-1,1]$ 使得

$$\varphi(f)=\langle f,g\rangle$$

对任一 $f\in C[-1,1]$ 成立.

注 本题表明，里斯表示定理（6.42）在对 $V$ 和 $\varphi$ 没有额外假设的无限维向量空间上不成立.

**23** 对所有 $u,v\in V$，定义 $d(u,v)=\|u-v\|$.

(a) 证明：$d$ 是 $V$ 上的度量.

(b) 证明：如果 $V$ 是有限维的，那么 $d$ 是 $V$ 上的完备度量（意即任一柯西序列都收敛）.

(c) 证明：$V$ 的每个有限维子空间都是 $V$（关于度量 $d$）的闭子集.

注 本题需要对度量空间比较熟悉.

**美国最高法院上的“正交”**

2010 年，法学教授理查德・弗里德曼上诉到美国最高法院的一个案子：

弗里德曼先生：我认为那个议题和这个议题完全正交. 因为州政府承认——
罗伯茨首席大法官：对不起，完全什么？
弗里德曼先生：正交. 就是直角、不相干、不相关.
罗伯茨首席大法官：噢.
斯卡利亚大法官：那个形容词是什么？我喜欢它.
弗里德曼先生：正交的.
罗伯茨首席大法官：正交的.
弗里德曼先生：对，对.
斯卡利亚大法官：正交的，嚯.（大笑）
肯尼迪大法官：我意识到这件案子给我们出了个难题.（大笑）

# 6C 正交补和最小化问题

## 正交补

**6.46 定义：正交补（orthogonal complement）、$U^\perp$**

若 $U$ 是 $V$ 的子集，那么 $U$ 的**正交补**，记作 $U^\perp$，是与 $U$ 中的每个向量都正交的所有 $V$ 中向量所构成的集合：

$$U^\perp = \{v \in V : 对于每个\ u \in U, \langle u, v\rangle = 0\}.$$

正交补 $U^\perp$ 同时依赖于 $V$ 和 $U$. 然而，我们总会由上下文明确得知内积空间 $V$ 的选取，于是我们可从记号中省去它.

**6.47 例：正交补**

- 若 $V = \mathbf{R}^3$，$U$ 是仅包含点 $(2,3,5)$ 的 $V$ 的子集，那么 $U^\perp$ 就是平面 $\{(x,y,z) \in \mathbf{R}^3 : 2x+3y+5z = 0\}$.
- 若 $V = \mathbf{R}^3$，$U$ 是平面 $\{(x,y,z) \in \mathbf{R}^3 : 2x+3y+5z = 0\}$，那么 $U^\perp$ 就是直线 $\{(2t,3t,5t) : t \in \mathbf{R}\}$.
- 更一般地说，若 $U$ 是 $\mathbf{R}^3$ 中过原点的平面，那么 $U^\perp$ 是过原点且垂直于 $U$ 的直线.
- 若 $U$ 是 $\mathbf{R}^3$ 中过原点的直线，那么 $U^\perp$ 是过原点且垂直于 $U$ 的平面.
- 若 $V = \mathbf{F}^5$ 且 $U = \{(a,b,0,0,0) \in \mathbf{F}^5 : a, b \in \mathbf{F}\}$，那么
$$U^\perp = \{(0,0,x,y,z) \in \mathbf{F}^5 : x,y,z \in \mathbf{F}\}.$$
- 若 $e_1,\dots,e_m,f_1,\dots,f_n$ 是 $V$ 的规范正交基，那么
$$\big(\operatorname{span}(e_1,\dots,e_m)\big)^\perp = \operatorname{span}(f_1,\dots,f_n).$$

我们首先介绍一些由定义即可推出的简单结论.

**6.48 正交补的性质**

(a) 若 $U$ 是 $V$ 的子集，那么 $U^\perp$ 是 $V$ 的子空间.

(b) $\{0\}^\perp = V$.

(c) $V^\perp = \{0\}$.

(d) 若 $U$ 是 $V$ 的子集，那么 $U \cap U^\perp \subseteq \{0\}$.

(e) 若 $G$ 和 $H$ 是 $V$ 的子集且 $G \subseteq H$，那么 $H^\perp \subseteq G^\perp$.

**证明**

(a) 设 $U$ 是 $V$ 的子集. 那么对每个 $u \in U$ 有 $\langle u, 0\rangle = 0$，从而 $0 \in U^\perp$. 设 $v, w \in U^\perp$. 若 $u \in U$，那么
$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle = 0 + 0 = 0.$$
于是 $v + w \in U^\perp$，表明 $U^\perp$ 对于加法封闭.

类似地，设 $\lambda \in \mathbf{F}$ 且 $v \in U^{\perp}$. 若 $u \in U$，那么

$$\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle = \overline{\lambda} \cdot 0 = 0.$$

于是 $\lambda v \in U^{\perp}$，表明 $U^{\perp}$ 对于标量乘法封闭. 于是 $U^{\perp}$ 是 $V$ 的子空间.

(b) 设 $v \in V$. 那么 $\langle 0, v\rangle = 0$，这表明 $v \in \{0\}^{\perp}$. 于是 $\{0\}^{\perp} = V$.

(c) 设 $v \in V^{\perp}$. 那么 $\langle v, v\rangle = 0$，这表明 $v = 0$. 于是 $V^{\perp} = \{0\}$.

(d) 设 $U$ 是 $V$ 的子集，且 $u \in U \cap U^{\perp}$. 那么 $\langle u, u\rangle = 0$，这表明 $u = 0$. 于是 $U \cap U^{\perp} \subseteq \{0\}$.

(e) 设 $G$ 和 $H$ 是 $V$ 的子集且 $G \subseteq H$. 设 $v \in H^{\perp}$. 那么对每个 $u \in H$ 有 $\langle u, v\rangle = 0$，这表明对每个 $u \in G$ 也有 $\langle u, v\rangle = 0$. 因而 $v \in G^{\perp}$. 于是 $H^{\perp} \subseteq G^{\perp}$. ■

回忆一下，若 $U$ 和 $W$ 是 $V$ 的子空间，那么如果 $V$ 中每个元素都能被唯一地写成 $U$ 中的一个向量加上 $W$ 中的一个向量的形式，那么 $V$ 就是 $U$ 和 $W$ 的直和（记为 $V = U \oplus W$）（见 1.41）. 这还等价于 $V = U + W$ 且 $U \cap W = \{0\}$（见 1.46）.

下面结论表明，$V$ 的每个有限维子空间都能引出 $V$ 的一个很自然的直和分解. 习题 16 举例说明了如果没有子空间 $U$ 是有限维这个前提条件，下面的结论可能没法成立.

**6.49 子空间及其正交补的直和**

设 $U$ 是 $V$ 的有限维子空间. 那么

$$V = U \oplus U^{\perp}.$$

**证明** 首先我们证明

$$V = U + U^{\perp}.$$

为此，设 $v \in V$. 令 $e_1, \ldots, e_m$ 是 $U$ 的规范正交基. 我们想将 $v$ 写成 $U$ 中的向量与正交于 $U$ 的向量之和.

我们有

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}. \tag{6.50}$$

令 $u$ 和 $w$ 定义如上式（在 6.26 的证明中我们也曾这样做）. 因为各 $e_k \in U$，所以我们可知 $u \in U$. 因为 $e_1, \ldots, e_m$ 是规范正交组，所以对各 $k = 1, \ldots, m$ 我们有

$$\begin{aligned}\langle w, e_k\rangle &= \langle v, e_k\rangle - \langle v, e_k\rangle \\ &= 0.\end{aligned}$$

从而 $w$ 与 $\operatorname{span}(e_1, \ldots, e_m)$ 中的每个向量都正交，这表明 $w \in U^{\perp}$. 因此我们可写出 $v = u + w$，其中 $u \in U$ 且 $w \in U^{\perp}$，这就完成了 $V = U + U^{\perp}$ 的证明.

由 6.48 (d)，我们可知 $U \cap U^{\perp} = \{0\}$，则由 $V = U + U^{\perp}$ 便可得 $V = U \oplus U^{\perp}$（见 1.46）. ■

现在我们可知如何由 $\dim U$ 计算出 $\dim U^{\perp}$ 了.

**6.51 正交补的维数**

设 $V$ 是有限维的，$U$ 是 $V$ 的子空间. 那么

$$\dim U^{\perp} = \dim V - \dim U.$$

**证明** 由 6.49 和 3.94 立刻可得 $\dim U^{\perp}$ 的公式. ■

下面结论是 6.49 的一个重要推论.

**6.52 正交补的正交补**

设 $U$ 是 $V$ 的一个有限维子空间. 那么

$$U = (U^{\perp})^{\perp}.$$

♡

**证明** 首先我们证明

$$U \subseteq (U^{\perp})^{\perp}. \tag{6.53}$$

为此, 设 $u \in U$. 那么对每个 $w \in U^{\perp}$ 都有 $\langle u, w\rangle = 0$ (由 $U^{\perp}$ 的定义即得). 因为 $u$ 与 $U^{\perp}$ 中的每个向量都正交, 所以我们有 $u \in (U^{\perp})^{\perp}$, 这就证明了式 (6.53).

为了证明这个包含关系反过来也成立, 设 $v \in (U^{\perp})^{\perp}$. 由 6.49, 我们可写出 $v = u + w$, 其中 $u \in U$ 且 $w \in U^{\perp}$, 则有 $v - u = w \in U^{\perp}$. 因为 $v \in (U^{\perp})^{\perp}$ 且 $u \in (U^{\perp})^{\perp}$【由式 (6.53)】, 故可得 $v - u \in (U^{\perp})^{\perp}$. 于是 $v - u \in U^{\perp} \cap (U^{\perp})^{\perp}$, 这表明 $v - u = 0$【由 6.48 (d)】, 进而表明 $v = u$, 因此 $v \in U$. 因此 $(U^{\perp})^{\perp} \subseteq U$, 结合式 (6.53) 即完成了整个证明. ■

> 习题 16 (a) 表明, 去掉 $U$ 是有限维这个前提条件后, 下面结论就不成立了.

设 $U$ 是 $V$ 的一个子空间, 且我们想证明 $U = V$. 在某些情况下, 最简单的证法是先证明与 $U$ 正交的向量只有 0, 然后运用下述结论. 例如, 下述结论在习题 4 中就很有用处.

**6.54 对 $V$ 的有限维子空间 $U$, 有 $U^{\perp} = \{0\} \iff U = V$**

设 $U$ 是 $V$ 的有限维子空间. 那么

$$U^{\perp} = \{0\} \iff U = V.$$

♡

**证明** 先设 $U^{\perp} = \{0\}$. 那么由 6.52, $U = (U^{\perp})^{\perp} = \{0\}^{\perp} = V$, 得证.

反之, 若 $U = V$, 则由 6.48 (c) 知 $U^{\perp} = V^{\perp} = \{0\}$. ■

我们现在为 $V$ 的每个有限维子空间 $U$ 定义算子 $P_U$.

**6.55 定义: 正交投影 (orthogonal projection)、$P_U$**

设 $U$ 是 $V$ 的一个有限维子空间. 将 $V$ 映成 $U$ 的**正交投影**是定义如下的算子 $P_U \in \mathcal{L}(V)$: 对每个 $v \in V$, 将其写成 $v = u + w$, 其中 $u \in U$ 且 $w \in U^{\perp}$, 然后令 $P_U v = u$.

♣

6.49 给出的直和分解式 $V = U \oplus U^{\perp}$ 表明, $v \in V$ 可以被唯一表示为 $v = u + w$ (其中 $u \in U$ 且 $w \in U^{\perp}$) 的形式. 由此可见 $P_U v$ 的定义是完善的. 请看 6.61 的证明中所附的示意图——你应该将这张描述 $P_U v$ 的图像牢记于心.

---

**6.56 例: 映成一维子空间的正交投影**

设 $u \in V$ 且 $u \neq 0$, $U$ 是 $V$ 的一维子空间, 定义为 $U = \operatorname{span}(u)$.

若 $v \in V$, 那么

$$v = \frac{\langle v, u\rangle}{\|u\|^2} u + \left( v - \frac{\langle v, u\rangle}{\|u\|^2} u \right),$$

转下页

式中右侧第一项属于 span$(u)$（从而属于 $U$），第二项正交于 $u$（从而属于 $U^\perp$）. 于是 $P_U v$ 等于上式右侧第一项. 换言之，对每个 $v \in V$，有

$$P_U v = \frac{\langle v, u\rangle}{\|u\|^2} u.$$

取 $v = u$，上式变为 $P_U u = u$；取 $v \in \{u\}^\perp$，上式则变为 $P_U v = 0$. 这是下面结果中 (b) 与 (c) 的具体例子.

**6.57 正交投影 $P_U$ 的性质**

设 $U$ 是 $V$ 的有限维子空间. 那么

(a) $P_U \in \mathcal{L}(V)$；

(b) 对每个 $u \in U$，都有 $P_U u = u$；

(c) 对每个 $w \in U^\perp$，都有 $P_U w = 0$；

(d) range $P_U = U$；

(e) null $P_U = U^\perp$；

(f) 对每个 $v \in V$，都有 $v - P_U v \in U^\perp$；

(g) $P_U^2 = P_U$；

(h) 对每个 $v \in V$，都有 $\|P_U v\| \le \|v\|$；

(i) 若 $e_1, \ldots, e_m$ 是 $U$ 的一个规范正交基且 $v \in V$，那么

$$P_U v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m.$$

**证明**

(a) 为证明 $P_U$ 是 $V$ 上的线性映射，设 $v_1, v_2 \in V$. 将 $v_1, v_2$ 写成

$$v_1 = u_1 + w_1 \quad 与 \quad v_2 = u_2 + w_2,$$

其中 $u_1, u_2 \in U$ 且 $w_1, w_2 \in U^\perp$. 从而 $P_U v_1 = u_1$ 且 $P_U v_2 = u_2$. 则有

$$v_1 + v_2 = (u_1 + u_2) + (w_1 + w_2),$$

其中 $u_1 + u_2 \in U$ 且 $w_1 + w_2 \in U^\perp$. 于是

$$P_U(v_1 + v_2) = u_1 + u_2 = P_U v_1 + P_U v_2.$$

类似地，设 $\lambda \in \mathbf{F}$ 且 $v \in V$. 将 $v$ 写成 $v = u + w$，其中 $u \in U$ 且 $w \in U^\perp$. 那么 $\lambda v = \lambda u + \lambda w$，其中 $\lambda u \in U$，$\lambda w \in U^\perp$. 因此 $P_U(\lambda v) = \lambda u = \lambda P_U v$.

因此 $P_U$ 是从 $V$ 到 $V$ 的线性映射.

(b) 设 $u \in U$. 我们可将 $u$ 写为 $u = u + 0$，其中 $u \in U$ 且 $0 \in U^\perp$. 于是 $P_U u = u$.

(c) 设 $w \in U^\perp$. 我们可将 $w$ 写为 $w = 0 + w$，其中 $0 \in U$ 且 $w \in U^\perp$. 于是 $P_U w = 0$.

(d) 由 $P_U$ 的定义得 range $P_U \subseteq U$. 又由 (b) 可知 $U \subseteq$ range $P_U$. 于是 range $P_U = U$.

(e) 由 (c) 可知 $U^\perp \subseteq$ null $P_U$. 为证明这个包含关系反过来也成立，注意到若 $v \in$ null $P_U$，那么 6.49 给出的分解式必为 $v = 0 + v$，其中 $0 \in U$ 且 $v \in U^\perp$. 于是 null $P_U \subseteq U^\perp$.

(f) 若 $v \in V$，且有 $v = u + w$（其中 $u \in U$ 且 $w \in U^\perp$），那么

$$v - P_U v = v - u = w \in U^\perp.$$

(g) 若 $v \in V$，且有 $v = u + w$（其中 $u \in U$ 且 $w \in U^{\perp}$），那么

$$(P_U^2)v = P_U(P_U v) = P_U u = u = P_U v.$$

(h) 若 $v \in V$，且有 $v = u + w$（其中 $u \in U$ 且 $w \in U^{\perp}$），那么

$$\|P_U v\|^2 = \|u\|^2 \le \|u\|^2 + \|w\|^2 = \|v\|^2,$$

式中最后一个等号源于毕达哥拉斯定理.

(i) 由 6.49 证明中的式 (6.50) 即可得 $P_U v$ 的表达式. ■

在上节中我们证明了里斯表示定理（6.42），它的关键内容是，有限维内积空间上的每个线性泛函，都可由与某个固定向量的内积式来表示. 了解不同的证明往往能带给我们新的理解. 于是，我们现在改用正交补（而非之前所用的规范正交基）来重新证明里斯表示定理.

下面重述的里斯表示定理使我们能将 $V$ 与 $V'$ 等同起来看. 我们只证明下述结论的“映成”部分，因为“一对一”部分的证明很常规——按 6.42 的证法即可.

这个新证法背后的直观想法是：若 $\varphi \in V'$，$v \in V$ 且 $\varphi(u) = \langle u, v\rangle$ 对所有 $u \in V$ 都成立，那么 $v \in (\operatorname{null}\varphi)^{\perp}$. 而由 6.51 和 3.21，$(\operatorname{null}\varphi)^{\perp}$ 是 $V$ 的一维子空间（除 $\varphi = 0$ 的平凡情形外）. 于是我们可通过取 $(\operatorname{null}\varphi)^{\perp}$ 中的任一非零元素并将其与一适当的标量相乘而得到 $v$. 下面证明正是这样做的.

**6.58 里斯表示定理再讨论**

设 $V$ 是有限维的. 对每个 $v \in V$，定义 $\varphi_v \in V'$ 为：对每个 $u \in V$，

$$\varphi_v(u) = \langle u, v\rangle.$$

那么 $v \mapsto \varphi_v$ 是将 $V$ 映成 $V'$ 的一对一函数.

**证明**

为证明 $v \mapsto \varphi_v$ 是满射，设 $\varphi \in V'$. 若 $\varphi = 0$，那么 $\varphi = \varphi_0$. 从而假设 $\varphi \ne 0$. 因此 $\operatorname{null}\varphi \ne V$，这表明 $(\operatorname{null}\varphi)^{\perp} \ne \{0\}$（由 6.49，其中取 $U = \operatorname{null}\varphi$）.

**注意：** 若 $\mathbf{F} = \mathbf{R}$，则函数 $v \mapsto \varphi_v$ 是从 $V$ 到 $V'$ 的线性映射. 然而，若 $\mathbf{F} = \mathbf{C}$，则该函数不是线性的，因为若 $\lambda \in \mathbf{C}$，则 $\varphi_{\lambda v} = \bar{\lambda}\varphi_v$.

令 $w \in (\operatorname{null}\varphi)^{\perp}$ 且 $w \ne 0$. 令

$$v = \frac{\overline{\varphi(w)}}{\|w\|^2} w. \tag{6.59}$$

那么 $v \in (\operatorname{null}\varphi)^{\perp}$，并且 $v \ne 0$（因为 $w \notin \operatorname{null}\varphi$）.

在式 (6.59) 两侧同取范数可得

$$\|v\| = \frac{|\varphi(w)|}{\|w\|}. \tag{6.60}$$

将 $\varphi$ 同时作用于式 (6.59) 两端并利用式 (6.60)，可得

$$\varphi(v) = \frac{|\varphi(w)|^2}{\|w\|^2} = \|v\|^2.$$

现在设 $u \in V$. 利用上式可得

$$u = \left(u - \frac{\varphi(u)}{\varphi(v)} v\right) + \frac{\varphi(u)}{\|v\|^2} v.$$

式中带括号的一项属于 null $\varphi$，因此与 $v$ 正交．于是将此式两边都与 $v$ 作内积可得

$$\langle u, v\rangle = \frac{\varphi(u)}{\|v\|^2}\langle v, v\rangle = \varphi(u).$$

因此 $\varphi = \varphi_v$，表明 $v \mapsto \varphi_v$ 是满射，则原命题得证．■

里斯表示定理的另一种不同证法见于习题 13.

## 最小化问题

这个最小化问题的解特别简洁，由此产生了内积空间在纯粹数学以外的很多重要应用.

我们常会遇到下面问题：给定 $V$ 的子空间 $U$ 与一点 $v \in V$，求出一点 $u \in U$ 使得 $\|v - u\|$ 尽可能小．下面结论表明，$u = P_U v$ 是该最小化问题的唯一解.

**6.61 到子空间的最短距离**

设 $U$ 是 $V$ 的有限维子空间，$v \in V$ 且 $u \in U$．那么

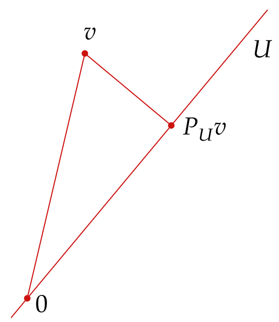
$$\|v - P_U v\| \le \|v - u\|.$$

进一步，上述不等式取得等号当且仅当 $u = P_U v$.

**证明** 我们有

$$\begin{aligned}\|v - P_U v\|^2 &\le \|v - P_U v\|^2 + \|P_U v - u\|^2 \qquad (6.62)\\ &= \|(v - P_U v) + (P_U v - u)\|^2\\ &= \|v - u\|^2,\end{aligned}$$

其中第一行成立是因为 $0 \le \|P_U v - u\|^2$，第二行源于毕达哥拉斯定理【可以运用它是因为由 6.57 (f) 知 $v - P_U v \in U^\perp$，又有 $P_U v - u \in U$】，简单计算即知第三行成立．两边开平方根就得到了我们欲证的不等式.




$P_U v$ 是 $U$ 中距离 $v$ 最近的点.

上面证明的不等式取得等号当且仅当式 (6.62) 取得等号，这等价于 $\|P_U v - u\| = 0$，又等价于 $u = P_U v$．■

我们常将上述结论与公式 6.57 (i) 相结合来计算出最小化问题的显式解，如下例所示.

---

**6.63 例：利用线性代数来逼近正弦函数**

假如我们想要求出次数不高于 5 的实系数多项式 $u$，使其在区间 $[-\pi, \pi]$ 上尽可能逼近正弦函数，这里的“逼近”指的是使

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \mathrm{d}x$$

尽可能小.

令 $C[-\pi, \pi]$ 表示定义在 $[-\pi, \pi]$ 上的全体连续实值函数所构成的实内积空间，其上内积定义为

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f g. \qquad (6.64)$$

转下页

令 $v \in C[-\pi, \pi]$ 是定义为 $v(x) = \sin x$ 的函数. 令 $U$ 表示由次数不高于 5 的所有实系数多项式所构成的 $C[-\pi, \pi]$ 的子空间. 现在可将问题重新表述为:

$$\text{求出 } u \in U \text{ 使得 } \|v - u\| \text{ 尽可能小.}$$

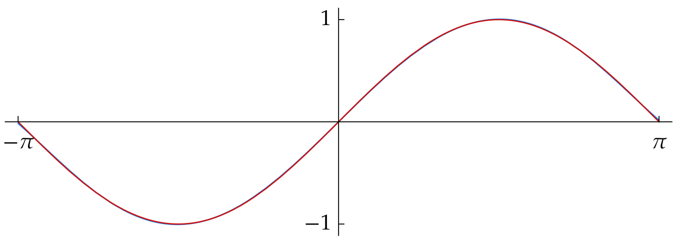
一台可演算积分的计算机对此很有用处.

为了计算出这个逼近问题的解,首先对 $U$ 的基 $1, x, x^2, x^3, x^4, x^5$ 用格拉姆-施密特过程【所用内积定义如式 (6.64)】,得到 $U$ 的一个规范正交基 $e_1, e_2, e_3, e_4, e_5, e_6$. 接着,再采用式 (6.64) 给出的内积并利用 6.57 (i) 来计算 $P_U v$(其中 $m = 6$). 计算得 $P_U v$ 就是定义如下的函数 $u$:

$$u(x) = 0.987862x - 0.155271x^3 + 0.00564312x^5, \tag{6.65}$$

这里我们把精确解中出现的含 $\pi$ 系数用足够准确的十进制近似值来代替. 由 6.61,上述多项式 $u$ 是利用不超过 5 次的多项式在区间 $[-\pi, \pi]$ 上对正弦函数的最佳逼近(此处“最佳逼近”是就 $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 dx$ 最小而言的).

为了观察逼近程度有多高,我们在下图中同时绘出正弦函数和式 (6.65) 给出的逼近函数 $u$ 在区间 $[-\pi, \pi]$ 上的图像.



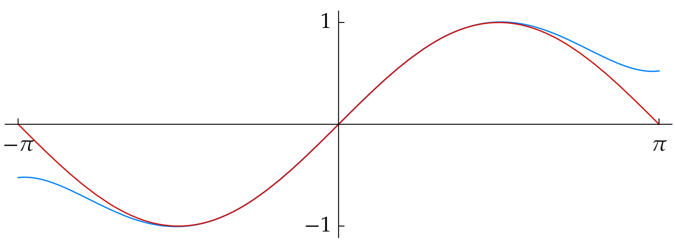

$[-\pi, \pi]$ 区间上正弦函数(红线)和式 (6.65) 给出的最佳逼近五次多项式 $u$(蓝线)的图像.

逼近式 (6.65) 准确到使得两函数的图像几乎一致——我们用肉眼只能看到一个图像!此处红色图像几乎覆盖掉了蓝色图像. 如果你在电子设备上查看这张图,你可以在 $\pi$ 或 $-\pi$ 附近将它放大至原先的 400%,以便观察两函数图像的细微差别.

用五次多项式来逼近正弦函数的另一种著名方法,就是利用定义如下的泰勒多项式(Taylor polynomial)$p$:

$$p(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}. \tag{6.66}$$

为了观察逼近程度有多高,我们在下图中同时绘出正弦函数和泰勒多项式 $p$ 在区间 $[-\pi, \pi]$ 上的图像.




$[-\pi, \pi]$ 区间上正弦函数(红线)和式 (6.66) 所示的泰勒多项式(蓝线)的图像.

转下页

在 $x$ 接近 0 时，泰勒多项式能很好地逼近 $\sin x$．然而上述图像表明，对于 $|x| > 2$，泰勒多项式就不那么准确了，特别是与式 (6.65) 相比．例如，取 $x = 3$，逼近式 (6.65) 对于 $\sin 3$ 的估计的误差约为 0.001，而泰勒多项式 (6.66) 对于 $\sin 3$ 的估计的误差约是 0.4．于是在 $x = 3$ 处，泰勒多项式的误差比式 (6.65) 误差大数百倍．线性代数帮助我们找到了正弦函数的新的逼近方法，这个方法改进了我们在微积分中学过的方法！

## 伪逆

设 $T \in \mathcal{L}(V, W)$ 且 $w \in W$．考虑这个问题：求出 $v \in V$ 使得

$$Tv = w.$$

例如，若 $V = \mathbf{F}^n$ 且 $W = \mathbf{F}^m$，那么上式可表示含有 $m$ 个线性方程，$n$ 个未知数 $v_1, \ldots, v_n$ 的方程组，其中 $v = (v_1, \ldots, v_n)$．

若 $T$ 可逆，那么上述方程的唯一解就是 $v = T^{-1}w$．然而，若 $T$ 不可逆，那么对于某些 $w \in W$，上述方程可能无解；对于另外一些 $w \in W$，上述方程可能有无穷多解．

当 $T$ 不可逆时，我们仍可尝试尽可能好地处理上述方程．例如，若上述方程无解，那么我们就不求解方程 $Tv - w = 0$，而是尝试解出使 $\|Tv - w\|$ 尽可能小的 $v \in V$．又例如，如果有无穷多个 $v \in V$ 满足上述方程，那么我们可以选出这些解中使得 $\|v\|$ 最小的那个．

伪逆就为我们尽可能好地求解上述方程提供了有力工具，即便 $T$ 并不可逆．为定义伪逆，我们需要下面的结论．

在接下来的两个证明中，我们将不加注解地使用这个结论：若 $V$ 是有限维的，且 $T \in \mathcal{L}(V, W)$，那么 $\operatorname{null} T$，$(\operatorname{null} T)^\perp$ 与 $\operatorname{range} T$ 都是有限维的．

**6.67 限制线性映射以获得既单又满的映射**

设 $V$ 是有限维的，且 $T \in \mathcal{L}(V, W)$．那么 $T|_{(\operatorname{null} T)^\perp}$ 是将 $(\operatorname{null} T)^\perp$ 映成 $\operatorname{range} T$ 的单射．

**证明** 设 $v \in (\operatorname{null} T)^\perp$ 且 $T|_{(\operatorname{null} T)^\perp} v = 0$．因此 $Tv = 0$，进而 $v \in (\operatorname{null} T) \cap (\operatorname{null} T)^\perp$，从而 $v = 0$【由 6.48 (d)】．因此 $\operatorname{null} T|_{(\operatorname{null} T)^\perp} = \{0\}$，这表明 $T|_{(\operatorname{null} T)^\perp}$ 是单射．

显然，$\operatorname{range} T|_{(\operatorname{null} T)^\perp} \subseteq \operatorname{range} T$．为证明这个包含关系反过来也成立，设 $w \in \operatorname{range} T$．因此存在 $v \in V$ 满足 $w = Tv$．由 6.49，存在 $u \in \operatorname{null} T$ 和 $x \in (\operatorname{null} T)^\perp$ 使得 $v = u + x$．则

$$T|_{(\operatorname{null} T)^\perp} x = Tx = Tv - Tu = w - 0 = w,$$

这表明 $w \in \operatorname{range} T|_{(\operatorname{null} T)^\perp}$．因此 $\operatorname{range} T \subseteq \operatorname{range} T|_{(\operatorname{null} T)^\perp}$，这就证明了 $\operatorname{range} T|_{(\operatorname{null} T)^\perp} = \operatorname{range} T$．■

现在我们可定义线性映射 $T$ 的伪逆 $T^\dagger$ 了．在下面定义（乃至之后的章节）中，将 $T|_{(\operatorname{null} T)^\perp}$ 看成把 $(\operatorname{null} T)^\perp$ 映成 $\operatorname{range} T$ 的可逆线性映射（上面证明了这一点）．

在 TEX 中输入 `T^\dagger` 即可打出 $T^\dagger$．

**6.68 定义：伪逆（pseudoinverse）、$T^\dagger$**

设 $V$ 是有限维的，$T \in \mathcal{L}(V,W)$. $T$ 的**伪逆** $T^\dagger \in \mathcal{L}(W,V)$ 是定义如下的从 $W$ 到 $V$ 的线性映射：对每个 $w \in W$，

$$T^\dagger w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T} w.$$

♣

回忆一下，若 $w \in (\operatorname{range} T)^\perp$，则 $P_{\operatorname{range} T} w = 0$；若 $w \in \operatorname{range} T$，则 $P_{\operatorname{range} T} w = w$. 于是，如果 $w \in (\operatorname{range} T)^\perp$，则 $T^\dagger w = 0$；若 $w \in \operatorname{range} T$，则 $T^\dagger w$ 是 $(\operatorname{null} T)^\perp$ 中唯一满足 $T(T^\dagger w) = w$ 的元素.

我们将看到，伪逆看上去很像一般的逆.

**6.69 伪逆的代数性质**

设 $V$ 是有限维的且 $T \in \mathcal{L}(V,W)$.

(a) 若 $T$ 可逆，则 $T^\dagger = T^{-1}$.

(b) $TT^\dagger = P_{\operatorname{range} T}$ = 将 $W$ 映成 $\operatorname{range} T$ 的正交投影.

(c) $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ = 将 $V$ 映成 $(\operatorname{null} T)^\perp$ 的正交投影.

♡

**证明**

(a) 设 $T$ 可逆. 那么 $(\operatorname{null} T)^\perp = V$，且 $\operatorname{range} T = W$. 从而 $T|_{(\operatorname{null} T)^\perp} = T$ 且 $P_{\operatorname{range} T}$ 是 $W$ 上的恒等算子. 因此 $T^\dagger = T^{-1}$.

(b) 设 $w \in \operatorname{range} T$. 于是

$$TT^\dagger w = T(T|_{(\operatorname{null} T)^\perp})^{-1} w = w = P_{\operatorname{range} T} w.$$

若 $w \in (\operatorname{range} T)^\perp$，那么 $T^\dagger w = 0$. 因此 $TT^\dagger w = 0 = P_{\operatorname{range} T} w$. 于是，在 $\operatorname{range} T$ 和 $(\operatorname{range} T)^\perp$ 上，$TT^\dagger$ 和 $P_{\operatorname{range} T}$ 都相同. 因此这两个线性映射相等（由 6.49）.

(c) 设 $v \in (\operatorname{null} T)^\perp$. 因为 $Tv \in \operatorname{range} T$，所以由 $T^\dagger$ 定义可知

$$T^\dagger (Tv) = (T|_{(\operatorname{null} T)^\perp})^{-1}(Tv) = v = P_{(\operatorname{null} T)^\perp} v.$$

若 $v \in \operatorname{null} T$，那么 $T^\dagger T v = 0 = P_{(\operatorname{null} T)^\perp} v$. 于是，在 $(\operatorname{null} T)^\perp$ 和 $\operatorname{null} T$ 上，$T^\dagger T$ 和 $P_{(\operatorname{null} T)^\perp}$ 都相同. 因此这两个线性映射相等（由 6.49）. ■

伪逆也被称作**摩尔-彭罗斯逆（Moore-Penrose inverse）**.

设 $T \in \mathcal{L}(V,W)$. 如果 $T$ 是满射，那么由上面结论的 (b) 可得，$TT^\dagger$ 是 $W$ 上的恒等算子. 如果 $T$ 是单射，那么由上面结论的 (c) 可得，$T^\dagger T$ 是 $V$ 上的恒等算子. 伪逆的更多代数性质见于习题 19 至 23.

对于 $T \in \mathcal{L}(V,W)$ 和 $w \in W$，我们现在回到求出满足方程 $Tv = w$ 的 $v \in V$ 这个问题上来. 前面提过，若 $T$ 是可逆的，那么 $v = T^{-1} w$ 就是方程的唯一解，但若 $T$ 不可逆，那么 $T^{-1}$ 没有定义. 然而，伪逆 $T^\dagger$ 是有定义的. 下面结论的 (a) 就会告诉我们，取 $v = T^\dagger w$ 可使 $Tv$ 尽可能接近 $w$. 于是伪逆给出了上述方程的所谓**最佳拟合（best fit）**.

结合下面的 (b) 以及 (a) 中不等式取等号的条件可知，在所有使得 $Tv$ 尽可能接近 $w$ 的向量 $v \in V$ 中，向量 $T^\dagger w$ 具有最小的范数.

**6.70 伪逆可给出最佳近似解或最优解**

设 $V$ 是有限维的，$T \in \mathcal{L}(V,W)$，$w \in W$.

(a) 若 $v \in V$，则

$$\|T(T^{\dagger}w) - w\| \leq \|Tv - w\|,$$

当且仅当 $v \in T^{\dagger}w + \operatorname{null} T$ 时，上式取得等号.

(b) 若 $v \in T^{\dagger}w + \operatorname{null} T$，则

$$\|T^{\dagger}w\| \leq \|v\|,$$

当且仅当 $v = T^{\dagger}w$ 时，上式取得等号.

**证明**

(a) 设 $v \in V$. 那么

$$Tv - w = (Tv - TT^{\dagger}w) + (TT^{\dagger}w - w).$$

上式第一个括号中的项属于 $\operatorname{range} T$. 因为算子 $TT^{\dagger}$ 是将 $W$ 映成 $\operatorname{range} T$ 的正交投影【由 6.69 (b)】，所以上式第二个括号中的项属于 $(\operatorname{range} T)^{\perp}$【见 6.57 (f)】.

于是，由毕达哥拉斯定理可知，上式第二个括号中的项的范数小于或等于 $\|Tv - w\|$，并且当且仅当上式第一个括号中的项等于 0 时方可取得等号. 因此上述不等式取得等号，当且仅当 $v - T^{\dagger}w \in \operatorname{null} T$，这等价于 $v \in T^{\dagger}w + \operatorname{null} T$. 这就完成了 (a) 的证明.

(b) 设 $v \in T^{\dagger}w + \operatorname{null} T$. 因此 $v - T^{\dagger}w \in \operatorname{null} T$. 我们有

$$v = (v - T^{\dagger}w) + T^{\dagger}w.$$

$T^{\dagger}$ 的定义表明，$T^{\dagger}w \in (\operatorname{null} T)^{\perp}$. 于是由毕达哥拉斯定理得 $\|T^{\dagger}w\| \leq \|v\|$，并且当且仅当 $v = T^{\dagger}w$ 时取得等号. ■

下章将给出 $T^{\dagger}$ 的一个计算公式（见 7.78）.

---

**6.71 例：$\mathbf{F}^4$ 到 $\mathbf{F}^3$ 的线性映射的伪逆**

设 $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ 定义为

$$T(a,b,c,d) = (a+b+c, 2c+d, 0).$$

该线性映射既不是单射也不是满射，但我们仍可计算其伪逆. 为此，首先注意到 $\operatorname{range} T = \{(x,y,0) : x, y \in \mathbf{F}\}$. 从而对每个 $(x,y,z) \in \mathbf{F}^3$，

$$P_{\operatorname{range} T}(x,y,z) = (x,y,0).$$

并且，

$$\operatorname{null} T = \{(a,b,c,d) \in \mathbf{F}^4 : a+b+c = 0 \text{ 且 } 2c+d = 0\}.$$

由 $\operatorname{null} T$ 中的两个向量构成的组 $(-1,1,0,0), (-1,0,1,-2)$ 张成 $\operatorname{null} T$，因为若 $(a,b,c,d) \in \operatorname{null} T$，那么

$$(a,b,c,d) = b(-1,1,0,0) + c(-1,0,1,-2).$$

转下页

又由于组 $(-1,1,0,0),(-1,0,1,-2)$ 线性无关，所以该组是 $\operatorname{null} T$ 的一个基.

现设 $(x,y,z)\in\mathbf{F}^3$. 那么

$$T^{\dagger}(x,y,z)=\left(T|_{(\operatorname{null}T)^{\perp}}\right)^{-1}P_{\operatorname{range}T}(x,y,z)=\left(T|_{(\operatorname{null}T)^{\perp}}\right)^{-1}(x,y,0). \tag{6.72}$$

上式右侧就是满足 $T(a,b,c,d)=(x,y,0)$ 和 $(a,b,c,d)\in(\operatorname{null}T)^{\perp}$ 的向量 $(a,b,c,d)\in\mathbf{F}^4$. 换言之，$a,b,c,d$ 必满足下列方程：

$$\begin{aligned} a+b+c&=x\\ 2c+d&=y\\ -a+b&=0\\ -a+c-2d&=0, \end{aligned}$$

其中前两个方程等价于方程 $T(a,b,c,d)=(x,y,0)$，后两个方程则是由于 $(a,b,c,d)$ 正交于 $\operatorname{null}T$ 的基向量 $(-1,1,0,0),(-1,0,1,-2)$. 将 $x,y$ 视为常数，并将 $a,b,c,d$ 视为未知量，上述方程组便含有四个方程、四个未知数，解之得

$$a=\frac{1}{11}(5x-2y),\ b=\frac{1}{11}(5x-2y),\ c=\frac{1}{11}(x+4y),\ d=\frac{1}{11}(-2x+3y).$$

因此由式 (6.72)，

$$T^{\dagger}(x,y,z)=\frac{1}{11}(5x-2y,5x-2y,x+4y,-2x+3y).$$

由上述 $T^{\dagger}$ 的计算公式可得，对所有 $(x,y,z)\in\mathbf{F}^3$，有 $TT^{\dagger}(x,y,z)=(x,y,0)$，这形象阐释了 6.69 (b) 中的 $TT^{\dagger}=P_{\operatorname{range}T}$ 这个式子.

## 习题 6C

**1** 设 $v_1,\dots,v_m\in V$. 证明：

$$\{v_1,\dots,v_m\}^{\perp}=\bigl(\operatorname{span}(v_1,\dots,v_m)\bigr)^{\perp}.$$

**2** 设 $U$ 是 $V$ 的子空间，且有一个基 $u_1,\dots,u_m$，且

$$u_1,\dots,u_m,v_1,\dots,v_n$$

是 $V$ 的基. 证明：如果对以上 $V$ 的基应用格拉姆-施密特过程，得到组 $e_1,\dots,e_m,f_1,\dots,f_n$；那么 $e_1,\dots,e_m$ 是 $U$ 的规范正交基，而 $f_1,\dots,f_n$ 是 $U^{\perp}$ 的规范正交基.

**3** 设 $U$ 是 $\mathbf{R}^4$ 的子空间，其定义为

$$U=\operatorname{span}\bigl((1,2,3,-4),(-5,4,3,2)\bigr).$$

求 $U$ 的一规范正交基，与 $U^{\perp}$ 的一规范正交基.

**4** 设 $e_1,\dots,e_n$ 是 $V$ 中一组向量，其满足：$\|e_k\|=1$ 对任一 $k=1,\dots,n$ 都成立以及

$$\|v\|^2=|\langle v,e_1\rangle|^2+\cdots+|\langle v,e_n\rangle|^2$$

对所有 $v\in V$ 成立. 证明：$e_1,\dots,e_n$ 是 $V$ 的规范正交基.

注　本题提出了 6.30 (b) 的逆命题.

**5** 设 $V$ 是有限维的，且 $U$ 为 $V$ 的子空间．证明：$P_{U^\perp} = I - P_U$，其中 $I$ 是 $V$ 中的恒等算子．

**6** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V, W)$．证明：

$$T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T}T.$$

**7** 设 $X$ 和 $Y$ 是 $V$ 的有限维子空间．证明：$P_X P_Y = 0$，当且仅当 $\langle x, y\rangle = 0$ 对所有 $x \in X$ 和 $y \in Y$ 成立．

**8** 设 $U$ 是 $V$ 的有限维子空间，且 $v \in V$．定义线性映射 $\varphi : U \to \mathbf{F}$ 为：

$$\varphi(u) = \langle u, v\rangle$$

对所有 $u \in U$ 成立．将里斯表示定理（6.42）应用于内积空间 $U$ 可得，存在唯一的向量 $w \in U$，使得

$$\varphi(u) = \langle u, w\rangle$$

对所有 $u \in U$ 成立．证明：$w = P_U v$．

**9** 设 $V$ 是有限维的．设 $P \in \mathcal{L}(V)$ 使得 $P^2 = P$ 且 null $P$ 中任一向量都正交于 range $P$ 中任一向量．证明：存在 $V$ 的子空间 $U$ 使得 $P = P_U$．

**10** 设 $V$ 是有限维的，$P \in \mathcal{L}(V)$ 使得 $P^2 = P$ 且

$$\|Pv\| \leq \|v\|$$

对任一 $v \in V$ 成立．证明存在 $V$ 的子空间 $U$ 使得 $P = P_U$．

**11** 设 $T \in \mathcal{L}(V)$ 且 $U$ 是 $V$ 的有限维子空间．证明：

$$U \text{ 在 } T \text{ 下不变} \iff P_U T P_U = T P_U.$$

**12** 设 $V$ 是有限维的，$T \in \mathcal{L}(V)$，且 $U$ 是 $V$ 的子空间．证明：

$$U \text{ 和 } U^\perp \text{ 都在 } T \text{ 下不变} \iff P_U T = T P_U.$$

**13** 设 $\mathbf{F} = \mathbf{R}$，且 $V$ 是有限维的．对任一 $v \in V$，令 $\varphi_v$ 表示 $V$ 上的线性泛函，其定义为

$$\varphi_v(u) = \langle u, v\rangle$$

对所有 $u \in V$ 成立．

(a) 证明：$v \mapsto \varphi_v$ 是从 $V$ 到 $V'$ 的单的线性映射．

(b) 利用 (a) 并结合对维数的计算，证明 $v \mapsto \varphi_v$ 是把 $V$ 映成 $V'$ 的同构．

注　本题的目的是，给出里斯表示定理（6.42 和 6.58）在 $\mathbf{F} = \mathbf{R}$ 条件下的另一证明．因此你不应该在解答中将里斯表示定理用作工具．

**14** 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基．解释为什么：在由里斯表示定理（6.58）提供的 $V'$ 与 $V$ 的等同关系下，$e_1, \ldots, e_n$ 的对偶基（见 3.112）是 $e_1, \ldots, e_n$．[3]

**15** 在 $\mathbf{R}^4$ 中，令

$$U = \operatorname{span}\big((1, 1, 0, 0), (1, 1, 1, 2)\big).$$

求 $u \in U$ 使得 $\|u - (1, 2, 3, 4)\|$ 尽可能小．

[3]译者完成这道题后留下了小小的提示：$e_k \mapsto \langle \cdot, e_k\rangle$．

**16** 设 $C[-1,1]$ 是区间 $[-1,1]$ 上的连续实值函数构成的向量空间，具有内积——对所有 $f,g \in C[-1,1]$ 成立的

$$\langle f,g\rangle = \int_{-1}^{1} fg.$$

令 $U$ 是 $C[-1,1]$ 的子空间，其定义为

$$U = \{f \in C[-1,1] : f(0) = 0\}.$$

(a) 证明：$U^{\perp} = \{0\}$.

(b) 证明：不假设有限维的情况下，6.49 和 6.52 不再成立.

**17** 求 $p \in \mathcal{P}_3(\mathbf{R})$ 使得 $p(0)=0, p'(0)=0$，且 $\int_0^1 |2+3x-p(x)|^2 \mathrm{d}x$ 尽可能小.

**18** 求 $p \in \mathcal{P}_5(\mathbf{R})$ 使 $\int_{-\pi}^{\pi} |\sin x - p(x)|^2 \mathrm{d}x$ 尽可能小.

注　式 (6.65) 这一多项式是对本题答案很不错的近似，但在这里你需要求出精确解，其中涉及 $\pi$ 的幂. 可演算符号积分的计算机应该会有帮助.

**19** 设 $V$ 是有限维的，且 $P \in \mathcal{L}(V)$ 是将 $V$ 映成其某个子空间的正交投影. 证明：$P^{\dagger} = P$.

**20** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V,W)$. 证明：

$$\operatorname{null} T^{\dagger} = (\operatorname{range} T)^{\perp} \quad \text{且} \quad \operatorname{range} T^{\dagger} = (\operatorname{null} T)^{\perp}.$$

**21** 设 $T \in \mathcal{L}(\mathbf{F}^3, \mathbf{F}^2)$ 定义为

$$T(a,b,c) = (a+b+c, 2b+3c).$$

(a) 对于 $(x,y) \in \mathbf{F}^2$，求 $T^{\dagger}(x,y)$ 的表达式.

(b) 由 (a) 中得到的 $T^{\dagger}$ 的表达式，验证：6.69 (b) 中的等式 $TT^{\dagger} = P_{\operatorname{range} T}$ 成立.

(c) 由 (a) 中得到的 $T^{\dagger}$ 的表达式，验证：6.69 (c) 中的等式 $T^{\dagger}T = P_{(\operatorname{null} T)^{\perp}}$ 成立.

**22** 设 $V$ 是有限维的，且 $T \in \mathcal{L}(V,W)$. 证明：

$$TT^{\dagger}T = T \quad \text{且} \quad T^{\dagger}TT^{\dagger} = T^{\dagger}.$$

注　以上公式在 $T$ 可逆时显然都成立，因为在那种情形下我们可以将 $T^{\dagger}$ 替换成 $T^{-1}$.

**23** 设 $V$ 和 $W$ 是有限维的，且 $T \in \mathcal{L}(V,W)$. 证明：

$$(T^{\dagger})^{\dagger} = T.$$

注　上式类似于 $(T^{-1})^{-1} = T$（$T$ 可逆时成立）.

# 第 7 章 内积空间上的算子

与内积空间有关的最深刻的结果，涉及我们现在所要讨论的主题——内积空间上的线性映射和算子. 我们将会看到，利用伴随的性质，可以证明一些很好的定理.

极其重要的谱定理将全面地描述实内积空间上的自伴算子和复内积空间上的正规算子. 然后我们将借助谱定理以理解正算子和幺正算子，这将引出幺正矩阵和矩阵分解. 谱定理也将引出相当常用的奇异值分解，奇异值分解又将引出极分解.

本书剩余部分最重要的结果都只在有限的维数下成立. 所以从现在开始我们假设 $V$ 和 $W$ 是有限维的.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 表示 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 和 $W$ 是 $\mathbf{F}$ 上的非零有限维内积空间.



图为利沃夫（Lviv）的集市广场. 由于国界线的变化，这座城市曾有过多个名字，也曾被划进过多个国家. 从 1772 年到 1918 年，这座城市位于奥地利，并被称为伦贝格（Lemberg）. 在第一次世界大战和第二次世界大战之间，这座城市位于波兰，并被称为利沃夫（Lwów）. 在此期间，利沃夫的数学家们，特别是斯特凡 · 巴拿赫（Stefan Banach，1892-1945）和他的同事们利用分析学的工具研究无限维向量空间，逐步建立了现代泛函分析的基本结论. 自第二次世界大战结束以来，利沃夫一直位于乌克兰，而后者在 1991 年宣告独立之前曾是苏联的一部分.

# 7A 自伴算子和正规算子

## 伴随

**7.1 定义：伴随（adjoint）、$T^*$**

设 $T \in \mathcal{L}(V, W)$. $T$ 的**伴随**是使得对任一 $v \in V$ 和任一 $w \in W$ 都有

$$\langle Tv, w\rangle = \langle v, T^* w\rangle$$

的函数 $T^* : W \to V$.

"伴随"这个词在线性代数里还有另一种含义，你要是在别的地方遇到了，要注意它和此处的"伴随"没有联系.

下面看看以上定义为什么是有意义的：设 $T \in \mathcal{L}(V, W)$，取定 $w \in W$，考虑 $V$ 上的线性泛函

$$v \mapsto \langle Tv, w\rangle,$$

它将 $v \in V$ 映射到 $\langle Tv, w\rangle$. 该线性泛函依赖于 $T$ 和 $w$. 根据里斯表示定理（6.42），$V$ 中存在唯一的向量，使得该线性泛函由与它的内积给出. 我们称这唯一的向量为 $T^* w$. 换句话说，$T^* w$ 是 $V$ 中使得对任一 $v \in V$ 都有

$$\langle Tv, w\rangle = \langle v, T^* w\rangle$$

的唯一向量.

上式中，左侧的内积是在 $W$ 上的，右侧的内积是在 $V$ 上的. 不过，我们对这两种内积用相同的记号 $\langle\cdot,\cdot\rangle$.

---

**7.2 例：从 $\mathbf{R}^3$ 到 $\mathbf{R}^2$ 的一线性映射之伴随**

定义 $T : \mathbf{R}^3 \to \mathbf{R}^2$ 为

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1).$$

为了计算 $T^*$，设 $(x_1, x_2, x_3) \in \mathbf{R}^3$ 和 $(y_1, y_2) \in \mathbf{R}^2$. 从而有

$$\begin{aligned}\big\langle T(x_1, x_2, x_3), (y_1, y_2)\big\rangle &= \big\langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\big\rangle \\ &= x_2 y_1 + 3x_3 y_1 + 2x_1 y_2 \\ &= \big\langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\big\rangle.\end{aligned}$$

根据上式以及伴随的定义可得

$$T^*(y_1, y_2) = (2y_2, y_1, 3y_1).$$

---

**7.3 例：值域维数至多为 1 的一线性映射之伴随**

取定 $u \in V$ 和 $x \in W$. 定义 $T \in \mathcal{L}(V, W)$ 为：对任一 $v \in V$ 都有

$$Tv = \langle v, u\rangle x.$$

转下页

为了计算 $T^*$，设 $v \in V$ 和 $w \in W$. 从而有

$$\begin{aligned}\langle Tv, w\rangle &= \big\langle \langle v, u\rangle x, w\big\rangle \\ &= \langle v, u\rangle\langle x, w\rangle \\ &= \big\langle v, \langle w, x\rangle u\big\rangle.\end{aligned}$$

因此，

$$T^*w = \langle w, x\rangle u.$$

---

在以上两例中，最后得到的 $T^*$ 不但是从 $W$ 到 $V$ 的函数，还是从 $W$ 到 $V$ 的线性映射. 下个结果表明，这在一般情况下都是成立的.

以上两例以及下面的证明用到了计算 $T^*$ 的常用方法：从 $\langle Tv, w\rangle$ 的表达式开始，然后处理一番，使得 $\langle \cdot, \cdot\rangle$ 的前一个位置里只有 $v$，那么后一个位置里就是 $T^*w$ 了.

**7.4 线性映射的伴随是线性映射**

如果 $T \in \mathcal{L}(V, W)$，那么 $T^* \in \mathcal{L}(W, V)$.

**证明** 设 $T \in \mathcal{L}(V, W)$. 如果 $v \in V$ 且 $w_1, w_2 \in W$，那么

$$\begin{aligned}\langle Tv, w_1 + w_2\rangle &= \langle Tv, w_1\rangle + \langle Tv, w_2\rangle \\ &= \langle v, T^*w_1\rangle + \langle v, T^*w_2\rangle \\ &= \langle v, T^*w_1 + T^*w_2\rangle.\end{aligned}$$

上式表明

$$T^*(w_1 + w_2) = T^*w_1 + T^*w_2.$$

如果 $v \in V$、$\lambda \in \mathbf{F}$ 且 $w \in W$，那么

$$\begin{aligned}\langle Tv, \lambda w\rangle &= \bar{\lambda}\langle Tv, w\rangle \\ &= \bar{\lambda}\langle v, T^*w\rangle \\ &= \langle v, \lambda T^*w\rangle.\end{aligned}$$

上式表明

$$T^*(\lambda w) = \lambda T^*w.$$

因此，$T^*$ 是线性映射，命题得证. ■

**7.5 伴随的性质**

设 $T \in \mathcal{L}(V, W)$，那么有：

(a) $(S + T)^* = S^* + T^*$ 对所有 $S \in \mathcal{L}(V, W)$ 成立；

(b) $(\lambda T)^* = \bar{\lambda} T^*$ 对所有 $\lambda \in \mathbf{F}$ 成立；

(c) $(T^*)^* = T$；

(d) $(ST)^* = T^*S^*$ 对所有 $S \in \mathcal{L}(W, U)$ 成立（这里 $U$ 是 $\mathbf{F}$ 上的有限维内积空间）；

(e) $I^* = I$，其中 $I$ 是 $V$ 上的恒等算子；

(f) 如果 $T$ 可逆，那么 $T^*$ 可逆且 $(T^*)^{-1} = (T^{-1})^*$.

**证明** 设 $v \in V$ 且 $w \in W$.

(a) 如果 $S \in \mathcal{L}(V, W)$，那么

$$\begin{aligned}\big\langle (S+T)v, w\big\rangle &= \langle Sv, w\rangle + \langle Tv, w\rangle \\ &= \langle v, S^*w\rangle + \langle v, T^*w\rangle \\ &= \langle v, S^*w + T^*w\rangle.\end{aligned}$$

因此 $(S+T)^*w = S^*w + T^*w$，命题得证.

(b) 如果 $\lambda \in \mathbf{F}$，那么

$$\big\langle (\lambda T)v, w\big\rangle = \lambda\langle Tv, w\rangle = \lambda\langle v, T^*w\rangle = \langle v, \bar{\lambda}T^*w\rangle.$$

因此 $(\lambda T)^*w = \bar{\lambda}T^*w$，命题得证.

(c) 因为

$$\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle,$$

所以 $(T^*)^*v = Tv$，命题得证.

(d) 设 $S \in \mathcal{L}(W, U)$ 和 $u \in U$，从而有

$$\big\langle (ST)v, u\big\rangle = \big\langle S(Tv), u\big\rangle = \big\langle Tv, S^*u\big\rangle = \big\langle v, T^*(S^*u)\big\rangle.$$

因此 $(ST)^*u = T^*(S^*u)$，命题得证.

(e) 设 $u \in V$. 从而有

$$\langle Iu, v\rangle = \langle u, v\rangle.$$

因此 $I^*v = v$，命题得证.

(f) 设 $T$ 可逆. 对等式 $T^{-1}T = I$ 两边取伴随，然后用 (d) 和 (e) 证得 $T^*(T^{-1})^* = I$. 类似地，由 $TT^{-1} = I$ 可得 $(T^{-1})^*T^* = I$. 因此，$(T^{-1})^*$ 是 $T^*$ 的逆，命题得证. ■

如果 $\mathbf{F} = \mathbf{R}$，那么根据以上结果的 (a) 和 (b)，$T \mapsto T^*$ 这一映射是从 $\mathcal{L}(V, W)$ 到 $\mathcal{L}(W, V)$ 的线性映射. 然而，如果 $\mathbf{F} = \mathbf{C}$，那么该映射不是线性的，这是由于 (b) 中出现的复共轭.

下面的结果揭示了线性映射同其伴随的零空间和值域之间的关系.

**7.6 $T^*$ 的零空间和值域**

设 $T \in \mathcal{L}(V, W)$，那么有：

(a) $\operatorname{null} T^* = (\operatorname{range} T)^\perp$；

(b) $\operatorname{range} T^* = (\operatorname{null} T)^\perp$；

(c) $\operatorname{null} T = (\operatorname{range} T^*)^\perp$；

(d) $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

**证明** 我们先证明 (a). 令 $w \in W$，从而有

$$\begin{aligned} w \in \operatorname{null} T^* &\iff T^*w = 0 \\ &\iff \langle v, T^*w\rangle = 0 \text{ 对所有 } v \in V \text{ 成立} \\ &\iff \langle Tv, w\rangle = 0 \text{ 对所有 } v \in V \text{ 成立} \\ &\iff w \in (\operatorname{range} T)^\perp.\end{aligned}$$

因此，$\operatorname{null} T^* = (\operatorname{range} T)^\perp$，(a) 得证.

如果我们对 (a) 的两侧求正交补，我们就得到了 (d)，其中用到了 6.52. 在 (a) 中将 $T$ 替换成 $T^*$，就得到了 (c)，其中用到了 7.5 (c). 最后，在 (d) 中将 $T$ 替换成 $T^*$，就得到了 (b). ■

我们很快会看到，接下来的定义与线性映射的伴随的矩阵密切相关.

**7.7 定义：共轭转置（conjugate transpose）、$A^*$**

$m \times n$ 矩阵 $A$ 的**共轭转置**是将其行列互换再对每个元素取复共轭得到的 $n \times m$ 矩阵 $A^*$. 换句话说，如果 $j \in \{1, \ldots, n\}$ 且 $k \in \{1, \ldots, m\}$，那么有

$$(A^*)_{j,k} = \overline{A_{k,j}}.$$

♣

**7.8 例：一个 $2 \times 3$ 矩阵的共轭转置**

$2 \times 3$ 矩阵 $\begin{pmatrix} 2 & 3+4\mathrm{i} & 7 \\ 6 & 5 & 8\mathrm{i} \end{pmatrix}$ 的共轭转置是 $3 \times 2$ 矩阵

$$\begin{pmatrix} 2 & 6 \\ 3-4\mathrm{i} & 5 \\ 7 & -8\mathrm{i} \end{pmatrix}.$$

若矩阵 $A$ 只含有实元素，则 $A^* = A^{\mathrm{t}}$，其中 $A^{\mathrm{t}}$ 表示 $A$ 的转置（行列互换得到的矩阵）.

接下来这条结果展示了如何从 $T$ 的矩阵计算出 $T^*$ 的矩阵. **注意：**关于非规范正交基，$T^*$ 的矩阵不一定等于 $T$ 的矩阵的共轭转置.

线性映射的伴随不依赖于基的选取. 因此我们经常强调的是线性映射的伴随，而不是矩阵的转置或共轭转置.

**7.9 $T^*$ 的矩阵等于 $T$ 的矩阵的共轭转置**

令 $T \in \mathcal{L}(V, W)$. 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，$f_1, \ldots, f_m$ 是 $W$ 的规范正交基. 那么 $\mathcal{M}(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n))$ 是 $\mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ 的共轭转置. 换句话说，

$$\mathcal{M}(T^*) = \big(\mathcal{M}(T)\big)^*.$$

♡

**证明** 在本证明中，我们把 $\mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ 和 $\mathcal{M}(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n))$ 这两个较长的表达式，分别写成 $\mathcal{M}(T)$ 和 $\mathcal{M}(T^*)$.

回顾一下，$\mathcal{M}(T)$ 的第 $k$ 列，是通过把 $Te_k$ 写成各 $f_j$ 的线性组合得到的——该线性组合中所用的标量就成为 $\mathcal{M}(T)$ 的第 $k$ 列. 因为 $f_1, \ldots, f_m$ 是 $W$ 的规范正交基，所以我们知道如何把 $Te_k$ 写成各 $f_j$ 的线性组合【见 6.30 (a)】:

$$Te_k = \langle Te_k, f_1\rangle f_1 + \cdots + \langle Te_k, f_m\rangle f_m.$$

因此

$$\mathcal{M}(T) \text{ 的 } j \text{ 行 } k \text{ 列元素是 } \langle Te_k, f_j\rangle.$$

在以上表述中，将 $T$ 替换成 $T^*$，并互换 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_m$，就证得 $\mathcal{M}(T^*)$ 的 $j$ 行 $k$ 列元素为 $\langle T^* f_k, e_j\rangle$. 这等于 $\langle f_k, Te_j\rangle$，进而等于 $\overline{\langle Te_j, f_k\rangle}$，而这就等于 $\mathcal{M}(T)$ 的 $k$ 行 $j$ 列元素的复共轭. 因此 $\mathcal{M}(T^*) = \big(\mathcal{M}(T)\big)^*$. ■

里斯表示定理（6.58 中所述）指出了 $V$ 和其对偶空间 $V'$（在 3.110 中定义）的等同关系. 在这一等同关系下，子集 $U \subseteq V$ 的正交补 $U^{\perp}$ 对应于 $U$ 的零化子 $U^0$. 如果 $U$ 是 $V$ 的子空间，那么在这一等同关系下，$U^{\perp}$ 和 $U^0$ 的维数公式就等同起来了——见 3.125 和 6.51.

比起零化子和对偶映射，正交补和伴随更容易处理. 所以在内积空间的背景下，就不需要处理零化子和对偶映射了.

设 $T: V \to W$ 是一线性映射. 在 $V$ 和 $V'$ 的等同关系与 $W$ 和 $W'$ 的等同关系之下，伴随映射 $T^*: W \to V$ 对应于对偶映射 $T': W' \to V'$（在 3.118 中定义），这也是你在本节习题 32 中所要验证的. 在这一等同关系下，$\operatorname{null} T^*$ 和 $\operatorname{range} T^*$ 的公式【7.6 (a) 和 (b)】就跟 $\operatorname{null} T'$ 和 $\operatorname{range} T'$ 的公式【3.128 (a) 和 3.130 (b)】完全一致. 此外，有关 $T^*$ 的矩阵的定理（7.9）也类似于有关 $T'$ 的矩阵的定理（3.132）.

## 自伴算子

现在，我们把注意力转向内积空间上的算子. 我们关注的不是从 $V$ 到 $W$ 的线性映射，而是从 $V$ 到 $V$ 的线性映射. 回顾一下，这样的线性映射就叫做算子.

**7.10 定义：自伴（self-adjoint）**

算子 $T \in \mathcal{L}(V)$ 称为**自伴的**，如果 $T = T^*$. ♣

如果 $T \in \mathcal{L}(V)$ 且 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，那么 $T$ 是自伴的当且仅当 $\mathcal{M}(T, (e_1, \ldots, e_n)) = \mathcal{M}(T, (e_1, \ldots, e_n))^*$. 这是由 7.9 推得的.

**7.11 例：根据 $T$ 的矩阵确定其是否自伴**

设 $c \in \mathbf{F}$ 且 $T$ 是 $\mathbf{F}^2$ 上的算子，其关于标准基的矩阵是

$$\mathcal{M}(T) = \begin{pmatrix} 2 & c \\ 3 & 7 \end{pmatrix}.$$

$T^*$ 关于标准基的矩阵是

$$\mathcal{M}(T^*) = \begin{pmatrix} 2 & 3 \\ \bar{c} & 7 \end{pmatrix}.$$

因此 $\mathcal{M}(T) = \mathcal{M}(T^*)$ 当且仅当 $c = 3$. 于是，算子 $T$ 自伴当且仅当 $c = 3$.

有个值得一记的好类比：$\mathcal{L}(V)$ 上的伴随扮演了类似于 $\mathbf{C}$ 上的复共轭的角色. 复数 $z$ 是实的当且仅当 $z = \bar{z}$；自伴算子（$T = T^*$）就类似于实数.

算子 $T \in \mathcal{L}(V)$ 自伴，当且仅当

$$\langle Tv, w \rangle = \langle v, Tw \rangle$$

对所有 $v, w \in V$ 成立.

我们将看到，以上所讨论的类比，会在自伴算子的一些重要性质上体现出来，首先是下一条结果中的特征值.

如果 $\mathbf{F} = \mathbf{R}$，那么根据定义每个特征值都是实的. 所以接下来这条结果只在 $\mathbf{F} = \mathbf{C}$ 的情形下引人兴趣.

**7.12 自伴算子的特征值**

自伴算子的每个特征值都是实的. ♡

**证明** 设 $T$ 是 $V$ 上的自伴算子. 令 $\lambda$ 是 $T$ 的特征值, 再令 $v$ 是 $V$ 中使得 $Tv = \lambda v$ 的非零向量. 那么有

$$\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\|v\|^2.$$

因此 $\lambda = \bar{\lambda}$, 意即 $\lambda$ 是实的, 命题得证. ■

接下来这条结果对于实内积空间而言是假命题. 举个例子, 考虑"绕原点逆时针旋转 90°"这一算子 $T \in \mathcal{L}(\mathbf{R}^2)$, 那么 $T(x, y) = (-y, x)$. 注意到, $Tv$ 正交于 $v$ 对任一 $v \in \mathbf{R}^2$ 都成立, 尽管 $T \neq 0$.

**7.13 对所有 $v$ 都有 $Tv$ 正交于 $v$ $\Longleftrightarrow$ $T = 0$ (假设 $\mathbf{F} = \mathbf{C}$)**

设 $V$ 是复内积空间以及 $T \in \mathcal{L}(V)$, 那么

$$\langle Tv, v\rangle = 0 \text{ 对任一 } v \in V \text{ 都成立} \iff T = 0.$$

**证明** 如果 $u, w \in V$, 那么有

$$\begin{aligned}\langle Tu, w\rangle = &\frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}\\ &+ \frac{\langle T(u+\mathrm{i}w), u+\mathrm{i}w\rangle - \langle T(u-\mathrm{i}w), u-\mathrm{i}w\rangle}{4}\mathrm{i}.\end{aligned}$$

这可以通过计算右侧来验证. 注意到, 右侧的每一项都具有 $\langle Tv, v\rangle$ ($v \in V$) 的形式.

现在设 $\langle Tv, v\rangle = 0$ 对任一 $v \in V$ 都成立, 那么上式意味着 $\langle Tu, w\rangle = 0$ 对所有 $u, w \in V$ 成立. 这意味着 $Tu = 0$ 对任一 $u \in V$ 都成立 (取 $w = Tu$). 于是 $T = 0$, 命题得证. ■

接下来这条结果对于实内积空间而言是假命题. 考虑实内积空间上任何一个非自伴的算子, 都能说明这点.

关于自伴算子如何表现得像实数, 下面这条结果给出了又一个好例子.

**7.14 $\langle Tv, v\rangle$ 对所有 $v$ 都是实的 $\Longleftrightarrow$ $T$ 是自伴的 (假设 $\mathbf{F} = \mathbf{C}$)**

设 $V$ 是复内积空间以及 $T \in \mathcal{L}(V)$, 那么

$$T \text{ 是自伴的} \iff \langle Tv, v\rangle \in \mathbf{R} \text{ 对任一 } v \in V \text{ 都成立.}$$

**证明** 如果 $v \in V$, 那么

$$\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}. \tag{7.15}$$

这样一来, 就有

$$\begin{aligned} T \text{ 是自伴的} &\iff T - T^* = 0\\ &\iff \langle (T - T^*)v, v\rangle = 0 \text{ 对任一 } v \in V \text{ 都成立}\\ &\iff \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ 对任一 } v \in V \text{ 都成立}\\ &\iff \langle Tv, v\rangle \in \mathbf{R} \text{ 对任一 } v \in V \text{ 都成立.}\end{aligned}$$

其中第二个等价关系通过对 $T - T^*$ 应用 7.13 得到, 第三个等价关系由式 (7.15) 得到. ■

在实内积空间 $V$ 上, 一个非零算子 $T$ 也许会对所有 $v \in V$ 都满足 $\langle Tv, v\rangle = 0$. 但是接下来这条结果表明, 这不会发生在自伴算子上.

**7.16 $T$ 自伴且 $\langle Tv, v\rangle = 0$ 对所有 $v$ 成立 $\Longleftrightarrow$ $T = 0$**

设 $T$ 是 $V$ 上的自伴算子，那么

$$\langle Tv, v\rangle = 0 \text{ 对任一 } v \in V \text{ 都成立} \iff T = 0.$$

**证明** 在 $V$ 是复内积空间的条件下，我们已经证明过这一定理，且无需假设 $T$ 自伴（见 7.13）. 因此我们可以假设 $V$ 是实内积空间. 如果 $u, w \in V$，那么有

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}. \tag{7.17}$$

这可以通过计算右侧来证明. 计算过程需用到下式：

$$\langle Tw, u\rangle = \langle w, Tu\rangle = \langle Tu, w\rangle,$$

其中第一个等号成立是因为 $T$ 自伴，第二个等号成立是因为我们在实内积空间里操作.

现在设 $\langle Tv, v\rangle = 0$ 对任一 $v \in V$ 都成立. 因为式 (7.17) 的右侧每一项都具有 $\langle Tv, v\rangle$（$v \in V$）的形式，所以这意味着 $\langle Tu, w\rangle = 0$ 对所有 $u, w \in V$ 成立. 这又意味着 $Tu = 0$ 对任一 $u \in V$ 都成立（取 $w = Tu$）. 于是，$T = 0$，命题得证. ■

## 正规算子

**7.18 定义：正规（normal）**

- 内积空间上的算子被称为**正规的**，如果它与它的伴随可交换.
- 换句话说，$T \in \mathcal{L}(V)$ 是正规的，如果 $TT^* = T^*T$.

每个自伴算子都是正规的，这是因为，如果 $T$ 自伴，那么 $T^* = T$，于是 $T$ 和 $T^*$ 可交换.

**7.19 例：正规但不自伴的算子**

令 $T$ 是 $\mathbf{F}^2$ 上的算子，其关于标准基的矩阵是

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

因此 $T(w, z) = (2w - 3z, 3w + 2z)$.

该算子 $T$ 并不是自伴的，因为其矩阵第 2 行第 1 列的元素（等于 3）不等于第 1 行第 2 列的元素（等于 $-3$）的复共轭.

$TT^*$ 的矩阵等于

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}, \quad \text{等于} \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

类似地，$T^*T$ 的矩阵等于

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \quad \text{等于} \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

因为 $TT^*$ 和 $T^*T$ 的矩阵相同，我们得知 $TT^* = T^*T$. 因此 $T$ 是正规的.

在下一节，我们将看到正规算子为什么值得特别关注. 接下来这条结果给出了对正规算子的一个很有用的刻画.

**7.20 $T$ 是正规的当且仅当 $Tv$ 和 $T^*v$ 的范数相同**

设 $T \in \mathcal{L}(V)$，那么有

$$T \text{ 是正规的} \iff \|Tv\| = \|T^*v\| \text{ 对任一 } v \in V \text{ 都成立.}$$

**证明**

$$\begin{aligned} T \text{ 是正规的} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)v, v \rangle = 0 \text{ 对任一 } v \in V \text{ 都成立} \\ &\iff \langle T^*Tv, v \rangle = \langle TT^*v, v \rangle \text{ 对任一 } v \in V \text{ 都成立} \\ &\iff \langle Tv, Tv \rangle = \langle T^*v, T^*v \rangle \text{ 对任一 } v \in V \text{ 都成立} \\ &\iff \|Tv\|^2 = \|T^*v\|^2 \text{ 对任一 } v \in V \text{ 都成立} \\ &\iff \|Tv\| = \|T^*v\| \text{ 对任一 } v \in V \text{ 都成立.} \end{aligned}$$

其中，我们用到了 7.16 来建立第二个等价关系（注意到算子 $T^*T - TT^*$ 是自伴的）. ■

下一条结果展示了以上结果的几个推论. 这里将下一条结果的 (e) 和本节习题 3 相比较. 那道习题指出每个算子的伴随的特征值（作为集合）都等于该算子特征值的复共轭 [1]，而没有提及特征向量，因为算子同它的伴随可以有不同的特征向量. 然而，下一条结果中的 (e) 表明，正规算子和它的伴随特征向量完全相同.

**7.21 正规算子的值域、零空间和特征向量**

设 $T \in \mathcal{L}(V)$ 是正规的，那么有：

(a) $\operatorname{null} T = \operatorname{null} T^*$；

(b) $\operatorname{range} T = \operatorname{range} T^*$；

(c) $V = \operatorname{null} T \oplus \operatorname{range} T$；

(d) 对任一 $\lambda \in \mathbf{F}$，$T - \lambda I$ 都是正规的；

(e) 如果 $v \in V$ 且 $\lambda \in \mathbf{F}$，那么 $Tv = \lambda v$ 当且仅当 $T^*v = \bar{\lambda} v$.

**证明**

(a) 设 $v \in V$，那么有

$$v \in \operatorname{null} T \iff \|Tv\| = 0 \iff \|T^*v\| = 0 \iff v \in \operatorname{null} T^*.$$

以上等价关系的中间那个由 7.20 得到. 因此，$\operatorname{null} T = \operatorname{null} T^*$.

(b)

$$\operatorname{range} T = (\operatorname{null} T^*)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*.$$

其中，第一个相等关系来自 7.6 (d)，第二个相等关系来自本结果的 (a)，第三个相等关系来自 7.6 (b).

[1] 意为：该算子的全体特征值可构成一个集合，对这集合中所有元素取复共轭，就会得到该算子的伴随的全体特征值构成的集合.

(c)

$$V = (\operatorname{null} T) \oplus (\operatorname{null} T)^\perp = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T.$$

其中，第一个相等关系来自 6.49，第二个相等关系来自 7.6 (b)，第三个相等关系来自本结果的 (b).

(d) 设 $\lambda \in \mathbf{F}$，那么有

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \bar{\lambda} I)\\ &= TT^* - \bar{\lambda} T - \lambda T^* + |\lambda|^2 I\\ &= T^*T - \bar{\lambda} T - \lambda T^* + |\lambda|^2 I\\ &= (T^* - \bar{\lambda} I)(T - \lambda I)\\ &= (T - \lambda I)^*(T - \lambda I).\end{aligned}$$

因此，$T - \lambda I$ 和它的伴随可交换．于是 $T - \lambda I$ 是正规的.

(e) 设 $v \in V$ 且 $\lambda \in \mathbf{F}$．那么 (d) 和 7.20 蕴涵

$$\|(T - \lambda I)v\| = \|(T - \lambda I)^* v\| = \|(T^* - \bar{\lambda} I)v\|.$$

因此，$\|(T - \lambda I)v\| = 0$ 当且仅当 $\|(T^* - \bar{\lambda} I)v\| = 0$．于是 $Tv = \lambda v$ 当且仅当 $T^* v = \bar{\lambda} v$. ■

因为每个自伴算子都是正规的，所以下一条结果也适用于自伴算子.

**7.22 正规算子的正交特征向量**

设 $T \in \mathcal{L}(V)$ 是正规的．那么 $T$ 的对应于不同特征值的特征向量正交.

**证明** 设 $\alpha, \beta$ 是 $T$ 的不同特征值，对应的特征向量是 $u, v$．因此 $Tu = \alpha u$ 且 $Tv = \beta v$．根据 7.21 (e)，有 $T^* v = \bar{\beta} v$．因此

$$\begin{aligned}(\alpha - \beta)\langle u, v\rangle &= \langle \alpha u, v\rangle - \langle u, \bar{\beta} v\rangle\\ &= \langle Tu, v\rangle - \langle u, T^* v\rangle\\ &= 0.\end{aligned}$$

因为 $\alpha \neq \beta$，所以上式表明 $\langle u, v\rangle = 0$．因此 $u$ 和 $v$ 正交，命题得证. ■

如下所述，下一条结果仅在 $\mathbf{F} = \mathbf{C}$ 条件下有意义．然而在本节习题 12 中，可以看到一个不论 $\mathbf{F} = \mathbf{C}$ 还是 $\mathbf{F} = \mathbf{R}$ 都有意义的版本.

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$．将 $\mathcal{L}(V)$ 类比为 $\mathbf{C}$，$\mathcal{L}(V)$ 上的伴随就如同 $\mathbf{C}$ 上的复共轭，那么由式 (7.24) 定义的算子 $A$ 和 $B$ 也就相当于 $T$ 的“实部”和“虚部”．这样一来，以下结果的非正式标题才有了意义.

**7.23 $T$ 是正规的 $\Longleftrightarrow$ $T$ 的实部和虚部可交换**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$．那么，$T$ 是正规的当且仅当存在可交换的自伴算子 $A$ 和 $B$ 使得 $T = A + \mathrm{i}B$.

**证明** 首先，设 $T$ 是正规的．令

$$A = \frac{T + T^*}{2} \quad 且 \quad B = \frac{T - T^*}{2\mathrm{i}}. \tag{7.24}$$

那么 $A$ 和 $B$ 是自伴的且 $T = A + \mathrm{i}B$. 可以很快计算得

$$AB - BA = \frac{T^*T - TT^*}{2\mathrm{i}}. \tag{7.25}$$

因为 $T$ 是正规的，所以上式右侧等于 0. 因此，算子 $A$ 和 $B$ 可交换，此方向得证.

为了证明另一方向的蕴涵关系，现在假设存在可交换的自伴算子 $A$ 和 $B$ 使得 $T = A + \mathrm{i}B$，那么 $T^* = A - \mathrm{i}B$. 将前面这两个式子相加再除以 2，得到式 (7.24) 中 $A$ 的式子；将前面这两个式子相减再除以 2i，得到式 (7.24) 中 $B$ 的式子. 再由式 (7.24) 可得式 (7.25). 因为 $B$ 和 $A$ 可交换，所以式 (7.25) 蕴涵 $T$ 是正规的，此方向得证. ■

## 习题 7A

**1** 设 $n$ 是一正整数. 定义 $T \in \mathcal{L}(\mathbf{F}^n)$ 为

$$T(z_1, \ldots, z_n) = (0, z_1, \ldots, z_{n-1}).$$

求 $T^*(z_1, \ldots, z_n)$ 的表达式.

**2** 设 $T \in \mathcal{L}(V, W)$. 证明：

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

**3** 设 $T \in \mathcal{L}(V)$ 和 $\lambda \in \mathbf{F}$. 证明：

$$\lambda \text{ 是 } T \text{ 的特征值} \iff \overline{\lambda} \text{ 是 } T^* \text{ 的特征值}.$$

**4** 设 $T \in \mathcal{L}(V)$ 且 $U$ 是 $V$ 的子空间. 证明：

$$U \text{ 在 } T \text{ 下不变} \iff U^\perp \text{ 在 } T^* \text{ 下不变}.$$

**5** 设 $T \in \mathcal{L}(V, W)$. 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，$f_1, \ldots, f_m$ 是 $W$ 的规范正交基. 证明：

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*f_1\|^2 + \cdots + \|T^*f_m\|^2.$$

注 上式中的 $\|Te_1\|^2, \ldots, \|Te_n\|^2$ 各项依赖于规范正交基 $e_1, \ldots, e_n$，但右侧不依赖于 $e_1, \ldots, e_n$. 因此上式表明，左侧的和不取决于 $e_1, \ldots, e_n$ 选择的是哪个规范正交基.

**6** 设 $T \in \mathcal{L}(V, W)$. 证明：

(a) $T$ 是单射 $\iff$ $T^*$ 是满射；

(b) $T$ 是满射 $\iff$ $T^*$ 是单射.

**7** 证明：如果 $T \in \mathcal{L}(V, W)$，那么

(a) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$；

(b) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

**8** 设 $A$ 为 $m \times n$ 矩阵，其元素在 $\mathbf{F}$ 中. 用第 7 题的 (b) 证明：$A$ 的行秩等于 $A$ 的列秩.

注 题述结果在 3.57 和 3.133 中已经证明过了，本题要求换种方式证明.

**9** 证明：$V$ 上的两自伴算子的乘积是自伴的，当且仅当这两个算子可交换.

**10** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 是自伴的，当且仅当

$$\langle Tv, v \rangle = \langle T^*v, v \rangle$$

对所有 $v \in V$ 成立.

**11** 定义算子 $S:\mathbf{F}^2\to\mathbf{F}^2$ 为 $S(w,z)=(-z,w)$.

(a) 求 $S^*$ 的表达式.

(b) 证明：$S$ 正规但不自伴.

(c) 求 $S$ 的所有特征值.

注 如果 $\mathbf{F}=\mathbf{R}$，那么 $S$ 是 $\mathbf{R}^2$ 上“逆时针旋转 $90°$”这一算子.

**12** 称算子 $B\in\mathcal{L}(V)$ 是**斜的**（**skew**）[2]，如果

$$B^*=-B.$$

设 $T\in\mathcal{L}(V)$，证明：$T$ 是正规的，当且仅当存在可交换算子 $A$ 和 $B$ 使得 $A$ 是自伴的、$B$ 是斜算子且 $T=A+B$.

**13** 设 $\mathbf{F}=\mathbf{R}$. 定义 $\mathcal{A}\in\mathcal{L}\big(\mathcal{L}(V)\big)$ 为：对所有 $T\in\mathcal{L}(V)$ 都有 $\mathcal{A}T=T^*$.

(a) 求 $\mathcal{A}$ 的所有特征值.

(b) 求 $\mathcal{A}$ 的最小多项式.

**14** 在 $\mathcal{P}_2(\mathbf{R})$ 上定义一内积为 $\langle p,q\rangle=\int_0^1 pq$. 定义算子 $T\in\mathcal{L}\big(\mathcal{P}_2(\mathbf{R})\big)$ 为

$$T(ax^2+bx+c)=bx.$$

(a) 证明：该内积下，$T$ 不是自伴的.

(b) $T$ 关于基 $1,x,x^2$ 的矩阵是

$$\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix}.$$

这个矩阵等于它的共轭转置，尽管 $T$ 不是自伴的. 解释为什么这不矛盾.

**15** 设 $T\in\mathcal{L}(V)$ 可逆. 证明：

(a) $T$ 自伴 $\Longleftrightarrow$ $T^{-1}$ 自伴；

(b) $T$ 正规 $\Longleftrightarrow$ $T^{-1}$ 正规.

**16** 设 $\mathbf{F}=\mathbf{R}$.

(a) 证明：$V$ 上的自伴算子构成的集合是 $\mathcal{L}(V)$ 的子空间.

(b) (a) 中这个子空间的维数是多少？（用 $\dim V$ 表示）

**17** 设 $\mathbf{F}=\mathbf{C}$. 证明：$V$ 上的全体自伴算子构成的集合不是 $\mathcal{L}(V)$ 的子空间.

**18** 设 $\dim V\geq 2$. 证明：$V$ 上的全体正规算子构成的集合不是 $\mathcal{L}(V)$ 的子空间.

**19** 设 $T\in\mathcal{L}(V)$ 且对任一 $v\in V$ 都有 $\|T^*v\|\leq\|Tv\|$. 证明：$T$ 是正规的.

注 本题在无限维内积空间上不成立，从而引出了所谓**亚正规算子**（**hyponormal operators**）. 关于此类算子，已有较完善的理论.

**20** 设 $P\in\mathcal{L}(V)$ 使得 $P^2=P$. 证明下列等价：

(a) $P$ 是自伴的.

(b) $P$ 是正规的.

(c) 存在 $V$ 的子空间 $U$ 使得 $P=P_U$.

---

[2] 斜？确实很像斜对称矩阵（满足 $A^t+A=0$ 的矩阵 $A$）. 而斜对称（skew-symmetric）矩阵又叫反对称（antisymmetric）矩阵，类似地，这种算子也叫做反自伴算子.

**21** 设 $D:\mathcal{P}_8(\mathbf{R})\to\mathcal{P}_8(\mathbf{R})$ 是定义为 $Dp=p'$ 的微分算子．证明：$\mathcal{P}_8(\mathbf{R})$ 上不存在能使 $D$ 成为正规算子的内积．

**22** 给出一例：算子 $T\in\mathcal{L}(\mathbf{R}^3)$ 正规但不自伴．

**23** 设 $T$ 是 $V$ 上的正规算子．设 $v,w\in V$ 满足下列各式：

$$\|v\|=\|w\|=2,\qquad Tv=3v,\qquad Tw=4w.$$

证明：$\|T(v+w)\|=10$.

**24** 设 $T\in\mathcal{L}(V)$ 且

$$a_0+a_1z+a_2z^2+\cdots+a_{m-1}z^{m-1}+z^m$$

是 $T$ 的最小多项式．证明：$T^*$ 的最小多项式是

$$\overline{a_0}+\overline{a_1}z+\overline{a_2}z^2+\cdots+\overline{a_{m-1}}z^{m-1}+z^m.$$

注　本题表明，如果 $\mathbf{F}=\mathbf{R}$，那么 $T^*$ 的最小多项式等于 $T$ 的最小多项式．

**25** 设 $T\in\mathcal{L}(V)$．证明：$T$ 可对角化，当且仅当 $T^*$ 可对角化．

**26** 给定 $u,x\in V$．定义 $T\in\mathcal{L}(V)$ 为：对任一 $v\in V$ 都有 $Tv=\langle v,u\rangle x$.

(a) 证明：如果 $V$ 是实向量空间，那么 $T$ 自伴当且仅当组 $u,x$ 线性相关．

(b) 证明：$T$ 正规，当且仅当组 $u,x$ 线性相关．

**27** 设 $T\in\mathcal{L}(V)$ 是正规的．证明：

$$\operatorname{null}T^k=\operatorname{null}T\quad\text{和}\quad\operatorname{range}T^k=\operatorname{range}T$$

对任一正整数 $k$ 都成立．

**28** 设 $T\in\mathcal{L}(V)$ 是正规的，证明：如果 $\lambda\in\mathbf{F}$，那么 $T$ 的最小多项式不是 $(x-\lambda)^2$ 的多项式倍．

**29** 证明或给出一反例：如果 $T\in\mathcal{L}(V)$ 且 $V$ 有一规范正交基 $e_1,\dots,e_n$ 使得 $\|Te_k\|=\|T^*e_k\|$ 对任一 $k=1,\dots,n$ 都成立，那么 $T$ 是正规的．

**30** 设 $T\in\mathcal{L}(\mathbf{F}^3)$ 是正规的且 $T(1,1,1)=(2,2,2)$．设 $(z_1,z_2,z_3)\in\operatorname{null}T$．证明：$z_1+z_2+z_3=0$.

**31** 给定一正整数 $n$．在由 $[-\pi,\pi]$ 上全体连续实值函数构成的、具有内积 $\langle f,g\rangle=\int_{-\pi}^{\pi}fg$ 的内积空间中，令

$$V=\operatorname{span}(1,\cos x,\cos 2x,\dots,\cos nx,\sin x,\sin 2x,\dots,\sin nx).$$

(a) 定义 $D\in\mathcal{L}(V)$ 为 $Df=f'$．证明 $D^*=-D$，并得出结论：$D$ 正规但不自伴．

(b) 定义 $T\in\mathcal{L}(V)$ 为 $Tf=f''$．证明：$T$ 是自伴的．

**32** 设 $T:V\to W$ 是一线性映射．证明：在 $V$ 和 $V'$ 的标准等同关系（见 6.58）以及相应的 $W$ 和 $W'$ 的等同关系下，伴随映射 $T^*:W\to V$ 对应于对偶映射 $T':W'\to V'$．更准确地说，证明：

$$T'(\varphi_w)=\varphi_{T^*w}$$

对所有 $w\in W$ 成立，其中 $\varphi_w$ 和 $\varphi_{T^*w}$ 如 6.58 所定义．

# 7B 谱定理

回顾一下：对角矩阵是对角线外处处为 0 的方阵；$V$ 上的算子若关于 $V$ 的某个基有对角矩阵就被称为可对角化的，而这种情况当且仅当 $V$ 有由该算子的特征向量组成的基时才会发生（见 5.55）.

$V$ 上最好的算子，就是关于 $V$ 中的某个**规范正交**基有对角矩阵的算子. 也正是这些算子，其特征向量可用来组成 $V$ 的规范正交基. 本节我们的目标是，证明用以刻画这些算子的谱定理——它指出，当 $\mathbf{F}=\mathbf{R}$ 时这些算子即自伴算子，当 $\mathbf{F}=\mathbf{C}$ 时这些算子即正规算子.

在对内积空间上的算子的研究中，谱定理也许是最有用的工具. 它在特定的无限维内积空间上的推广——例如，见作者的《测度、积分与实分析》（*Measure, Integration & Real Analysis*）一书的章节 10D——在泛函分析中起到了关键作用.

由于谱定理的结论依赖于 $\mathbf{F}$，所以我们将谱定理划成两块，分别称作实谱定理和复谱定理.

## 实谱定理

为证明实谱定理，我们需要两个引理. 它们在实内积空间和复内积空间上都成立，但是在复谱定理的证明中不需要用到.

对于接下来这条结果，你只要想到实系数二次多项式，就能猜到它是正确的，甚至还能给出它的证明. 具体而言，设 $b,c\in\mathbf{R}$ 且 $b^2<4c$，令 $x$ 是一实数，从而有

$$x^2+bx+c=\left(x+\frac{b}{2}\right)^2+\left(c-\frac{b^2}{4}\right)>0.$$

特别地，$x^2+bx+c$ 是一可逆实数（这不过是把“非 0”往复杂了说）. 将实数 $x$ 替换为自伴算子（回想一下实数和自伴算子之间的类比），就引出了下面这条结果.

这种配方法可以用于推导二次求根公式.

**7.26 可逆二次表达式**

设 $T\in\mathcal{L}(V)$ 是自伴的，且 $b,c\in\mathbf{R}$ 使得 $b^2<4c$. 那么

$$T^2+bT+cI$$

是可逆算子.

**证明** 令 $v$ 是 $V$ 中的非零向量，那么

$$\begin{aligned}\langle(T^2+bT+cI)v,v\rangle&=\langle T^2v,v\rangle+b\langle Tv,v\rangle+c\langle v,v\rangle\\&=\langle Tv,Tv\rangle+b\langle Tv,v\rangle+c\|v\|^2\\&\geq\|Tv\|^2-|b|\,\|Tv\|\,\|v\|+c\|v\|^2\\&=\left(\|Tv\|-\frac{|b|\,\|v\|}{2}\right)^2+\left(c-\frac{b^2}{4}\right)\|v\|^2\\&>0,\end{aligned}$$

其中第三行成立是由柯西-施瓦兹不等式(6.14)得到的.最后一个不等号意味着 $(T^2+bT+cI)v \neq 0$. 因此 $T^2+bT+cI$ 是单射，这意味着它是可逆的(见 3.65). ■

接下来这条结果将是我们证明实谱定理所用的关键工具.

**7.27 自伴算子的最小多项式**

设 $T \in \mathcal{L}(V)$ 是自伴的. 那么 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$, 其中 $\lambda_1,\ldots,\lambda_m \in \mathbf{R}$.

**证明** 首先，设 $\mathbf{F}=\mathbf{C}$. $T$ 的最小多项式的零点即 $T$ 的特征值【根据 5.27 (a)】. 而 $T$ 的所有特征值都是实的(根据 7.12). 因此代数基本定理的版本二(见 4.13)告诉我们 $T$ 的最小多项式具有待证的形式.

现在假设 $\mathbf{F}=\mathbf{R}$. 由 $\mathbf{R}$ 上的多项式分解(见 4.16)可得，存在 $\lambda_1,\ldots,\lambda_m \in \mathbf{R}$ 和 $b_1,\ldots,b_N$, $c_1,\ldots,c_N \in \mathbf{R}$(对任一 $k$ 都有 $b_k^2 < 4c_k$)使得 $T$ 的最小多项式等于

$$(z-\lambda_1)\cdots(z-\lambda_m)(z^2+b_1z+c_1)\cdots(z^2+b_Nz+c_N). \tag{7.28}$$

其中 $m$ 或 $N$ 中可能会有一个等于 0，也就是没有相应形式的项. 这样一来，就有

$$(T-\lambda_1 I)\cdots(T-\lambda_m I)\bigl(T^2+b_1T+c_1I\bigr)\cdots\bigl(T^2+b_NT+c_NI\bigr)=0.$$

如果 $N>0$，那么我们可以对上式两边同时乘以 $T^2+b_NT+c_NI$(由 7.26 可知其为可逆算子)的逆，这样得到的 $T$ 的多项式仍等于 0，但其次数比式 (7.28) 低 2，这和最小多项式"次数最小"的性质相违背. 所以必然有 $N=0$，这意味着式 (7.28) 中的最小多项式具有的形式是 $(z-\lambda_1)\cdots(z-\lambda_m)$，命题得证. ■

以上结果连同 5.27 (a) 蕴涵了"每个自伴算子都有特征值"这一命题. 事实上，我们在接下来的结果中还会看到，自伴算子有足够的特征向量以形成一个基.

下面这条结果，是线性代数中的主要定理之一，全面地描述了实内积空间上的自伴算子.

**7.29 实谱定理(real spectral theorem)**

设 $\mathbf{F}=\mathbf{R}$ 且 $T \in \mathcal{L}(V)$. 那么下列等价:

(a) $T$ 是自伴的.

(b) $T$ 关于 $V$ 的某个规范正交基有对角矩阵.

(c) $V$ 有由 $T$ 的特征向量构成的规范正交基.

**证明** 首先假设 (a) 成立，那么 $T$ 是自伴的. 根据 6.37 和 7.27 这两条有关最小多项式的结果，可得 $T$ 关于 $V$ 的某个规范正交基有上三角矩阵. 关于这一规范正交基，$T^*$ 的矩阵是 $T$ 的矩阵的转置. 然而，$T^*=T$. 因此，$T$ 的矩阵的转置等于 $T$ 的矩阵. 又因为 $T$ 的矩阵是上三角的，所以这意味着该矩阵对角线上方和下方的元素都为 0. 于是，$T$ 关于这一规范正交基的矩阵是对角矩阵. 因此，(a) 蕴涵 (b).

相反，现在假设 (b) 成立，那么 $T$ 关于 $V$ 的某个规范正交基有对角矩阵. 这个对角矩阵和它的转置相等. 所以关于那个基，$T^*$ 的矩阵等于 $T$ 的矩阵. 于是，$T^*=T$，证得 (b) 蕴涵 (a).

(b) 和 (c) 之间的等价关系，由定义可得【或者参见 5.55 (a) 和 (b) 等价的证明】. ■

**7.30 例：一算子的特征向量所成的规范正交基**

考虑 $\mathbf{R}^3$ 上的算子 $T$，其矩阵（关于标准基）为

$$\begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}.$$

该矩阵等于自身的转置且其元素都为实数，所以 $T$ 是自伴的．可以验证，

$$\frac{(1,-1,0)}{\sqrt{2}}, \frac{(1,1,1)}{\sqrt{3}}, \frac{(1,1,-2)}{\sqrt{6}}$$

是 $\mathbf{R}^3$ 的规范正交基，且由 $T$ 的特征向量组成．关于这个基，$T$ 的矩阵是对角矩阵

$$\begin{pmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{pmatrix}.$$

实谱定理同时应用于不止一个算子的版本，见本节习题 17.

## 复谱定理

接下来这条结果，全面地描述了复内积空间上的正规算子.

**7.31 复谱定理（complex spectral theorem）**

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$．那么下列等价：

(a) $T$ 是正规的.

(b) $T$ 关于 $V$ 的某个规范正交基有对角矩阵.

(c) $V$ 有由 $T$ 的特征向量构成的规范正交基.

**证明** 首先假设 (a) 成立，那么 $T$ 是正规的．由舒尔定理（6.38），存在 $V$ 的一规范正交基 $e_1,\dots,e_n$，使得 $T$ 关于它有上三角矩阵．于是我们可以写出

$$\mathcal{M}(T,(e_1,\dots,e_n)) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}. \tag{7.32}$$

我们将证明这个矩阵实际上是对角矩阵.

我们从以上这个矩阵可以得到

$$\begin{aligned} \|Te_1\|^2 &= |a_{1,1}|^2, \\ \|T^*e_1\|^2 &= |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2. \end{aligned}$$

因为 $T$ 是正规的，所以 $\|Te_1\| = \|T^*e_1\|$（见 7.20）．因此由以上两个等式可得，式 (7.32) 中的矩阵第一行除了第一个元素 $a_{1,1}$ 可能不为 0 外都为 0.

这样一来，式 (7.32) 蕴涵了

$$\|Te_2\|^2 = |a_{2,2}|^2$$

（因为我们在上一段证明了 $a_{1,2}=0$）和

$$\|T^*e_2\|^2=\left|a_{2,2}\right|^2+\left|a_{2,3}\right|^2+\cdots+\left|a_{2,n}\right|^2.$$

因为 $T$ 是正规的，所以 $\|Te_2\|=\|T^*e_2\|$. 因此由以上两个等式可得，式 (7.32) 中的矩阵第二行除了对角线元素 $a_{2,2}$ 可能不为 0 外都为 0.

按这种方式继续下去，我们可以得到，矩阵 (7.32) 的所有非对角线元素都等于 0. 因此 (b) 成立，也就完成了 (a) 蕴涵 (b) 的证明.

现在假设 (b) 成立，那么 $T$ 关于 $V$ 的某个规范正交基有一对角矩阵. 取 $T$ 的矩阵的共轭转置，可以得到 $T^*$（关于同一个基）的矩阵；于是 $T^*$ 也有对角矩阵. 任意两个对角矩阵都可交换，从而 $T$ 和 $T^*$ 可交换，意即 $T$ 是正规的. 换句话说，(a) 成立. 这也就完成了 (b) 蕴涵 (a) 的证明.

(b) 和 (c) 的等价关系，由定义可得（也见 5.55）. ■

关于前述结果中的 (a) 蕴涵 (b)，可在本节习题 13 和 20 找到其他方式的证明.

本节习题 14 和 15 通过将定义空间表达成特征空间的正交直和，来解释实谱定理和复谱定理.

复谱定理同时应用于不止一个算子的版本，见习题 16.

复谱定理的主要结论是：有限维复内积空间上的每个正规算子，都可由规范正交基对角化. 接下来的例子呈现了这一点.

---

**7.33 例：一算子的特征向量所成的规范正交基**

考虑算子 $T\in\mathcal{L}(\mathbf{C}^2)$，其定义为：$T(w,z)=(2w-3z,3w+2z)$. $T$（关于标准基）的矩阵为

$$\begin{pmatrix}2&-3\\3&2\end{pmatrix}.$$

正如我们在例 7.19 中所见，$T$ 是正规算子.

可以验证，

$$\frac{1}{\sqrt{2}}(\mathrm{i},1),\ \frac{1}{\sqrt{2}}(-\mathrm{i},1)$$

是 $\mathbf{C}^2$ 的规范正交基，且由 $T$ 的特征向量组成. 关于这个基，$T$ 的矩阵是对角矩阵

$$\begin{pmatrix}2+3\mathrm{i}&0\\0&2-3\mathrm{i}\end{pmatrix}.$$

---

## 习题 7B

**1** 证明：复内积空间上的正规算子是自伴的，当且仅当它的所有特征值是实的.

注 在假设算子正规的前提下，本题加强了自伴算子和实数之间的类比.

**2** 设 $\mathbf{F}=\mathbf{C}$. 设 $T\in\mathcal{L}(V)$ 是正规的且只有一个特征值. 证明：$T$ 是恒等算子的标量倍.

**3** 设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$ 是正规的. 证明：$T$ 的特征值所构成的集合包含于 $\{0,1\}$，当且仅当存在 $V$ 的子空间 $U$ 使得 $T=P_U$.

**4** 证明：复内积空间上的正规算子是斜的（意即等于其伴随的负[3]），当且仅当它的所有特征值都是纯虚数（意即它们的实部都等于 0）.

**5** 证明或给出一反例：如果 $T \in \mathcal{L}(\mathbf{C}^3)$ 是可对角化的算子，那么 $T$（关于通常的内积）是正规的.

**6** 设 $V$ 是复内积空间，$T \in \mathcal{L}(V)$ 是使得 $T^9 = T^8$ 的正规算子. 证明：$T$ 是自伴的且 $T^2 = T$.

**7** 给出一例：复内积空间上的算子 $T$ 使得 $T^9 = T^8$ 而 $T^2 \neq T$.

**8** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 是正规的，当且仅当 $T$ 的每个特征向量都是 $T^*$ 的特征向量.

**9** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 是正规的，当且仅当存在一多项式 $p \in \mathcal{P}(\mathbf{C})$ 使得 $T^* = p(T)$.

**10** 设 $V$ 是复内积空间. 证明：$V$ 上每个正规算子都有平方根.

注　算子 $S \in \mathcal{L}(V)$ 称为 $T \in \mathcal{L}(V)$ 的**平方根（square root）**，如果 $S^2 = T$. 关于算子的平方根，我们将在章节 7C 和 8C 讨论更多的内容.

**11** 证明：$V$ 上的每个自伴算子都有立方根.

注　算子 $S \in \mathcal{L}(V)$ 称为 $T \in \mathcal{L}(V)$ 的**立方根（cube root）**，如果 $S^3 = T$.

**12** 设 $V$ 是复向量空间，且 $T \in \mathcal{L}(V)$ 是正规的. 证明：如果 $S$ 是 $V$ 上与 $T$ 可交换的算子，那么 $S$ 和 $T^*$ 可交换.

注　本题的结果称为富格里德定理（Fuglede's theorem）.

**13** 不用复谱定理，而是用舒尔定理应用于两个可交换算子的版本（6B 节习题 20 中，取 $\mathcal{E} = \{T, T^*\}$），给出该命题的另一证明：如果 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ 是正规的，那么 $T$ 关于 $V$ 的某个规范正交基有对角矩阵.

**14** 设 $\mathbf{F} = \mathbf{R}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 是自伴的，当且仅当 $T$ 的对应于不同特征值的特征向量两两正交，且 $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$，其中 $\lambda_1, \ldots, \lambda_m$ 表示 $T$ 的所有互异特征值.

**15** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 是正规的，当且仅当 $T$ 的对应于不同特征值的特征向量两两正交，且 $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$，其中 $\lambda_1, \ldots, \lambda_m$ 表示 $T$ 的所有互异特征值.

**16** 设 $\mathbf{F} = \mathbf{C}$ 且 $\mathcal{E} \subseteq \mathcal{L}(V)$. 证明：$V$ 有一个规范正交基使得 $\mathcal{E}$ 中每个元素关于它都有对角矩阵，当且仅当对于所有 $S, T \in \mathcal{E}$ 都有 $S$ 和 $T$ 是可交换的正规算子.

注　本题将复谱定理扩展到了可交换正规算子的集合中.

**17** 设 $\mathbf{F} = \mathbf{R}$ 且 $\mathcal{E} \subseteq \mathcal{L}(V)$. 证明：$V$ 有一个规范正交基使得 $\mathcal{E}$ 中每个元素关于它都有对角矩阵，当且仅当对于所有 $S, T \in \mathcal{E}$ 都有 $S$ 和 $T$ 是可交换的自伴算子.

注　本题将实谱定理扩展到了可交换自伴算子的集合中.

**18** 给出一例：实内积空间 $V$，算子 $T \in \mathcal{L}(V)$，以及实数 $b, c$（$b^2 < 4c$）使得

$$T^2 + bT + cI$$

不是可逆的.

注　本题表明，即使是在实向量空间中，7.26 中"$T$ 是自伴的"这一条件也不能删去.

**19** 设 $T \in \mathcal{L}(V)$ 是自伴的，$U$ 是 $V$ 的在 $T$ 下不变的子空间.

(a) 证明：$U^\perp$ 在 $T$ 下不变.

(b) 证明；$T|_U \in \mathcal{L}(U)$ 是自伴的.

(c) 证明：$T|_{U^\perp} \in \mathcal{L}(U^\perp)$ 是自伴的.

[3]也就是 $T \in \mathcal{L}(V), T^* = -T$.

**20** 设 $T \in \mathcal{L}(V)$ 是正规的，$U$ 是 $V$ 的在 $T$ 下不变的子空间.

(a) 证明：$U^{\perp}$ 在 $T$ 下不变.

(b) 证明：$U$ 在 $T^*$ 下不变.

(c) 证明：$(T|_U)^* = (T^*)|_U$.

(d) 证明：$T|_U \in \mathcal{L}(U)$ 和 $T|_{U^{\perp}} \in \mathcal{L}(U^{\perp})$ 是正规算子.

注 本题可用于给出复谱定理的又一证明（对 $\dim V$ 用数学归纳法，而且要用到"$T$ 有特征向量"这一结果）.

**21** 设 $T$ 是有限维内积空间上的自伴算子，且 2 和 3 是 $T$ 仅有的特征值. 证明：

$$T^2 - 5T + 6I = 0.$$

**22** 给出一例：$T \in \mathcal{L}(\mathbf{C}^3)$，2 和 3 是 $T$ 仅有的特征值，而 $T^2 - 5T + 6I \neq 0$.

**23** 设 $T \in \mathcal{L}(V)$ 是自伴的，$\lambda \in \mathbf{F}$ 且 $\epsilon > 0$. 设存在 $v \in V$，使得 $\|v\| = 1$ 且

$$\|Tv - \lambda v\| < \epsilon.$$

证明：$T$ 有特征值 $\lambda'$ 使得 $|\lambda - \lambda'| < \epsilon$.

注 本题表明，对于自伴算子而言，一个数如果接近于满足使其成为特征值的等式，那么它就接近于特征值.

**24** 设 $U$ 是有限维向量空间，$T \in \mathcal{L}(U)$.

(a) 设 $\mathbf{F} = \mathbf{R}$. 证明：$T$ 是可对角化的，当且仅当存在 $U$ 的一个基使得 $T$ 关于这个基的矩阵等于其转置.

(b) 设 $\mathbf{F} = \mathbf{C}$. 证明：$T$ 是可对角化的，当且仅当存在 $U$ 的一个基使得 $T$ 关于这个基的矩阵与其共轭转置是可交换的.

注 本题在 5.55 所示的可对角化的一系列等价条件之上又添加了一条.

**25** 设 $T \in \mathcal{L}(V)$ 且 $V$ 有一规范正交基 $e_1, \ldots, e_n$ 由 $T$ 的对应于特征值 $\lambda_1, \ldots, \lambda_n$ 的特征向量构成. 证明：如果 $k \in \{1, \ldots, n\}$，那么伪逆 $T^{\dagger}$ 满足等式

$$T^{\dagger} e_k = \begin{cases} \frac{1}{\lambda_k} e_k & \text{若 } \lambda_k \neq 0, \\ 0 & \text{若 } \lambda_k = 0. \end{cases}$$

## 7C 正算子

**7.34 定义：正算子（positive operator）**

算子 $T \in \mathcal{L}(V)$ 称为**正的**，如果 $T$ 是自伴的且对所有 $v \in V$ 有

$$\langle Tv, v\rangle \geq 0.$$

♣

如果 $V$ 是复向量空间，那么以上定义中"$T$ 是自伴的"这一条件可以去掉（根据 7.14）.

---

**7.35 例：正算子**

(a) 令算子 $T \in \mathcal{L}(\mathbf{F}^2)$ 的矩阵（关于标准基）是 $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. 那么 $T$ 是自伴的，而且 $\big\langle T(w,z),(w,z)\big\rangle = 2|w|^2 - 2\operatorname{Re}(w\bar{z}) + |z|^2 = |w-z|^2 + |w|^2 \geq 0$ 对所有 $(w,z) \in \mathbf{F}^2$ 成立. 因此 $T$ 是正算子.

(b) 如果 $U$ 是 $V$ 的子空间，那么正交投影 $P_U$ 是正算子. 你应该验证一下.

(c) 如果 $T \in \mathcal{L}(V)$ 是自伴的，且 $b, c \in \mathbf{R}$ 使得 $b^2 < 4c$，那么 $T^2 + bT + cI$ 是正算子，如 7.26 的证明所示.

---

**7.36 定义：平方根（square root）**

算子 $R$ 称为算子 $T$ 的**平方根**，如果 $R^2 = T$.

♣

---

**7.37 例：算子的平方根**

如果 $T \in \mathcal{L}(\mathbf{F}^3)$ 定义为 $T(z_1, z_2, z_3) = (z_3, 0, 0)$，那么定义为 $R(z_1, z_2, z_3) = (z_2, z_3, 0)$ 的算子 $R \in \mathcal{L}(\mathbf{F}^3)$ 是 $T$ 的一个平方根，因为 $R^2 = T$（你可以验证）.

---

因为正算子对应于非负数，所以更好的术语应该是"非负算子".然而，算子理论家们一贯称之为正算子，所以我们也遵循习惯.一些数学家用的是**"正半定算子"（positive semidefinite operator）**，意思和"正算子"一样.

接下来这条结果中的正算子的特性，对应于 $\mathbf{C}$ 中非负数的特性. 具体而言，$z \in \mathbf{C}$ 非负当且仅当它有非负平方根，这对应于条件 (d)；$z$ 非负当且仅当它有实平方根，这对应于条件 (e)；$z$ 非负当且仅当存在 $w \in \mathbf{C}$ 使得 $z = \bar{w}w$，这对应于条件 (f). 等价于"某算子是正算子"的又一条件，见本节习题 20.

**7.38 正算子的特性**

令 $T \in \mathcal{L}(V)$. 那么下列等价：

(a) $T$ 是正算子.

(b) $T$ 自伴且所有特征值非负.

(c) 关于 $V$ 的某个规范正交基，$T$ 的矩阵是对角矩阵且对角线上仅有非负数.

(d) $T$ 有正平方根.

(e) $T$ 有自伴平方根.

(f) 存在某个 $R \in \mathcal{L}(V)$ 使得 $T = R^*R$.

♡

**证明** 我们将证明：(a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (e) $\Longrightarrow$ (f) $\Longrightarrow$ (a).

首先假设 (a) 成立，那么 $T$ 是正的，也就意味着 $T$ 是自伴的（根据正算子的定义）. 为证明 (b) 中的另一条件，设 $\lambda$ 是 $T$ 的特征值. 令 $v$ 是 $T$ 的对应于 $\lambda$ 的特征向量. 那么

$$0 \le \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda \langle v, v\rangle .$$

因此，$\lambda$ 是非负数. 于是，(b) 成立，也就表明 (a) 蕴涵 (b).

现在假设 (b) 成立，那么 $T$ 自伴且所有特征值非负. 由谱定理（7.29 和 7.31），$V$ 有一规范正交基 $e_1, \dots, e_n$ 由 $T$ 的特征向量组成. 令 $\lambda_1, \dots, \lambda_n$ 是 $T$ 的对应于 $e_1, \dots, e_n$ 的特征值，那么每个 $\lambda_k$ 都是非负数. $T$ 关于 $e_1, \dots, e_n$ 的矩阵是以 $\lambda_1, \dots, \lambda_n$ 为对角线元素的对角矩阵，这表明 (b) 蕴涵 (c).

现在假设 (c) 成立. 设 $e_1, \dots, e_n$ 是 $V$ 的规范正交基，使得 $T$ 关于这个基的矩阵是对角矩阵，且对角线上是非负数 $\lambda_1, \dots, \lambda_n$. 由线性映射引理（3.4）可得，存在 $R \in \mathcal{L}(V)$ 使得

$$Re_k = \sqrt{\lambda_k} e_k$$

对任一 $k = 1, \dots, n$ 都成立. 你应自行验证，$R$ 是正算子. 此外，有 $R^2 e_k = \lambda_k e_k = Te_k$ 对任一 $k$ 都成立. 这意味着 $R^2 = T$. 因此，$R$ 是 $T$ 的正平方根. 于是，(d) 成立，这表明 (c) 蕴涵 (d).

每个正算子都是自伴的（根据正算子的定义），于是 (d) 蕴涵 (e).

现在假设 (e) 成立，意即 $V$ 上存在一自伴算子 $R$ 使得 $T = R^2$. 那么 $T = R^*R$（因为 $R^* = R$）. 于是 (e) 蕴涵 (f).

最后，假设 (f) 成立. 令 $R \in \mathcal{L}(V)$ 使得 $T = R^*R$. 那么 $T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$. 于是 $T$ 自伴. 离证明 (a) 成立就差一步——注意到

$$\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle \ge 0$$

对任一 $v \in V$ 都成立. 因此 $T$ 是正的，这表明 (f) 蕴涵 (a). ■

每个非负数都有唯一非负平方根. 接下来这条结果表明正算子也有类似的性质.

**7.39 每个正算子都只有一个正平方根**

$V$ 上的每个正算子都有唯一正平方根.

**证明** 设 $T \in \mathcal{L}(V)$ 是正的. 设 $v \in V$ 是 $T$ 的特征向量. 于是存在实数 $\lambda \ge 0$ 使得 $Tv = \lambda v$.

令 $R$ 是 $T$ 的正平方根. 我们将证明 $Rv = \sqrt{\lambda} v$，这意味着 $R$ 对 $T$ 的特征向量的作用是唯一确定的. 因为 $V$ 中存在由 $T$ 的特征向量组成的基（根据谱定理），所以这意味着 $R$ 是唯一确定的.

> 一个正算子可以有无穷多个平方根（但是它们当中只能有一个是正的）. 例如，如果 $\dim V > 1$ 的话，$V$ 上的恒等算子就有无穷多个平方根.

为了证明 $Rv = \sqrt{\lambda} v$，注意到，谱定理指出 $V$ 中存在由 $R$ 的特征向量组成的规范正交基 $e_1, \dots, e_n$. 因为 $R$ 是正算子，所以它的所有特征值非负. 因此存在非负数 $\lambda_1, \dots, \lambda_n$ 使得 $Re_k = \sqrt{\lambda_k} e_k$ 对任一 $k = 1, \dots, n$ 都成立.

因为 $e_1, \dots, e_n$ 是 $V$ 的基，所以我们可以写出

$$v = a_1 e_1 + \cdots + a_n e_n ,$$

其中 $a_1,\ldots,a_n\in\mathbf{F}$. 因此

$$Rv=a_1\sqrt{\lambda_1}e_1+\cdots+a_n\sqrt{\lambda_n}e_n.$$

于是,

$$\lambda v=Tv=R^2v=a_1\lambda_1e_1+\cdots+a_n\lambda_ne_n.$$

上式表明

$$a_1\lambda e_1+\cdots+a_n\lambda e_n=a_1\lambda_1e_1+\cdots+a_n\lambda_ne_n.$$

因此, $a_k(\lambda-\lambda_k)=0$ 对任一 $k=1,\ldots,n$ 都成立. 于是

$$v=\sum_{\{k:\lambda_k=\lambda\}}a_ke_k.$$

因此,

$$Rv=\sum_{\{k:\lambda_k=\lambda\}}a_k\sqrt{\lambda}e_k=\sqrt{\lambda}v,$$

命题得证. ■

由于上面这条结果,下面定义的这个记号就有了意义.

**7.40 记号: $\sqrt{T}$**

对于正算子 $T$, $\sqrt{T}$ 表示 $T$ 的唯一正平方根. ♣

**7.41 例:正算子的平方根**

在(具有通常的欧几里得内积的) $\mathbf{R}^2$ 上定义算子 $S,T$ 为

$$S(x,y)=(x,2y)\quad 和\quad T(x,y)=(x+y,x+y).$$

那么,关于 $\mathbf{R}^2$ 的标准基,我们有

$$\mathcal{M}(S)=\begin{pmatrix}1&0\\0&2\end{pmatrix}\quad 和\quad \mathcal{M}(T)=\begin{pmatrix}1&1\\1&1\end{pmatrix}. \tag{7.42}$$

这两个矩阵都等于自身的转置,因此 $S$ 和 $T$ 都是自伴的.

如果 $(x,y)\in\mathbf{R}^2$,那么

$$\langle S(x,y),(x,y)\rangle=x^2+2y^2\geq 0$$

且

$$\langle T(x,y),(x,y)\rangle=x^2+2xy+y^2=(x+y)^2\geq 0.$$

因此, $S$ 和 $T$ 都是正算子.

$\mathbf{R}^2$ 的标准基是由 $S$ 的特征向量组成的规范正交基. 注意到

$$\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$$

是 $T$ 的特征向量所成的规范正交基,其中第一个特征向量对应的特征值为 2,第二个对应的特征值为 0. 因此 $\sqrt{T}$ 的特征向量与 $T$ 的相同,而特征值为 $\sqrt{2}$ 和 0.

转下页

通过证明以下这些矩阵的平方就是式 (7.42) 中的矩阵，并且以下每个矩阵都是正算子关于标准基的矩阵，你可以验证：

$$\mathcal{M}\left(\sqrt{S}\right)=\begin{pmatrix}1 & 0\\ 0 & \sqrt{2}\end{pmatrix} \quad 和 \quad \mathcal{M}\left(\sqrt{T}\right)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}.$$

接下来这条结果的表述并不涉及平方根，但是这个简洁的证明很好地利用了正算子的平方根.

**7.43 $T$ 是正的且 $\langle Tv, v\rangle = 0 \Longrightarrow Tv = 0$**

设 $T$ 是 $V$ 上的正算子且 $v \in V$ 使得 $\langle Tv, v\rangle = 0$，那么 $Tv = 0$.

**证明**

$$0=\langle Tv, v\rangle=\left\langle\sqrt{T}\sqrt{T}v, v\right\rangle=\left\langle\sqrt{T}v, \sqrt{T}v\right\rangle=\left\|\sqrt{T}v\right\|^2.$$

于是，$\sqrt{T}v = 0$. 所以 $Tv=\sqrt{T}\left(\sqrt{T}v\right)=0$，命题得证. ■

## 习题 7C

**1** 设 $T \in \mathcal{L}(V)$. 证明：如果 $T$ 和 $-T$ 都是正算子，那么 $T = 0$.

**2** 设算子 $T \in \mathcal{L}(\mathbf{F}^4)$（关于标准基）的矩阵是

$$\begin{pmatrix}2 & -1 & 0 & 0\\ -1 & 2 & -1 & 0\\ 0 & -1 & 2 & -1\\ 0 & 0 & -1 & 2\end{pmatrix}.$$

证明：$T$ 是可逆正算子.

**3** 设 $n$ 是正整数，算子 $T \in \mathcal{L}(\mathbf{F}^n)$（关于标准基）的矩阵的各元素全为 1. 证明：$T$ 是正算子.

**4** 设 $n$ 是整数，$n > 1$. 证明：存在 $n \times n$ 矩阵 $A$，其所有元素都为正数且 $A = A^*$，但是 $\mathbf{F}^n$ 上（关于标准基）以 $A$ 为矩阵的算子不是正算子.

**5** 设 $T \in \mathcal{L}(V)$ 是自伴的. 证明：$T$ 是正算子，当且仅当对于 $V$ 的任一规范正交基 $e_1, \ldots, e_n$，都有 $\mathcal{M}(T, (e_1, \ldots, e_n))$ 的对角线元素全为非负数.

**6** 证明：$V$ 上的两正算子之和是正算子.

**7** 设 $S \in \mathcal{L}(V)$ 是可逆正算子，$T \in \mathcal{L}(V)$ 是正算子. 证明：$S + T$ 可逆.

**8** 设 $T \in \mathcal{L}(V)$. 证明：$T$ 是正算子，当且仅当伪逆 $T^\dagger$ 是正算子.

**9** 设 $T \in \mathcal{L}(V)$ 是正算子，$S \in \mathcal{L}(W, V)$. 证明：$S^*TS$ 是 $W$ 上的正算子.

**10** 设 $T$ 是 $V$ 上的正算子. 设 $v, w \in V$ 使得

$$Tv = w \quad 且 \quad Tw = v.$$

证明：$v = w$.

**11** 设 $T$ 是 $V$ 上的正算子，$U$ 是 $V$ 的在 $T$ 下不变的子空间．证明：$T|_U \in \mathcal{L}(U)$ 是 $U$ 上的正算子．

**12** 设 $T \in \mathcal{L}(V)$ 是正算子．证明：对于任一正整数 $k$，$T^k$ 都是正算子．

**13** 设 $T \in \mathcal{L}(V)$ 是自伴的，$\alpha \in \mathbf{R}$.

(a) 证明：$T - \alpha I$ 是正算子，当且仅当 $\alpha$ 小于或等于 $T$ 的每个特征值.

(b) 证明：$\alpha I - T$ 是正算子，当且仅当 $\alpha$ 大于或等于 $T$ 的每个特征值.

**14** 设 $T$ 是 $V$ 上的正算子，$v_1, \ldots, v_m \in V$．证明：

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle T v_k, v_j \rangle \geq 0.$$

**15** 设 $T \in \mathcal{L}(V)$ 是自伴的．证明：存在正算子 $A, B \in \mathcal{L}(V)$ 使得

$$T = A - B \quad 且 \quad \sqrt{T^*T} = A + B \quad 且 \quad AB = BA = 0.$$

**16** 设 $T$ 是 $V$ 上的正算子．证明：

$$\operatorname{null} \sqrt{T} = \operatorname{null} T \quad 且 \quad \operatorname{range} \sqrt{T} = \operatorname{range} T.$$

**17** 设 $T \in \mathcal{L}(V)$ 是正算子．证明：存在实系数多项式 $p$ 使得 $\sqrt{T} = p(T)$.

**18** 设 $S$ 和 $T$ 是 $V$ 上的正算子．证明：$ST$ 是正算子，当且仅当 $S$ 和 $T$ 可交换.

**19** 证明：$\mathbf{F}^2$ 上的恒等算子具有无穷多个自伴的平方根.

**20** 设 $T \in \mathcal{L}(V)$，$e_1, \ldots, e_n$ 是 $V$ 的规范正交基. 证明：$T$ 是正算子，当且仅当存在 $v_1, \ldots, v_n \in V$ 使得

$$\langle T e_k, e_j \rangle = \langle v_k, v_j \rangle$$

对所有 $j, k = 1, \ldots, n$ 成立.

注 $\{\langle T e_k, e_j \rangle\}_{j,k=1,\ldots,n}$ 就是 $T$ 关于规范正交基 $e_1, \ldots, e_n$ 的矩阵元素.

**21** 设 $n$ 是正整数．$n \times n$ 的**希尔伯特矩阵（Hilbert matrix）**，是第 $j$ 行第 $k$ 列元素为 $\frac{1}{j+k-1}$ 的 $n \times n$ 矩阵．设算子 $T \in \mathcal{L}(V)$ 关于 $V$ 的某个规范正交基的矩阵是 $n \times n$ 希尔伯特矩阵．证明：$T$ 是可逆正算子.

注 例如：$4 \times 4$ 希尔伯特矩阵是

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

**22** 设 $T \in \mathcal{L}(V)$ 是正算子，$u \in V$ 满足 $\|u\| = 1$ 且 $\|Tu\| \geq \|Tv\|$ 对所有 $\|v\| = 1$ 的 $v \in V$ 成立．证明：$u$ 是 $T$ 的特征向量，对应于其最大的特征值.

**23** 对于 $T \in \mathcal{L}(V)$ 和 $u, v \in V$，定义 $\langle u, v \rangle_T$ 为 $\langle u, v \rangle_T = \langle Tu, v \rangle$.

(a) 设 $T \in \mathcal{L}(V)$．证明：$\langle \cdot, \cdot \rangle_T$ 是 $V$ 上的内积，当且仅当 $T$ 是（关于原始内积 $\langle \cdot, \cdot \rangle$ 的）可逆正算子.

(b) 证明：$V$ 上的任一内积都具有 $\langle \cdot, \cdot \rangle_T$ 的形式，其中 $T \in \mathcal{L}(V)$ 是某个可逆正算子.

**24** 设 $S$ 和 $T$ 是 $V$ 上的正算子．证明：

$$\operatorname{null}(S+T) = \operatorname{null} S \cap \operatorname{null} T.$$

**25** 令 $T$ 是 7A 节习题 31 (b) 中的二阶求导算子．证明：$-T$ 是正算子．

# 7D 等距映射、幺正算子和矩阵分解

## 等距映射

保持范数的线性映射实在重要，值得有个名字.

**7.44 定义：等距映射（isometry）**

线性映射 $S \in \mathcal{L}(V,W)$ 被称为**等距映射**，如果

$$\|Sv\| = \|v\|$$

对任一 $v \in V$ 都成立．换句话说，保持范数的线性映射就是等距映射.

如果 $S \in \mathcal{L}(V,W)$ 是等距映射且 $v \in V$ 使得 $Sv = 0$，那么 $\|v\| = \|Sv\| = \|0\| = 0$，这意味着 $v = 0$．因此每个等距映射都是单射.

希腊语单词 **isos** 意为相等、**metron** 意为度量，所以 **isometry** 的字面意思就是度量相等.

**7.45 例：将规范正交基映射到规范正交组 $\Longrightarrow$ 等距映射**

设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，$g_1, \ldots, g_n$ 是 $W$ 中的规范正交组．令 $S \in \mathcal{L}(V,W)$ 是使得对任一 $k = 1, \ldots, n$ 都有 $Se_k = g_k$ 的线性映射．下面证明 $S$ 是等距映射．设 $v \in V$．那么有

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n \tag{7.46}$$

以及

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 . \tag{7.47}$$

其中我们用到了 6.30 (b)．将 $S$ 作用于式 (7.46) 的两侧，得到

$$Sv = \langle v, e_1 \rangle Se_1 + \cdots + \langle v, e_n \rangle Se_n = \langle v, e_1 \rangle g_1 + \cdots + \langle v, e_n \rangle g_n .$$

因此，

$$\|Sv\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 . \tag{7.48}$$

对比式 (7.47) 和式 (7.48)，可知 $\|v\| = \|Sv\|$．因此，$S$ 是等距映射.

下一条结果给出了“某线性映射是等距映射”的等价条件．(a) 和 (c) 的等价关系表明，线性映射是等距映射当且仅当它是保持内积的．(a) 和 (d) 的等价关系表明，线性映射是等距映射当且仅当它将某个规范正交基映射为规范正交组．因此例 7.45 涵盖了所有等距映射．更进一步还有，线性映射是等距映射，当且仅当它将每个规范正交基都映射为规范正交组——因为 (a) 的成立与否并不依赖于基 $e_1, \ldots, e_n$ 的选取.

接下来这条结果中，(a) 和 (e) 的等价关系表明，线性映射是等距映射当且仅当它（关于任一规范正交基）的矩阵的列形成一规范正交组．这里我们将 $m \times n$ 矩阵的列等同于 $\mathbf{F}^m$ 的元素，然后应用 $\mathbf{F}^m$ 上的欧几里得内积.

**7.49 等距映射的特性**

设 $S \in \mathcal{L}(V, W)$. 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，$f_1, \ldots, f_m$ 是 $W$ 的规范正交基. 那么下列等价：

(a) $S$ 是等距映射.

(b) $S^*S = I$.

(c) $\langle Su, Sv \rangle = \langle u, v \rangle$ 对所有 $u, v \in V$ 成立.

(d) $Se_1, \ldots, Se_n$ 是 $W$ 中的规范正交组.

(e) $\mathcal{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ 的列形成 $\mathbf{F}^m$ 中关于欧几里得内积的规范正交组.

**证明** 首先假设 (a) 成立，那么 $S$ 是等距映射. 如果 $v \in V$，那么

$$\langle (I - S^*S)v, v \rangle = \langle v, v \rangle - \langle S^*Sv, v \rangle = \|v\|^2 - \langle Sv, Sv \rangle = \|v\|^2 - \|Sv\|^2 = 0.$$

于是自伴算子 $I - S^*S$ 等于 0（根据 7.16）. 因此 $S^*S = I$，证明了 (a) 蕴涵 (b).

现在假设 (b) 成立，那么 $S^*S = I$. 如果 $u, v \in V$，那么

$$\langle Su, Sv \rangle = \langle S^*Su, v \rangle = \langle Iu, v \rangle = \langle u, v \rangle,$$

证明了 (b) 蕴涵 (c).

现在假设 (c) 成立，那么 $\langle Su, Sv \rangle = \langle u, v \rangle$ 对所有 $u, v \in V$ 成立. 因此，如果 $j, k \in \{1, \ldots, n\}$，那么

$$\langle Se_j, Se_k \rangle = \langle e_j, e_k \rangle.$$

于是，$Se_1, \ldots, Se_n$ 是 $W$ 中的规范正交组，证明了 (c) 蕴涵 (d).

现在假设 (d) 成立，那么 $Se_1, \ldots, Se_n$ 是 $W$ 中的规范正交组.令 $A = \mathcal{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$. 如果 $k, r \in \{1, \ldots, n\}$，那么

$$\sum_{j=1}^{m} A_{j,k} \overline{A_{j,r}} = \left\langle \sum_{j=1}^{m} A_{j,k} f_j, \sum_{j=1}^{m} A_{j,r} f_j \right\rangle = \langle Se_k, Se_r \rangle = \begin{cases} 1 & \text{若 } k = r, \\ 0 & \text{若 } k \neq r. \end{cases} \tag{7.50}$$

式 (7.50) 的左侧是 $A$ 的第 $k$ 列和第 $r$ 列在 $\mathbf{F}^m$ 中的内积. 因此 $A$ 的列形成了 $\mathbf{F}^m$ 中的规范正交组，证明了 (d) 蕴涵 (e).

现在假设 (e) 成立，那么上一段定义的矩阵 $A$ 的列形成 $\mathbf{F}^m$ 中的规范正交组.那么式 (7.50) 表明 $Se_1, \ldots, Se_n$ 是 $W$ 中的规范正交组.所以，根据例 7.45（$Se_1, \ldots, Se_n$ 相当于其中的 $g_1, \ldots, g_n$），可得 $S$ 是等距映射，证明了 (e) 蕴涵 (a). ■

“某线性映射是等距映射”的更多等价条件，见本节习题 1、11.

## 幺正算子

在本小节，我们将注意力限制在从向量空间到自身的映射. 换句话说，我们讨论算子.

**7.51 定义：幺正算子（unitary operator）**

算子 $S \in \mathcal{L}(V)$ 被称为**幺正的**，如果 $S$ 是可逆等距映射.

前面提到过，每个等距映射都是单射. 而有限维向量空间上的每个单射算子都是可逆的（见 3.65），“$V$ 是有限维内积空间”又是本章总成立的一个假设，因此我们可以从上面的定义中移去“可逆”一词，而不改变其含义. 但这里保留“可逆”这个不必要的词，是为了定义的一致性，毕竟读者在学习不一定有限维的内积空间时，遇到的应该还是上面这个定义.

虽然对于有限维内积空间上的算子而言，幺正（unitary）和等距（isometry）是同一个意思；但是要记得，幺正算子将向量空间映射到自身，而等距映射将向量空间映射到另一（可能不同的）向量空间.

**7.52 例：$\mathbf{R}^2$ 的旋转**

设 $\theta \in \mathbf{R}$ 且 $S$ 是 $\mathbf{F}^2$ 上的算子，其关于 $\mathbf{F}^2$ 的标准基的矩阵是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

该矩阵的这两列形成了 $\mathbf{F}^2$ 中的规范正交组，于是 $S$ 是等距映射【根据 7.49 中 (a) 和 (e) 的等价关系】. 因此，$S$ 是幺正算子.

如果 $\mathbf{F} = \mathbf{R}$，那么 $S$ 是“绕 $\mathbf{R}^2$ 的原点逆时针旋转 $\theta$ 弧度”这一算子. 观察到这一点，能让我们用另一种方式来思考 $S$ 为什么是等距映射——因为绕 $\mathbf{R}^2$ 的原点旋转是保持范数的.

下一条结果（7.53）列出了“某算子是幺正算子”的几个等价条件. 7.49 中所有等价于“某线性映射是等距映射”的条件都应该纳入其中；除此之外，7.53 中出现了另一些条件，这是因为我们将讨论的范围限制在从向量空间到自身的线性映射当中. 例如，7.49 表明线性映射 $S \in \mathcal{L}(V,W)$ 是等距映射当且仅当 $S^*S = I$，而 7.53 表明算子 $S \in \mathcal{L}(V)$ 是幺正算子当且仅当 $S^*S = SS^* = I$.

7.49 与 7.53 的差异还有：7.49 (d) 提到的是规范正交组，而 7.53 (d) 提到的是规范正交基；7.49 (e) 提到的是 $\mathcal{M}(T)$ 的列，而 7.53 (e) 提到的是 $\mathcal{M}(T)$ 的行；7.49 (e) 中的 $\mathcal{M}(T)$ 是关于 $V$ 中一个规范正交基和 $W$ 中一个规范正交基的，而 7.53 (e) 中的 $\mathcal{M}(T)$ 是仅关于 $V$ 中一个基的——这个基“身兼二职”.

**7.53 幺正算子的特性**

设 $S \in \mathcal{L}(V)$. 设 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基. 那么下列等价：

(a) $S$ 是幺正算子.

(b) $S^*S = SS^* = I$.

(c) $S$ 可逆且 $S^{-1} = S^*$.

(d) $Se_1, \ldots, Se_n$ 是 $V$ 的规范正交基.

(e) $\mathcal{M}(S, (e_1, \ldots, e_n))$ 的行形成 $\mathbf{F}^n$ 中关于欧几里得内积的规范正交基.

(f) $S^*$ 是幺正算子.

**证明** 首先假设 (a) 成立，那么 $S$ 是幺正算子. 于是，根据 7.49 中 (a) 和 (b) 的等价关系，有

$$S^*S = I.$$

等式两侧同时右乘 $S^{-1}$ 可得 $S^* = S^{-1}$. 因此 $SS^* = SS^{-1} = I$，待证等式得证，进而 (a) 蕴涵 (b) 得证.

可逆性和逆的定义表明，(b) 蕴涵 (c).

现在假设 (c) 成立，那么 $S$ 可逆且 $S^{-1} = S^*$. 因此，$S^*S = I$. 于是，根据 7.49 中 (b) 和 (d) 的等价关系，可得 $Se_1, \ldots, Se_n$ 是 $V$ 中的规范正交组. 这个组的长度又等于 $\dim V$，因此 $Se_1, \ldots, Se_n$ 是 $V$ 的规范正交基，证明了 (c) 蕴涵 (d).

现在假设 (d) 成立，那么 $Se_1, \ldots, Se_n$ 是 $V$ 的规范正交基. 7.49 中 (a) 和 (d) 的等价关系表明 $S$ 是幺正算子. 因此

$$(S^*)^*S^* = SS^* = I.$$

其中最后一个等号成立，是因为我们已经证明了本结果中的 (a) 蕴涵 (b). 上式和 7.49 中 (a) 和 (b) 的等价关系表明 $S^*$ 是等距映射. 因此，$\mathcal{M}(S^*, (e_1, \ldots, e_n))$ 的列形成了 $\mathbf{F}^n$ 的规范正交基【根据 7.49 中 (a) 和 (e) 的等价关系】. $\mathcal{M}(S, (e_1, \ldots, e_n))$ 的行是 $\mathcal{M}(S^*, (e_1, \ldots, e_n))$ 的列的复共轭. 因此，$\mathcal{M}(S, (e_1, \ldots, e_n))$ 的行形成 $\mathbf{F}^n$ 的规范正交基，证明了 (d) 蕴涵 (e).

现在假设 (e) 成立，那么 $\mathcal{M}(S^*, (e_1, \ldots, e_n))$ 的列形成 $\mathbf{F}^n$ 的规范正交基. 7.49 中 (a) 和 (e) 之间的等价关系表明，$S^*$ 是等距映射，证明了 (e) 蕴涵 (f).

现在假设 (f) 成立，那么 $S^*$ 是幺正算子. 在这一结果中，我们已经证明了的蕴涵关系链表明 (a) 蕴涵 (f). 对 $S^*$ 应用这一结果，可证明 $(S^*)^*$ 是幺正算子，也就证明了 (f) 蕴涵 (a).

我们证明了 (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (e) $\Longrightarrow$ (f) $\Longrightarrow$ (a)，从而完成了证明. ■

回顾一下我们在 $\mathbf{C}$ 与 $\mathcal{L}(V)$ 之间所做的类比. 在这一类比下，复数 $z$ 对应于算子 $S \in \mathcal{L}(V)$，$\bar{z}$ 对应于 $S^*$. 实数（$z = \bar{z}$）对应于自伴算子（$S = S^*$），而非负数对应于（没取对名字的）正算子.

$\mathbf{C}$ 的另一个重要的子集是单位圆（unit circle），由满足 $|z| = 1$ 的复数 $z$ 构成. $|z| = 1$ 这一条件等价于 $\bar{z}z = 1$ 这一条件. 在我们的类比下，这对应于 $S^*S = I$ 这一条件，等价于 $S$ 是幺正算子. 于是，这一类比表明，$\mathbf{C}$ 中的单位圆对应于幺正算子构成的集合. 在接下来的两条结果中，这一类比体现在了幺正算子的特征值上. 这一类比的另一例子见习题 15.

**7.54 幺正算子的特征值绝对值是 1**

设 $\lambda$ 是幺正算子的特征值，那么 $|\lambda| = 1$.

**证明** 设 $S \in \mathcal{L}(V)$ 是幺正算子且 $\lambda$ 是 $S$ 的特征值. 令 $v \in V$（$v \neq 0$）使得 $Sv = \lambda v$. 那么有

$$|\lambda|\,\|v\| = \|\lambda v\| = \|Sv\| = \|v\|.$$

因此，$|\lambda| = 1$，命题得证. ■

接下来这条结果，以复谱定理为主要工具，刻画了有限维复内积空间上的幺正算子.

**7.55 对复内积空间上的幺正算子的描述**

设 $\mathbf{F} = \mathbf{C}$ 且 $S \in \mathcal{L}(V)$. 那么下列等价：

(a) $S$ 是幺正算子.

(b) 存在 $V$ 的一规范正交基由 $S$ 的特征向量组成，其对应的特征值的绝对值都是 1.

**证明** 假设 (a) 成立，那么 $S$ 是幺正算子. 7.53 中 (a) 和 (b) 的等价关系表明 $S$ 是正规的. 因此，由复谱定理（7.31）可得，存在 $V$ 的一规范正交基由 $S$ 的特征向量组成. 而 $S$ 的每个特征值的绝对值都是 1（根据 7.54），也就完成了 (a) 蕴涵 (b) 的证明.

现在假设 (b) 成立. 令 $e_1, \ldots, e_n$ 是 $V$ 的规范正交基，由 $S$ 的特征向量组成，其对应特征值 $\lambda_1, \ldots, \lambda_n$ 的绝对值都为 1. 那么 $Se_1, \ldots, Se_n$ 也是 $V$ 的规范正交基，这是因为

$$\langle Se_j, Se_k \rangle = \langle \lambda_j e_j, \lambda_k e_k \rangle = \lambda_j \overline{\lambda_k} \langle e_j, e_k \rangle = \begin{cases} 0 & \text{若 } j \neq k, \\ 1 & \text{若 } j = k, \end{cases}$$

对所有 $j, k = 1, \ldots, n$ 成立. 因此，由 7.53 中 (a) 和 (d) 的等价关系可得，$S$ 是幺正的，证明了 (b) 蕴涵 (a). ■

## QR 分解

在本小节，我们将注意力从算子转向矩阵. 这有助于你培养将算子等同于方阵的思维方式（在选取了算子的定义空间的基之后）. 你还会更加熟练地在算子和方阵之间，将种种概念和结果来回转换.

当我们从 $n \times n$ 矩阵而非算子出发时，除非另有说明，否则假设相关的算子都在（具有欧几里得内积的）$\mathbf{F}^n$ 上，并且它们的矩阵都是关于 $\mathbf{F}^n$ 的标准基得到的.

我们首先做如下定义，将幺正算子的概念转换为幺正矩阵.

**7.56 定义：幺正矩阵（unitary matrix）**

一 $n \times n$ 矩阵被称为**幺正的**，如果它的列形成 $\mathbf{F}^n$ 中的规范正交组. ♣

在以上定义中，我们可以把"$\mathbf{F}^n$ 中的规范正交组"替换为"$\mathbf{F}^n$ 的规范正交基"，因为 $n$ 维内积空间中的长度为 $n$ 的每个规范正交组都是规范正交基. 如果 $S \in \mathcal{L}(V)$ 且 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$ 是 $V$ 的规范正交基，那么 $S$ 是幺正算子当且仅当 $\mathcal{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_n))$ 是幺正矩阵，这可由 7.49 中 (a) 和 (e) 的等价关系得到. 还要注意，我们可以将定义中的"列"替换为"行"，这用到了 7.53 中 (a) 和 (e) 的等价关系.

接下来这条结果，给出了"某方阵是幺正矩阵"的一些等价条件，其证明留作读者的习题. 在 (c) 中，$Qv$ 表示 $Q$ 和 $v$ 的矩阵乘积，这里将 $\mathbf{F}^n$ 的元素等同于 $n \times 1$ 矩阵（有时也称为列向量）. 下面的 (c) 中的范数是 $\mathbf{F}^n$ 上通常的欧几里得范数，其来自欧几里得内积. 在 (d) 中，$Q^*$ 表示矩阵 $Q$ 的共轭转置，对应于其所关联的算子的伴随.

**7.57 幺正矩阵的特性**

设 $Q$ 是 $n \times n$ 矩阵. 那么下列等价：

(a) $Q$ 是幺正矩阵.

(b) $Q$ 的行形成 $\mathbf{F}^n$ 中的规范正交组.

(c) $\|Qv\| = \|v\|$ 对任一 $v \in \mathbf{F}^n$ 都成立.

(d) $Q^*Q = QQ^* = I$，$I$ 是对角线上为 1、其余处处为 0 的 $n \times n$ 矩阵. ♡

下面陈述并证明的 QR 分解，是被广泛使用的 QR 算法（此处不作讨论）中的主要工具，该算法用于求出方阵特征值和特征向量的良好近似. 在下面的结果中，如果矩阵 $A$ 属于 $\mathbf{F}^{n,n}$，那么矩阵 $Q$ 和 $R$ 也属于 $\mathbf{F}^{n,n}$.

**7.58 QR 分解（QR factorization）**

设 $A$ 是各列线性无关的方阵. 那么存在唯一一对矩阵 $Q$ 和 $R$，其中 $Q$ 是幺正的，而 $R$ 是上三角的且对角线上仅有正数，使得

$$A = QR.$$

**证明** 令 $v_1, \dots, v_n$ 表示 $A$ 的列，并将其视为 $\mathbf{F}^n$ 的元素. 对 $v_1, \dots, v_n$ 应用格拉姆-施密特过程（6.32），得到 $\mathbf{F}^n$ 的规范正交基 $e_1, \dots, e_n$，使得

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k) \tag{7.59}$$

对任一 $k = 1, \dots, n$ 都成立. 令 $R$ 是 $n \times n$ 矩阵，定义为

$$R_{j,k} = \langle v_k, e_j \rangle,$$

其中 $R_{j,k}$ 表示 $R$ 的第 $j$ 行第 $k$ 列的元素. 如果 $j > k$，那么 $e_j$ 正交于 $\operatorname{span}(e_1, \dots, e_k)$ 从而正交于 $v_k$【根据式 (7.59)】. 换句话说，如果 $j > k$，那么 $\langle v_k, e_j \rangle = 0$. 因此，$R$ 是上三角矩阵.

令 $Q$ 是各列为 $e_1, \dots, e_n$ 的幺正矩阵. 如果 $k \in \{1, \dots, n\}$，那么 $QR$ 的第 $k$ 列等于 $Q$ 各列的线性组合，其系数来自 $R$ 的第 $k$ 列——见 3.51 (a). 于是 $QR$ 的第 $k$ 列等于

$$\langle v_k, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k,$$

也就等于 $v_k$【根据 6.30 (a)】，即 $A$ 的第 $k$ 列. 因此，$A = QR$，即待证等式.

格拉姆-施密特过程的定义式（见 6.32）表明，每个 $v_k$ 都等于 $e_k$ 的正数倍加上 $e_1, \dots, e_{k-1}$ 的线性组合. 因此 $\langle v_k, e_k \rangle$ 是正数，于是 $R$ 的对角线上所有元素都是正数，这一点也得证.

最后，为证明 $Q$ 和 $R$ 是唯一的，假设我们还有 $A = \widehat{Q}\widehat{R}$，其中 $\widehat{Q}$ 是幺正的，而 $\widehat{R}$ 是上三角的且对角线上仅有正数. 令 $q_1, \dots, q_n$ 表示 $\widehat{Q}$ 的列. 像上面那样看待矩阵乘法，可以得到，每个 $v_k$ 都是 $q_1, \dots, q_k$ 的线性组合，其系数来自 $\widehat{R}$ 的第 $k$ 列. 这意味着 $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(q_1, \dots, q_k)$ 且 $\langle v_k, q_k \rangle > 0$. 而满足这些条件的规范正交组具有唯一性（见 6B 节习题 10），这表明 $q_k = e_k$ 对任一 $k = 1, \dots, n$ 都成立. 于是，$\widehat{Q} = Q$，也就意味着 $\widehat{R} = R$，从而完成了唯一性的证明. ■

QR 分解的证明表明，将格拉姆-施密特过程应用于待分解矩阵的列，可以得到其分解式中幺正矩阵的列. 接下来这个例子，基于我们刚刚完成的证明，呈现了 QR 分解的计算过程.

**7.60 例：$3 \times 3$ 矩阵的 QR 分解**

下面求矩阵

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix}$$

的 QR 分解. 我们按照 7.58 的证明，令 $v_1, v_2, v_3$ 等于 $A$ 的列：

$$v_1 = (1, 0, 0), \quad v_2 = (2, 1, 3), \quad v_3 = (1, -4, 2).$$

转下页

对 $v_1, v_2, v_3$ 应用格拉姆-施密特过程，得到规范正交组

$$e_1 = (1,0,0), \quad e_2 = \left(0, \frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right), \quad e_3 = \left(0, -\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right).$$

顺着 7.58 的证明，令 $Q$ 是以 $e_1, e_2, e_3$ 为列的幺正矩阵：

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}.$$

再按 7.58 的证明，令 $R$ 是 $3 \times 3$ 矩阵，第 $j$ 行第 $k$ 列的元素为 $\langle v_k, e_j \rangle$，得到

$$R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix}.$$

注意到 $R$ 确实是对角线上仅有正数的上三角矩阵，这正是 QR 分解所保证的.

现在，做矩阵乘法即可验证 $A = QR$ 就是待求的分解：

$$QR = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix} = A.$$

因此，$A = QR$，正如我们预期的那般.

---

在下一小节科列斯基分解（7.63）的证明中，QR 分解将是主要工具. QR 分解的另一个很好的应用，见阿达马不等式（9.66）的证明.

若一矩阵有 QR 分解，那么求解该矩阵对应的线性方程组就可以用这一分解式而不用高斯消元法. 具体来说，设 $A$ 是各列线性无关的 $n \times n$ 方阵，$b \in \mathbf{F}^n$，我们想要求解方程 $Ax = b$，其中 $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$（我们还是照常把 $\mathbf{F}^n$ 的元素等同于 $n \times 1$ 列向量）.

设 $A = QR$，其中 $Q$ 是幺正的，而 $R$ 是上三角的且对角线上仅有正数（$Q$ 和 $R$ 可以用格拉姆-施密特过程由 $A$ 得到，如 7.58 的证明所示）. 方程 $Ax = b$ 等价于 $QRx = b$. 等式两边同时左乘 $Q^*$，再用上 7.57 (d)，就可得到方程

$$Rx = Q^* b.$$

因为矩阵 $Q^*$ 是 $Q$ 的共轭转置，所以计算 $Q^*b$ 是很简单的. 因为 $R$ 是对角线上为正数的上三角矩阵，所以以上方程所表示的线性方程组，求解起来很快——先求 $x_n$，然后求 $x_{n-1}$，依次类推.

## 科列斯基分解

我们以从内积角度对可逆正算子的刻画作为本小节的开始.

**7.61 可逆正算子**

自伴算子 $T \in \mathcal{L}(V)$ 是可逆正算子，当且仅当 $\langle Tv, v \rangle > 0$ 对任意非零 $v \in V$ 都成立.

**证明** 首先假设 $T$ 是可逆正算子. 如果 $v \in V$ 且 $v \neq 0$，那么由 $T$ 可逆则有 $Tv \neq 0$. 这意味着 $\langle Tv, v\rangle \neq 0$（根据 7.43）. 于是 $\langle Tv, v\rangle > 0$.

为了证明另一方向的蕴涵关系，现在假设 $\langle Tv, v\rangle > 0$ 对任意非零 $v \in V$ 都成立. 那么 $Tv \neq 0$ 对任意非零 $v \in V$ 都成立. 于是 $T$ 是单射，从而 $T$ 可逆，命题得证. ■

接下来的定义将以上结果转换为矩阵语言. 这里我们用的是 $\mathbf{F}^n$ 上通常的欧几里得内积，并且将 $\mathbf{F}^n$ 中的元素等同于 $n \times 1$ 列向量.

**7.62 定义：正定（positive definite）**

矩阵 $B \in \mathbf{F}^{n,n}$ 称为**正定的**，如果 $B^* = B$ 且

$$\langle Bx, x\rangle > 0$$

对任意非零 $x \in \mathbf{F}^n$ 都成立. ♣

矩阵是上三角的当且仅当它的共轭转置是下三角的（意即对角线上方所有元素为 0）. 下面的这个分解，将正定矩阵写成下三角矩阵和它的共轭转置的乘积，它在数值线性代数[4]中具有重要意义.

接下来这条结果完全是关于矩阵的，不过它的证明也利用了算子和矩阵两者结论之间的等同关系. 在下面这条结果中，如果矩阵 $B$ 属于 $\mathbf{F}^{n,n}$，那么矩阵 $R$ 也属于 $\mathbf{F}^{n,n}$.

**7.63 科列斯基分解（Cholesky factorization）**

设 $B$ 是正定矩阵. 那么存在唯一一个对角线上仅含正数的上三角矩阵 $R$ 使得

$$B = R^*R.$$

♡

**证明** 因为 $B$ 是正定的，所以存在与 $B$ 大小相同的可逆矩阵 $A$ 使得 $B = A^*A$【根据 7.38 中 (a) 和 (f) 之间的等价关系】.

令 $A = QR$ 是 $A$ 的 QR 分解（见 7.58），其中 $Q$ 是幺正的，而 $R$ 是上三角的且对角线上仅有正数. 那么，$A^* = R^*Q^*$.

因此，

$$B = A^*A = R^*Q^*QR = R^*R,$$

即待证等式.

这一分解由安德烈-路易·科列斯基（André-Louis Cholesky，1875-1918）发现，在他去世后发表于 1924 年.

为证明该结果的唯一性部分，设 $S$ 是对角线上仅含正数的上三角矩阵，且 $B = S^*S$. 因为矩阵 $B$ 可逆，所以矩阵 $S$ 可逆（见 3D 节的习题 11）. 在等式 $B = S^*S$ 两侧右乘 $S^{-1}$ 可得 $BS^{-1} = S^*$.

[4]原文 computational linear algebra，直译为“计算线性代数”，但在中文里，这个词有点像动宾结构的短语，不太像名词，因而不太常用，同义且更常见的名称是“数值线性代数”（numerical linear algebra）. 因此本书统一翻译成“数值线性代数”.

令 $A$ 是该证明第一段的那个矩阵 $A$，那么有

$$\begin{aligned}(AS^{-1})^*(AS^{-1}) &= (S^*)^{-1}A^*AS^{-1}\\ &= (S^*)^{-1}BS^{-1}\\ &= (S^*)^{-1}S^*\\ &= I.\end{aligned}$$

因此，$AS^{-1}$ 是幺正的.

于是，$A = (AS^{-1})S$ 这一分解，将 $A$ 化为一个幺正矩阵和对角线上仅含正数的一个上三角矩阵的乘积. 而正如 7.58 中所述，QR 分解具有唯一性，这也就意味着 $S = R$. ■

以上证明的第一段中，我们本可以将 $A$ 选定为 $B$ 的平方根当中唯一的正定矩阵（见 7.39）. 然而这一证明选择了更一般的 $A$，这是因为对于特定的正定矩阵 $B$，求出不那么特殊的 $A$ 可能更容易.

## 习题 7D

**1** 设 $\dim V \geq 2$，$S \in \mathcal{L}(V,W)$. 证明：$S$ 是等距映射，当且仅当对于 $V$ 中任一长度为 2 的规范正交组 $e_1, e_2$，都有 $Se_1, Se_2$ 是 $W$ 中的规范正交组.

**2** 设 $T \in \mathcal{L}(V,W)$ 且 $T \neq 0$. 证明：$T$ 是一等距映射的标量倍，当且仅当 $T$ 保持正交性.

注 “$T$ 保持正交性”的意思是，对于所有满足 $\langle u,v\rangle = 0$ 的 $u,v \in V$ 都有 $\langle Tu,Tv\rangle = 0$.

**3** (a) 证明：$V$ 上的两幺正算子之积是幺正算子.

(b) 证明：$V$ 上的幺正算子的逆是幺正算子.

注 本题表明，$V$ 上的幺正算子构成的集合是一个群，这个群上的运算就是两算子之间通常的乘法.

**4** 设 $\mathbf{F} = \mathbf{C}$，且 $A, B \in \mathcal{L}(V)$ 是自伴的. 证明：$A+\mathrm{i}B$ 是幺正的，当且仅当 $AB = BA$ 且 $A^2+B^2 = I$.

**5** 设 $S \in \mathcal{L}(V)$. 证明以下是等价的.

(a) $S$ 是自伴的幺正算子.

(b) $S = 2P - I$，其中 $P$ 是 $V$ 上的某个正交投影.

(c) 存在 $V$ 的子空间 $U$，使得 $Su = u$ 对于任一 $u \in U$ 都成立而 $Sw = -w$ 对任一 $w \in U^\perp$ 都成立.

**6** 设 $T_1, T_2$ 都是 $\mathbf{F}^3$ 上以 2、5、7 为特征值的正规算子. 证明：存在幺正算子 $S \in \mathcal{L}(\mathbf{F}^3)$ 使得 $T_1 = S^*T_2S$.

**7** 给出一例：两个自伴算子 $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$，其特征值都是 2、5、7，但是不存在一幺正算子 $S \in \mathcal{L}(\mathbf{F}^4)$ 使得 $T_1 = S^*T_2S$. 务必解释，为什么不存在满足所需性质的幺正算子.

**8** 证明或给出一反例：如果 $S \in \mathcal{L}(V)$ 且存在 $V$ 的一规范正交基 $e_1, \ldots, e_n$ 使得任一 $e_k$ 都有 $\|Se_k\| = 1$，那么 $S$ 是幺正算子.

**9** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 设 $T$ 的每个特征值的绝对值都是 1，且 $\|Tv\| \leq \|v\|$ 对任一 $v \in V$ 都成立. 证明：$T$ 是幺正算子.

**10** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ 是使得 $\|Tv\| \leq \|v\|$ 对所有 $v \in V$ 都成立的自伴算子.

(a) 证明：$I - T^2$ 是正算子.

(b) 证明：$T + \mathrm{i}\sqrt{I - T^2}$ 是幺正算子.

**11** 设 $S \in \mathcal{L}(V)$. 证明：$S$ 是幺正算子，当且仅当

$$\{Sv : v \in V \text{ 且 } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}.$$

**12** 证明或给出一反例：如果 $S \in \mathcal{L}(V)$ 可逆且 $\|S^{-1}v\| = \|Sv\|$ 对任一 $v \in V$ 都成立，那么 $S$ 是幺正的.

**13** 解释为什么：复数构成的方阵的列可组成 $\mathbf{C}^n$ 中的规范正交组，当且仅当其行可组成 $\mathbf{C}^n$ 中的规范正交组.

**14** 设 $v \in V$，$\|v\| = 1$，$b \in \mathbf{F}$. 又设 $\dim V \geq 2$. 证明：存在幺正算子 $S \in \mathcal{L}(V)$ 使得 $\langle Sv, v\rangle = b$，当且仅当 $|b| \leq 1$.

**15** 设 $T$ 是 $V$ 上的幺正算子，$T - I$ 可逆.

(a) 证明：$(T + I)(T - I)^{-1}$ 是斜算子（意即等于其伴随的负）.

(b) 证明：如果 $\mathbf{F} = \mathbf{C}$，那么 $\mathrm{i}(T + I)(T - I)^{-1}$ 是自伴算子.

注　函数 $z \mapsto \mathrm{i}(z + 1)(z - 1)^{-1}$ 将 $\mathbf{C}$ 中的单位圆（除了 1 这个点）映到 $\mathbf{R}$ 中. 因此 (b) 呈现了幺正算子和 $\mathbf{C}$ 中的单位圆之间的类比，以及自伴算子和 $\mathbf{R}$ 之间的类比.

**16** 设 $\mathbf{F} = \mathbf{C}$，且 $T \in \mathcal{L}(V)$ 是自伴的. 证明：$(T + \mathrm{i}I)(T - \mathrm{i}I)^{-1}$ 是幺正算子且 1 不是该算子的特征值.

**17** 解释为什么 7.57 给出的幺正矩阵的特性是成立的.

**18** 方阵 $A$ 被称为**对称的**（**symmetric**），如果它等于它的转置. 证明：如果 $A$ 是元素为实数的对称矩阵，那么存在元素为实数的幺正矩阵 $Q$ 使得 $Q^*AQ$ 是对角矩阵. [5]

**19** 设 $n$ 是正整数. 本题中，我们采用这一记号：将 $\mathbf{C}^n$ 的元素 $z$ 记为 $z = (z_0, z_1, \ldots, z_{n-1})$. 定义 $\mathbf{C}^n$ 上的线性泛函 $\omega_0, \omega_1, \ldots, \omega_{n-1}$ 为

$$\omega_j(z_0, z_1, \ldots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \mathrm{e}^{-2\pi \mathrm{i} jm/n}.$$

**离散傅里叶变换**（**discrete Fourier transform**）即定义为

$$\mathcal{F}z = \big(\omega_0(z), \omega_1(z), \ldots, \omega_{n-1}(z)\big)$$

的算子 $\mathcal{F} : \mathbf{C}^n \to \mathbf{C}^n$.

(a) 证明：$\mathcal{F}$ 是 $\mathbf{C}^n$ 上的幺正算子.

(b) 证明：如果 $(z_0, \ldots, z_{n-1}) \in \mathbf{C}^n$ 且规定 $z_n$ 等于 $z_0$，那么

$$\mathcal{F}^{-1}(z_0, z_1, \ldots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \ldots, z_1).$$

(c) 证明：$\mathcal{F}^4 = I$.

注　离散傅里叶变换在数据分析中有许多重要应用. 通常的傅里叶变换涉及形如

$$\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-2\pi \mathrm{i} t x}\, \mathrm{d}x$$

的表达式，其中 $f$ 是定义在 $\mathbf{R}$ 上的复值可积函数.

**20** 设 $A$ 是各列线性无关的方阵. 证明：存在唯一的矩阵 $R$ 和 $Q$ 使得 $A = RQ$，其中 $R$ 是下三角的且对角线上仅有正数，而 $Q$ 是幺正的.

[5]在其他文献中，“元素为实数的对称矩阵”常称为“实对称矩阵”；“元素为实数的幺正矩阵”常称为“正交矩阵”.

# 7E 奇异值分解

## 奇异值

本节我们将用到以下这条结果.

**7.64 $T^*T$ 的性质**

设 $T \in \mathcal{L}(V, W)$，那么

(a) $T^*T$ 是 $V$ 上的正算子；

(b) $\operatorname{null} T^*T = \operatorname{null} T$；

(c) $\operatorname{range} T^*T = \operatorname{range} T^*$；

(d) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

**证明**

(a) 因为

$$(T^*T)^* = T^*(T^*)^* = T^*T,$$

所以 $T^*T$ 是自伴的.

如果 $v \in V$，那么

$$\langle (T^*T)v, v\rangle = \langle T^*(Tv), v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0.$$

因此，$T^*T$ 是正算子.

(b) 首先设 $v \in \operatorname{null} T^*T$，那么

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = \langle 0, v\rangle = 0.$$

因此 $Tv = 0$，证明了 $\operatorname{null} T^*T \subseteq \operatorname{null} T$.

另一方向的包含关系是显然的，因为如果 $v \in V$ 且 $Tv = 0$ 就有 $T^*Tv = 0$.

因此，$\operatorname{null} T^*T = \operatorname{null} T$，这就完成了 (b) 的证明.

(c) 我们已经从 (a) 中知道了 $T^*T$ 是自伴的. 那么

$$\operatorname{range} T^*T = (\operatorname{null} T^*T)^{\perp} = (\operatorname{null} T)^{\perp} = \operatorname{range} T^*,$$

其中第一个和最后一个相等关系来自 7.6，而第二个相等关系来自 (b).

(d) 注意到

$$\dim \operatorname{range} T = \dim (\operatorname{null} T^*)^{\perp} = \dim W - \dim \operatorname{null} T^* = \dim \operatorname{range} T^*,$$

其中第一个相等关系来自 7.6 (d)，第二个来自 6.51，最后一个来自线性映射基本定理（3.21），从而验证了 (d) 中的第一个等式.

等式 $\dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$ 由 (c) 可得. ■

算子的特征值可以反映算子的一些性质. 还有一种被称为“奇异值”的数，也是很有用的. 下面提到的特征空间和记号 $E$（在示例中会用到）已在 5.52 中定义.

**7.65 定义：奇异值（singular values）**

设 $T \in \mathcal{L}(V, W)$. $T$ 的**奇异值**是 $T^*T$ 的特征值的非负平方根，按降序排列，而且每个奇异值的出现次数，等于 $T^*T$ 对应特征空间的维数.

**7.66 例：$\mathbf{F}^4$ 上一算子的奇异值**

定义 $T \in \mathcal{L}(\mathbf{F}^4)$ 为 $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$. 计算可得

$$T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4),$$

你应该自行验证这一点. 那么，我们用 $\mathbf{F}^4$ 的标准基对角化 $T^*T$，可得 $T^*T$ 的特征值是 9、4、0. 以及，对应于这些特征值的特征空间维数是

$$\dim E(9, T^*T) = 2, \qquad \dim E(4, T^*T) = 1, \qquad \dim E(0, T^*T) = 1.$$

取 $T^*T$ 的这些特征值的非负平方根，并用到以上的维数信息，我们可以得出结论：$T$ 的奇异值是 3、3、2、0.

$T$ 仅有的特征值是 $-3$ 和 0. 所以在这个例子中，特征值并没有包含出现在 $T$ 的定义中的 2，而奇异值中却包含了 2.

**7.67 例：从 $\mathbf{F}^4$ 到 $\mathbf{F}^3$ 的一线性映射的奇异值**

设 $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$（关于标准基）的矩阵是

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

你可以验证，$T^*T$ 的矩阵是

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix},$$

特征值是 25、2、0；而 $\dim E(25, T^*T) = 1$，$\dim E(2, T^*T) = 1$，$\dim E(0, T^*T) = 2$. 所以，$T$ 的奇异值是 5、$\sqrt{2}$、0、0.

本节习题 2 给出了正奇异值的一种刻画.

**7.68 正奇异值的作用**

设 $T \in \mathcal{L}(V, W)$，那么

(a) $T$ 是单射 $\Longleftrightarrow$ 0 不是 $T$ 的奇异值；

(b) $T$ 的正奇异值个数等于 $\dim \operatorname{range} T$；

(c) $T$ 是满射 $\Longleftrightarrow$ $T$ 的正奇异值个数等于 $\dim W$.

**证明** 线性映射 $T$ 是单射当且仅当 $\operatorname{null} T = \{0\}$，当且仅当 $\operatorname{null} T^*T = \{0\}$【根据 7.64 (b)】，当且仅当 0 不是 $T^*T$ 的特征值，当且仅当 0 不是 $T$ 的奇异值，这就完成了 (a) 的证明.

将谱定理应用于 $T^*T$ 可得 $\dim\operatorname{range} T^*T$ 等于 $T^*T$ 的正特征值个数（考虑重复）. 那么，由 7.64 (d) 可得 $\dim\operatorname{range} T$ 等于 $T$ 的正奇异值个数，证明了 (b).

利用 (b) 和 2.39 可证明 (c) 成立. ■

以下表格对比了特征值和奇异值.

| **由特征值构成的组** | **由奇异值构成的组** |
|---|---|
| 背景：向量空间 | 背景：内积空间 |
| 仅对从向量空间到自身的线性映射有定义 | 对于从内积空间到可能不同的内积空间的线性映射都有定义 |
| 组中元素可以是任意实数（若 $\mathbf{F}=\mathbf{R}$）或复数（若 $\mathbf{F}=\mathbf{C}$） | 组中元素必为非负数 |
| 如果 $\mathbf{F}=\mathbf{R}$，可以为空组 | 组的长度等于定义空间维数 |
| 包含 $0 \iff$ 算子不可逆 | 包含 $0 \iff$ 线性映射不是单射 |
| 没有标准顺序，特别是当 $\mathbf{F}=\mathbf{C}$ 时 | 总是降序排列 |

接下来这条结果，用奇异值很好地刻画了等距映射.

**7.69 所有奇异值都等于 1 是等距映射的特征**

设 $S \in \mathcal{L}(V,W)$，那么

$$S \text{ 是等距映射} \iff S \text{ 的所有奇异值都等于 } 1.$$

**证明**

$$\begin{aligned} S \text{ 是等距映射} &\iff S^*S = I \\ &\iff S^*S \text{ 的所有特征值都等于 } 1 \\ &\iff S \text{ 的所有奇异值都等于 } 1. \end{aligned}$$

其中，第一个等价关系来自 7.49，第二个由谱定理（7.29 或 7.31）应用于自伴算子 $S^*S$ 可得. ■

## 线性映射和矩阵的奇异值分解

接下来这条结果表明，对于从 $V$ 到 $W$ 的每个线性映射，都可以用它的奇异值以及 $V$ 和 $W$ 中的规范正交组，给出非常干净利落的描述. 在下一节，我们将看到奇异值分解（也经常称作 SVD）的一些重要应用.

奇异值分解在数值线性代数中很有用，因为有很好的工具可用于逼近正算子的特征值和特征向量，而正算子 $T^*T$ 的特征值和特征向量又能引出奇异值分解.

**7.70 奇异值分解（singular value decomposition）**

设 $T \in \mathcal{L}(V,W)$ 且 $T$ 的正奇异值是 $s_1,\ldots,s_m$. 那么存在 $V$ 中的规范正交组 $e_1,\ldots,e_m$ 和 $W$ 中的规范正交组 $f_1,\ldots,f_m$ 使得

$$Tv = s_1\langle v,e_1\rangle f_1 + \cdots + s_m\langle v,e_m\rangle f_m \tag{7.71}$$

对任一 $v \in V$ 都成立.

**证明** 令 $s_1,\ldots,s_n$ 表示 $T$ 的奇异值（则 $n=\dim V$）. 因为 $T^*T$ 是正算子【见 7.64 (a)】, 由谱定理可得, 存在 $V$ 中的规范正交基 $e_1,\ldots,e_n$ 使得

$$T^*Te_k = s_k^2 e_k \tag{7.72}$$

对任一 $k=1,\ldots,n$ 都成立.

对于任一 $k=1,\ldots,m$, 令

$$f_k = \frac{Te_k}{s_k}. \tag{7.73}$$

如果 $j,k\in\{1,\ldots,m\}$, 那么

$$\langle f_j,f_k\rangle = \frac{1}{s_js_k}\langle Te_j,Te_k\rangle = \frac{1}{s_js_k}\langle e_j,T^*Te_k\rangle = \frac{s_k}{s_j}\langle e_j,e_k\rangle = \begin{cases}0 & 若\ j\neq k,\\ 1 & 若\ j=k.\end{cases}$$

那么, $f_1,\ldots,f_m$ 是 $W$ 中的规范正交组.

如果 $k\in\{1,\ldots,n\}$ 且 $k>m$, 那么 $s_k=0$, 从而 $T^*Te_k=0$（根据 7.72）, 也就意味着 $Te_k=0$【根据 7.64 (b)】.

设 $v\in V$, 那么

$$\begin{aligned}Tv &= T\big(\langle v,e_1\rangle e_1+\cdots+\langle v,e_n\rangle e_n\big)\\ &= \langle v,e_1\rangle Te_1+\cdots+\langle v,e_m\rangle Te_m\\ &= s_1\langle v,e_1\rangle f_1+\cdots+s_m\langle v,e_m\rangle f_m.\end{aligned}$$

其中, 最后一个下标从第一行的 $n$ 变为第二行的 $m$ 是因为如果 $k>m$ 就有 $Te_k=0$（见上一段）, 而第三行是根据式 (7.73) 得到的. 上式就是我们预期的结果. ■

设 $T\in\mathcal{L}(V,W)$, $T$ 的正奇异值是 $s_1,\ldots,s_m$, 而 $e_1,\ldots,e_m$ 和 $f_1,\ldots,f_m$ 是奇异值分解（7.70）中的那两组向量. 规范正交组 $e_1,\ldots,e_m$ 可以扩充为 $V$ 的规范正交基 $e_1,\ldots,e_{\dim V}$, 规范正交组 $f_1,\ldots,f_m$ 可以扩充为 $W$ 的规范正交基 $f_1,\ldots,f_{\dim W}$. 公式 (7.71) 表明

$$Te_k = \begin{cases} s_kf_k & 若\ 1\le k\le m,\\ 0 & 若\ m<k\le\dim V.\end{cases}$$

那么 $T$ 关于规范正交基 $(e_1,\ldots,e_{\dim V})$ 和 $(f_1,\ldots,f_{\dim W})$ 的矩阵具有如下简单形式:

$$\mathcal{M}\big(T,(e_1,\ldots,e_{\dim V}),(f_1,\ldots,f_{\dim W})\big)_{j,k} = \begin{cases} s_k & 若\ 1\le j=k\le m,\\ 0 & 其他.\end{cases}$$

如果 $\dim V=\dim W$（比如, 如果 $W=V$ 就是这样）, 那么上一段描述的矩阵就是对角矩阵. 如果我们像下面这样推广对角矩阵的定义, 使其适用于那些不一定是方阵的矩阵, 那么我们就证明了一个极好的结果——每个从 $V$ 到 $W$ 的线性映射关于恰当的规范正交基都有对角矩阵.

**7.74 定义：对角矩阵（diagonal matrix）**

$M\times N$ 矩阵 $A$ 被称为**对角矩阵**, 如果除了 $A_{k,k}$（$k=1,\ldots,\min\{M,N\}$）可能不为 0 以外, 所有元素都为 0. ♣

以下表格对比了谱定理（7.29 和 7.31）和奇异值分解（7.70）.

| **谱定理** | **奇异值分解** |
|---|---|
| 只描述自伴算子（$\mathbf{F} = \mathbf{R}$ 的情形下）或正规算子（$\mathbf{F} = \mathbf{C}$ 的情形下） | 描述任意从内积空间到可能不同的内积空间的线性映射 |
| 得到一个规范正交基 | 得到两个规范正交组，一个是定义空间的，一个是值域的，而且即使在值域等于定义空间的情况下也不一定是同一个 |
| 取决于是 $\mathbf{F} = \mathbf{R}$ 还是 $\mathbf{F} = \mathbf{C}$ 而有不同的证明 | 不管是 $\mathbf{F} = \mathbf{R}$ 还是 $\mathbf{F} = \mathbf{C}$，证明都是一样的 |

奇异值分解让我们能以一种新的方式来理解线性映射的伴随和逆. 具体而言，接下来的结果表明，给出线性映射 $T \in \mathcal{L}(V, W)$ 的奇异值分解之后，只需交换各 $e$ 和各 $f$ 的角色【见式 (7.77)】，我们就可以得到 $T$ 的伴随；类似地，通过交换各 $e$ 和各 $f$ 的角色再把 $T$ 的每个正奇异值 $s_k$ 替换成 $1/s_k$【见式 (7.78)】，我们就可以得到 $T$ 的伪逆 $T^\dagger$（见 6.68）.

回想一下，如果 $T$ 可逆，那么以下的式 (7.78) 中的伪逆 $T^\dagger$ 就等于逆 $T^{-1}$【见 6.69 (a)】.

**7.75 伴随和伪逆的奇异值分解**

设 $T \in \mathcal{L}(V, W)$ 且 $T$ 的正奇异值是 $s_1, \ldots, s_m$. 设 $e_1, \ldots, e_m$ 和 $f_1, \ldots, f_m$ 是 $V$ 和 $W$ 中的规范正交组，使得对任一 $v \in V$ 都有

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m. \tag{7.76}$$

那么，对任一 $w \in W$ 都有

$$T^* w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m \tag{7.77}$$

和

$$T^\dagger w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m. \tag{7.78}$$

♡

**证明** 如果 $v \in V$ 且 $w \in W$，那么

$$\begin{aligned}
\langle Tv, w \rangle &= \bigl\langle s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m, w \bigr\rangle \\
&= s_1 \langle v, e_1 \rangle \langle f_1, w \rangle + \cdots + s_m \langle v, e_m \rangle \langle f_m, w \rangle \\
&= \bigl\langle v, s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m \bigr\rangle.
\end{aligned}$$

这意味着

$$T^* w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m,$$

证明了式 (7.77).

为证式 (7.78)，设 $w \in W$. 令

$$v = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m.$$

将 $T$ 作用于上式两侧，得到

$$\begin{aligned}
Tv &= \frac{\langle w, f_1 \rangle}{s_1} T e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} T e_m \\
&= \langle w, f_1 \rangle f_1 + \cdots + \langle w, f_m \rangle f_m \\
&= P_{\operatorname{range} T} w.
\end{aligned}$$

其中，第二行成立是因为式 (7.76) 蕴涵了若 $k = 1, \ldots, m$ 则 $Te_k = s_k f_k$，最后一行成立是因为式 (7.76) 蕴涵了 $f_1, \ldots, f_m$ 张成 $\operatorname{range} T$ 从而是 $\operatorname{range} T$ 的规范正交基【这样就可以用到 6.57 (i) 了】. 有了上式，并且观察到 $v \in (\operatorname{null} T)^\perp$【见本节习题 8 (b)】，再结合 $T^\dagger w$ 的定义（见 6.68），可得 $v = T^\dagger w$，证明了式 (7.78).

**7.79 例：求奇异值分解**

定义 $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ 为 $T(x_1, x_2, x_3, x_4) = (-5x_4, 0, x_1 + x_2)$. 我们想求 $T$ 的奇异值分解. $T$（关于标准基）的矩阵是

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

那么，正如例 7.67 讨论过的那样，$T^*T$ 的矩阵是

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix},$$

正特征值是 25 和 2；而 $\dim E(25, T^*T) = 1$，$\dim E(2, T^*T) = 1$. 所以，$T$ 的正奇异值是 5 和 $\sqrt{2}$.

所以，为求出 $T$ 的奇异值分解，我们必须求出 $\mathbf{F}^4$ 中的一规范正交组 $e_1, e_2$ 和 $\mathbf{F}^3$ 中的一规范正交组 $f_1, f_2$ 使得

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

对所有 $v \in \mathbf{F}^4$ 成立.

$E(25, T^*T)$ 的一规范正交基是向量 $(0, 0, 0, 1)$，$E(2, T^*T)$ 的一规范正交基是向量 $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)$. 所以，根据 7.70 的证明，我们可以取

$$e_1 = (0, 0, 0, 1) \quad 和 \quad e_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$$

以及

$$f_1 = \frac{Te_1}{5} = (-1, 0, 0) \quad 和 \quad f_2 = \frac{Te_2}{\sqrt{2}} = (0, 0, 1).$$

然后，不出所料，我们可以看到 $e_1, e_2$ 就是 $\mathbf{F}^4$ 中的规范正交组，$f_1, f_2$ 就是 $\mathbf{F}^3$ 中的规范正交组，而且

$$Tv = 5\langle v, e_1\rangle f_1 + \sqrt{2}\langle v, e_2\rangle f_2$$

对所有 $v \in \mathbf{F}^4$ 成立. 那么，我们就求出了 $T$ 的奇异值分解.

接下来的结果，将奇异值分解从线性映射背景下迁移到矩阵背景下. 具体而言，它给出了对任意矩阵的分解，将其化为三个性质良好的矩阵的乘积. 这里的证明过程利用线性映射奇异值分解的语言给出了这三个矩阵的具体构造方法.

接下来这条结果中，“列规范正交”一词应该解释为，各列关于标准欧几里得内积是规范正交的.

**7.80 奇异值分解（SVD）的矩阵版本**

设 $A$ 是 $p \times n$ 矩阵（秩 $m \geq 1$）. 那么，存在列规范正交的 $p \times m$ 矩阵 $B$，对角线上为正数的 $m \times m$ 对角矩阵 $D$ 以及列规范正交的 $n \times m$ 矩阵 $C$ 使得

$$A = BDC^*.$$

**证明** 令线性映射 $T : \mathbf{F}^n \to \mathbf{F}^p$ 关于标准基的矩阵等于 $A$. 那么，$\dim \operatorname{range} T = m$（根据 3.78）. 令

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \tag{7.81}$$

是 $T$ 的奇异值分解. 令

$$\begin{aligned} B &= \text{各列为 } f_1, \ldots, f_m \text{ 的 } p \times m \text{ 矩阵,} \\ D &= \text{对角线元素为 } s_1, \ldots, s_m \text{ 的 } m \times m \text{ 对角矩阵,} \\ C &= \text{各列为 } e_1, \ldots, e_m \text{ 的 } n \times m \text{ 矩阵.} \end{aligned}$$

令 $u_1, \ldots, u_m$ 表示 $\mathbf{F}^m$ 的标准基. 如果 $k \in \{1, \ldots, m\}$，那么

$$(AC - BD)u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0.$$

所以，$AC = BD$.

在前式两侧右乘 $C^*$（$C$ 的共轭转置），得到

$$ACC^* = BDC^*.$$

注意到 $C^*$ 的行是 $e_1, \ldots, e_m$ 的复共轭. 因此，如果 $k \in \{1, \ldots, m\}$，那么由矩阵乘法的定义可得 $C^* e_k = u_k$，于是 $CC^* e_k = e_k$. 因此，$ACC^* v = Av$ 对所有 $v \in \operatorname{span}(e_1, \ldots, e_m)$ 成立.

如果 $v \in \left(\operatorname{span}(e_1, \ldots, e_m)\right)^{\perp}$，那么（由 7.81 可得）$Av = 0$，（由矩阵乘法的定义可得）$C^* v = 0$. 于是，$ACC^* v = Av$ 对所有 $v \in \left(\operatorname{span}(e_1, \ldots, e_m)\right)^{\perp}$ 成立.

因为 $ACC^*$ 和 $A$ 在 $\operatorname{span}(e_1, \ldots, e_m)$ 和 $\left(\operatorname{span}(e_1, \ldots, e_m)\right)^{\perp}$ 上都是一致的，我们可以得出结论 $ACC^* = A$. 因此前面那个行间公式就变成了

$$A = BDC^*,$$

命题得证. ■

注意，以上结果中的矩阵 $A$ 有 $pn$ 个元素. 相比之下，以上的矩阵 $B, D, C$ 总共有

$$m(p + m + n)$$

个元素. 因此，如果 $p$ 和 $n$ 是很大的数而秩 $m$ 比 $p$ 和 $n$ 小得多，那么为了表示 $A$ 而必须存储在计算机中的元素个数就比 $pn$ 小得多.

## 习题 7E

**1** 设 $T \in \mathcal{L}(V,W)$. 证明：$T=0$，当且仅当 $T$ 的所有奇异值都为 0.

**2** 设 $T \in \mathcal{L}(V,W)$，$s>0$. 证明：$s$ 是 $T$ 的奇异值，当且仅当存在非零向量 $v \in V$ 和非零向量 $w \in W$，使得

$$Tv = sw \quad \text{且} \quad T^*w = sv.$$

注 满足以上两式的向量 $v, w$ 被称为**施密特对（Schmidt pair）**. 埃哈德·施密特在 1907 年引入了奇异值的概念.

**3** 给出一例：$T \in \mathcal{L}(\mathbf{C}^2)$，0 是 $T$ 的唯一特征值，而 $T$ 的奇异值是 5 和 0.

**4** 设 $T \in \mathcal{L}(V,W)$，$s_1$ 是 $T$ 的最大奇异值，$s_n$ 是 $T$ 的最小奇异值. 证明：

$$\{\|Tv\| : v \in V \text{ 且 } \|v\|=1\} = [s_n, s_1].$$

**5** 设 $T \in \mathcal{L}(\mathbf{C}^2)$ 定义为 $T(x,y)=(-4y,x)$. 求 $T$ 的奇异值.

**6** 求定义为 $Dp=p'$ 的微分算子 $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ 的奇异值. 其中 $\mathcal{P}_2(\mathbf{R})$ 上的内积如例 6.34 所示.

**7** 设 $T \in \mathcal{L}(V)$ 是自伴的，或者设 $\mathbf{F}=\mathbf{C}$ 且 $T \in \mathcal{L}(V)$ 是正规的. 令 $\lambda_1, \ldots, \lambda_n$ 是 $T$ 的特征值，每个特征值的出现次数等于其对应特征空间的维数. 证明：$T$ 的奇异值是按降序排列后的 $|\lambda_1|, \ldots, |\lambda_n|$.

**8** 设 $T \in \mathcal{L}(V,W)$. 设 $s_1 \geq s_2 \geq \cdots \geq s_m > 0$，$e_1, \ldots, e_m$ 是 $V$ 中的规范正交组，$f_1, \ldots, f_m$ 是 $W$ 中的规范正交组，使得

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

对任一 $v \in V$ 都成立.

(a) 证明：$f_1, \ldots, f_m$ 是 $\operatorname{range} T$ 的规范正交基.

(b) 证明：$e_1, \ldots, e_m$ 是 $(\operatorname{null} T)^{\perp}$ 的规范正交基.

(c) 证明：$s_1, \ldots, s_m$ 是 $T$ 的正奇异值.

(d) 证明：如果 $k \in \{1, \ldots, m\}$，那么 $e_k$ 是 $T^*T$ 的特征向量，对应特征值是 $s_k^2$.

(e) 证明：

$$TT^*w = s_1^2 \langle w, f_1 \rangle f_1 + \cdots + s_m^2 \langle w, f_m \rangle f_m$$

对所有 $w \in W$ 成立.

**9** 设 $T \in \mathcal{L}(V,W)$. 证明：$T$ 和 $T^*$ 的正奇异值相同.

**10** 设 $T \in \mathcal{L}(V,W)$ 的奇异值是 $s_1, \ldots, s_n$. 证明：如果 $T$ 是可逆线性映射，那么 $T^{-1}$ 的奇异值是

$$\frac{1}{s_n}, \ldots, \frac{1}{s_1}.$$

**11** 设 $T \in \mathcal{L}(V,W)$，$v_1, \ldots, v_n$ 是 $V$ 的规范正交基. 令 $s_1, \ldots, s_n$ 表示 $T$ 的奇异值.

(a) 证明：$\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.

(b) 证明：如果 $W=V$ 且 $T$ 是正算子，那么

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

注 见 7A 节习题 5 的小注部分.

**12** (a) 给出一例：一个有限维向量空间，其上一算子 $T$，使得 $T^2$ 的奇异值不等于 $T$ 的奇异值的平方.

(b) 设 $T \in \mathcal{L}(V)$ 是正规的. 证明：$T^2$ 的奇异值等于 $T$ 的奇异值的平方.

**13** 设 $T_1, T_2 \in \mathcal{L}(V)$. 证明：$T_1$ 和 $T_2$ 的奇异值相同，当且仅当存在幺正算子 $S_1, S_2 \in \mathcal{L}(V)$ 使得 $T_1 = S_1 T_2 S_2$.

**14** 设 $T \in \mathcal{L}(V, W)$. 令 $s_n$ 表示 $T$ 的最小奇异值. 证明：$s_n \|v\| \leq \|Tv\|$ 对任一 $v \in V$ 都成立.

**15** 设 $T \in \mathcal{L}(V)$，$s_1 \geq \cdots \geq s_n$ 是 $T$ 的奇异值. 证明：如果 $\lambda$ 是 $T$ 的特征值，那么 $s_1 \geq |\lambda| \geq s_n$.

**16** 设 $T \in \mathcal{L}(V, W)$. 证明：$(T^*)^\dagger = (T^\dagger)^*$.

注 将本题结果与可逆线性映射的类似结果【见 7.5 (f)】比较一下.

**17** 设 $T \in \mathcal{L}(V)$. 证明：$T$ 是自伴的，当且仅当 $T^\dagger$ 是自伴的.

| | |
|---|---|
| Matrices unfold | 矩阵藏奥秘， |
| Singular values gleam like stars | SVD 解玄机. |
| Order in chaos shines | 左右奇异向， |
| | 数据显真迹. |

左：由 ChatGPT 所作，输入提示词为“haiku about SVD”（主题为 SVD 的俳句）

右：由文心一言所作，输入提示词为“就‘奇异值分解’写一首五言绝句”

# 7F 奇异值分解的推论

## 线性映射的范数

奇异值分解引出了接下来的 $\|Tv\|$ 的上界.

**7.82 $\|Tv\|$ 的上界**

设 $T \in \mathcal{L}(V, W)$. 令 $s_1$ 是 $T$ 的最大奇异值. 那么

$$\|Tv\| \leq s_1 \|v\|$$

对所有 $v \in V$ 成立.

$\|Tv\|$ 的一个下界, 见 7E 节的习题 14.

**证明** 令 $s_1, \ldots, s_m$ 表示 $T$ 的正奇异值, 令 $e_1, \ldots, e_m$ 是 $V$ 中的规范正交组, $f_1, \ldots, f_m$ 是 $W$ 中的规范正交组, 且由此可给出 $T$ 的奇异值分解. 那么

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \tag{7.83}$$

对所有 $v \in V$ 成立. 于是, 若 $v \in V$ 则有

$$\begin{aligned}\|Tv\|^2 &= s_1^2 |\langle v, e_1 \rangle|^2 + \cdots + s_m^2 |\langle v, e_m \rangle|^2 \\ &\leq s_1^2 \left( |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \right) \\ &\leq s_1^2 \|v\|^2.\end{aligned}$$

其中最后一个不等关系由贝塞尔不等式(6.26)得到.不等式两侧取平方根,可得 $\|Tv\| \leq s_1 \|v\|$,命题得证. ■

设 $T \in \mathcal{L}(V, W)$ 且 $s_1$ 是 $T$ 的最大奇异值. 以上结果表明

$$\|Tv\| \leq s_1 \text{ 对所有满足 } \|v\| \leq 1 \text{ 的 } v \in V \text{ 成立.} \tag{7.84}$$

在式 (7.83) 中, 取 $v = e_1$, 可得 $Te_1 = s_1 f_1$. 因为 $\|f_1\| = 1$, 所以这意味着 $\|Te_1\| = s_1$. 那么, 由于 $\|e_1\| = 1$, 7.84 中的不等式可以引出等式

$$\max \left\{ \|Tv\| : v \in V \text{ 且 } \|v\| \leq 1 \right\} = s_1. \tag{7.85}$$

上式促使我们直接将 $T$ 的范数定义为上式的左侧, 而不需要提到(右侧的)奇异值或者奇异值分解, 如下所示.

**7.86 定义: 线性映射的范数 (norm of a linear map)、$\|\cdot\|$**

设 $T \in \mathcal{L}(V, W)$. 那么, $T$ 的**范数**, 记为 $\|T\|$, 定义为

$$\|T\| = \max \left\{ \|Tv\| : v \in V \text{ 且 } \|v\| \leq 1 \right\}.$$

一般地, 由非负数构成的无限集不一定存在最大值. 然而, 7.86 前的讨论表明, 从 $V$ 到 $W$ 的线性映射 $T$ 的范数定义中的最大值确实存在(且等于 $T$ 的最大奇异值).

关于**"范数"**一词和 $\|\cdot\|$ 这一记号, 我们现在有两种不同用法: 第一种用法跟 $V$ 上的内积相关, 我们定义的是 $\|v\| = \sqrt{\langle v, v \rangle}$ 对任一 $v \in V$ 都成立; 第二种用法, 则是用我们刚刚为

$T \in \mathcal{L}(V,W)$ 所作的 $\|T\|$ 的定义. $T \in \mathcal{L}(V,W)$ 的范数 $\|T\|$ 通常不是通过对 $T$ 和其本身取内积得到的（见习题 21）. 从上下文和所用的符号中，你应该能够判断出其中的范数是何种含义.

$\mathcal{L}(V,W)$ 上的范数性质列出如下，和内积空间上的范数性质（见 6.9 和 6.17）看起来是一样的. (d) 中的不等式称为**三角不等式**，这就和我们对 $V$ 上的范数用的术语是一样的. 至于反向三角不等式，见本节习题 1.

**7.87 线性映射范数的基本性质**

设 $T \in \mathcal{L}(V,W)$，那么

(a) $\|T\| \geq 0$；

(b) $\|T\| = 0 \iff T = 0$；

(c) $\|\lambda T\| = |\lambda|\,\|T\|$ 对所有 $\lambda \in \mathbf{F}$ 成立；

(d) $\|S+T\| \leq \|S\| + \|T\|$ 对所有 $S \in \mathcal{L}(V,W)$ 成立.

**证明**

(a) 因为 $\|Tv\| \geq 0$ 对任一 $v \in V$ 都成立，所以由 $T$ 的定义可得 $\|T\| \geq 0$.

(b) 设 $\|T\| = 0$. 那么 $Tv = 0$ 对所有满足 $\|v\| \leq 1$ 的 $v \in V$ 成立. 如果 $u \in V$（$u \neq 0$），那么

$$Tu = \|u\|\, T\left(\frac{u}{\|u\|}\right) = 0.$$

其中最后一个相等关系成立是因为 $u/\|u\|$ 范数为 1. 因为 $Tu = 0$ 对所有 $u \in V$ 成立，所以我们有 $T = 0$.

反之，如果 $T = 0$，那么 $Tv = 0$ 对所有 $v \in V$ 成立，于是 $\|T\| = 0$.

(c) 设 $\lambda \in \mathbf{F}$，那么

$$\begin{aligned}\|\lambda T\| &= \max\left\{\|\lambda Tv\| : v \in V \text{ 且 } \|v\| \leq 1\right\}\\ &= |\lambda| \max\left\{\|Tv\| : v \in V \text{ 且 } \|v\| \leq 1\right\}\\ &= |\lambda|\,\|T\|.\end{aligned}$$

(d) 设 $S \in \mathcal{L}(V,W)$. 由 $\|S+T\|$ 的定义可知，存在 $v \in V$ 使得 $\|v\| \leq 1$ 且 $\|S+T\| = \|(S+T)v\|$. 这样一来，

$$\|S+T\| = \|(S+T)v\| = \|Sv+Tv\| \leq \|Sv\| + \|Tv\| \leq \|S\| + \|T\|,$$

完成了 (d) 的证明. ■

对于 $S,T \in \mathcal{L}(V,W)$，$\|S-T\|$ 这个量常常被称为 $S$ 和 $T$ 之间的距离. 不正式地说，$\|S-T\|$ 这个数比较小就意味着 $S$ 和 $T$ 比较接近. 例如本节习题 9 中指出，对于任一 $T \in \mathcal{L}(V)$，总可以找到与 $T$ 任意接近的可逆算子.

**7.88 $\|T\|$ 的多种表达式**

设 $T \in \mathcal{L}(V,W)$，那么：

(a) $\|T\| = T$ 的最大奇异值；

(b) $\|T\| = \max\left\{\|Tv\| : v \in V \text{ 且 } \|v\| = 1\right\}$；

(c) $\|T\| =$ 使得 $\|Tv\| \leq c\,\|v\|$ 对所有 $v \in V$ 成立的最小数 $c$.

**证明**

(a) 见 7.85.

(b) 令 $v \in V$，$0 < \|v\| \le 1$. 令 $u = v/\|v\|$. 那么有

$$\|u\| = \left\|\frac{v}{\|v\|}\right\| = 1 \quad \text{和} \quad \|Tu\| = \left\|T\left(\frac{v}{\|v\|}\right)\right\| = \frac{\|Tv\|}{\|v\|} \ge \|Tv\|.$$

因此，求 $\|Tv\|$ 在 $\|v\| \le 1$ 条件下的最大值时，我们可以只关注 $V$ 中范数为 1 的那些向量，也就是证明了 (b).

(c) 设 $v \in V$ 且 $v \ne 0$. 那么由 $\|T\|$ 的定义可得

$$\left\|T\left(\frac{v}{\|v\|}\right)\right\| \le \|T\|,$$

从而可得

$$\|Tv\| \le \|T\|\,\|v\|. \tag{7.89}$$

现在设 $c \ge 0$ 且 $\|Tv\| \le c\|v\|$ 对所有 $v \in V$ 成立. 这意味着，

$$\|Tv\| \le c$$

对所有满足 $\|v\| \le 1$ 的 $v \in V$ 成立. 遍历所有满足 $\|v\| \le 1$ 的 $v \in V$，取得的以上不等式左侧最大值 $\|T\| \le c$. 因此，$\|T\|$ 是使得 $\|Tv\| \le c\|v\|$ 对所有 $v \in V$ 成立的最小数 $c$. ■

在处理线性映射的范数时，你可能会经常用到不等式 (7.89).

给定一线性映射 $T$ 关于某个规范正交基的矩阵，要计算 $T$ 的范数近似值，7.88 (a) 可能是最有用的. 用矩阵乘法可以很快算出 $T^*T$ 的矩阵，用计算机可以求出 $T^*T$ 的最大特征值的近似值（为此已有极好的数值算法），然后开平方根，再用 7.88 (a) 便可得到 $T$ 的范数的近似值（这通常无法确切计算）.

你应该自行验证下面例子中的所有结论.

---

**7.90 例：范数**

- 如果 $I$ 表示 $V$ 上通常的恒等算子，那么 $\|I\| = 1$.
- 如果 $T \in \mathcal{L}(\mathbf{F}^n)$ 且 $T$ 关于 $\mathbf{F}^n$ 的标准基的矩阵各元素全为 1，那么 $\|T\| = n$.
- 如果 $T \in \mathcal{L}(V)$ 且 $V$ 有一规范正交基由 $T$ 的对应于特征值 $\lambda_1, \ldots, \lambda_n$ 的特征向量组成，那么 $\|T\|$ 是 $|\lambda_1|, \ldots, |\lambda_n|$ 中的最大值.
- 设算子 $T \in \mathcal{L}(\mathbf{R}^5)$（关于标准基）的 $5 \times 5$ 矩阵第 $j$ 行第 $k$ 列的元素是 $1/(j^2 + k)$. 通用的数学软件显示，$T$ 的最大奇异值约等于 0.8，而最小的约等于 $10^{-6}$. 因此，$\|T\| \approx 0.8$ 以及（用 7E 节的习题 10 所述结论可得）$\|T^{-1}\| \approx 10^6$. 而求出这些范数的确切表达式是不可能的.

---

接下来这条结论表明，线性映射和它的伴随范数相同.

**7.91 伴随的范数**

设 $T \in \mathcal{L}(V, W)$，那么 $\|T^*\| = \|T\|$.

**证明** 设 $w \in W$，那么

$$\|T^*w\|^2 = \langle T^*w, T^*w\rangle = \langle TT^*w, w\rangle \le \|TT^*w\|\,\|w\| \le \|T\|\,\|T^*w\|\,\|w\|.$$

以上不等式意味着

$$\|T^*w\| \le \|T\|\,\|w\|,$$

结合 7.88 (c) 可得 $\|T^*\| \le \|T\|$.

将 $\|T^*\| \le \|T\|$ 中的 $T$ 替换成 $T^*$，再用上 $(T^*)^* = T$，可得 $\|T\| \le \|T^*\|$. 因此，$\|T^*\| = \|T\|$，命题得证. ■

你可能还想用 7E 节的习题 9 构造以上结果的另一证明. 那道习题指出，线性映射和它的伴随的正奇异值相同.

## 用具有较低维值域的线性映射进行逼近

接下来的结果是奇异值分解的精彩应用. 它说的是，削去奇异值分解中前 $k$ 项之后的项，从而用值域维数至多为 $k$ 的线性映射，得到线性映射的最佳逼近. 具体而言，接下来这条结果中，线性映射 $T_k$ 满足这样的性质——$\dim \operatorname{range} T_k = k$，并且在值域维数至多为 $k$ 的所有线性映射中，$T_k$ 到 $T$ 的距离最小. 这条结果引出了用以压缩巨大矩阵并同时保留其重要信息的算法.

**7.92 用值域维数至多为 $k$ 的线性映射得到的最佳逼近**

设 $T \in \mathcal{L}(V, W)$ 且 $s_1 \ge \cdots \ge s_m$ 是 $T$ 的正奇异值. 设 $1 \le k < m$，那么有

$$\min\left\{\|T - S\| : S \in \mathcal{L}(V, W) \text{ 且 } \dim \operatorname{range} S \le k\right\} = s_{k+1}.$$

进一步，如果

$$Tv = s_1 \langle v, e_1\rangle f_1 + \cdots + s_m \langle v, e_m\rangle f_m$$

是 $T$ 的奇异值分解，而 $T_k \in \mathcal{L}(V, W)$ 定义为

$$T_k v = s_1 \langle v, e_1\rangle f_1 + \cdots + s_k \langle v, e_k\rangle f_k$$

对任一 $v \in V$ 都成立，那么 $\dim \operatorname{range} T_k = k$ 且 $\|T - T_k\| = s_{k+1}$.

**证明** 如果 $v \in V$，那么

$$\begin{aligned}
\|(T - T_k)v\|^2 &= \|s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m \langle v, e_m\rangle f_m\|^2 \\
&= s_{k+1}^2 |\langle v, e_{k+1}\rangle|^2 + \cdots + s_m^2 |\langle v, e_m\rangle|^2 \\
&\le s_{k+1}^2 \left(|\langle v, e_{k+1}\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\right) \\
&\le s_{k+1}^2 \|v\|^2.
\end{aligned}$$

因此，$\|T - T_k\| \le s_{k+1}$. 又由 $(T - T_k)e_{k+1} = s_{k+1} f_{k+1}$，可得 $\|T - T_k\| = s_{k+1}$.

设 $S \in \mathcal{L}(V, W)$ 且 $\dim \operatorname{range} S \le k$. 则长度为 $k+1$ 的组 $Se_1, \ldots, Se_{k+1}$ 线性相关. 因此，存在不全为 0 的 $a_1, \ldots, a_{k+1} \in \mathbf{F}$ 使得

$$a_1 S e_1 + \cdots + a_{k+1} S e_{k+1} = 0.$$

这样一来，因为 $a_1,\ldots,a_{k+1}$ 不全为 0，所以 $a_1e_1+\cdots+a_{k+1}e_{k+1}\neq 0$. 我们有

$$\begin{aligned}\|(T-S)(a_1e_1+\cdots+a_{k+1}e_{k+1})\|^2 &= \|T(a_1e_1+\cdots+a_{k+1}e_{k+1})\|^2\\ &= \|s_1a_1f_1+\cdots+s_{k+1}a_{k+1}f_{k+1}\|^2\\ &= s_1^2|a_1|^2+\cdots+s_{k+1}^2|a_{k+1}|^2\\ &\geq s_{k+1}^2\left(|a_1|^2+\cdots+|a_{k+1}|^2\right)\\ &= s_{k+1}^2\|a_1e_1+\cdots+a_{k+1}e_{k+1}\|^2.\end{aligned}$$

因为 $a_1e_1+\cdots+a_{k+1}e_{k+1}\neq 0$，所以以上不等式就意味着

$$\|T-S\|\geq s_{k+1}.$$

因此，在 $\dim\operatorname{range}S\leq k$ 的 $S\in\mathcal{L}(V,W)$ 当中，$S=T_k$ 使得 $\|T-S\|$ 最小. ■

奇异值分解用于最佳逼近的其他例子，见本节习题 22、27. 前一题是求给定维数的子空间，使得限制在其上的线性映射尽可能小；而后一题是求一幺正算子，使其跟给定的算子尽可能近.

## 极分解

回顾一下，我们在 7.54 前讨论了满足 $|z|=1$ 的复数 $z$ 和幺正算子之间的类比. 我们继续用这个类比，并注意到除 0 以外每个复数 $z$ 都可以写成这样的形式：

$$\begin{aligned}z &= \left(\frac{z}{|z|}\right)|z|\\ &= \left(\frac{z}{|z|}\right)\sqrt{\bar{z}z},\end{aligned}$$

其中第一个因子，也就是 $z/|z|$，绝对值为 1.

根据上段，我们这样猜测：每个算子 $T\in\mathcal{L}(V)$ 都可以写成一幺正算子乘以 $\sqrt{T^*T}$ 的形式. 这样猜确实是对的. 相应的结果就称为极分解，它很漂亮地描述了 $V$ 上的任意算子.

注意，如果 $T\in\mathcal{L}(V)$，那么 $T^*T$ 是正算子【如 7.64 (a) 中所示】. 因此，算子 $\sqrt{T^*T}$ 是有意义的，而且明确定义为 $V$ 上的正算子.

我们将要陈述并证明的极分解，说的是 $V$ 上的每个算子都是一幺正算子和一正算子的乘积. 这样，我们就能将 $V$ 上的任意算子写成两个性质良好的算子的乘积，并且这两种算子我们既能全面地描述也能很清楚地理解. 幺正算子由 $\mathbf{F}=\mathbf{C}$ 情形下的 7.55 描述，而正算子由实谱定理（7.29）和复谱定理（7.31）描述.

具体而言，考虑 $\mathbf{F}=\mathbf{C}$ 的情形，设

$$T=S\sqrt{T^*T}$$

是算子 $T\in\mathcal{L}(V)$ 的极分解，其中 $S$ 是一幺正算子. 那么就存在 $V$ 的规范正交基使得 $S$ 关于其有对角矩阵，也存在 $V$ 的规范正交基使得 $\sqrt{T^*T}$ 关于其有对角矩阵. **注意：**可能并不存在一个规范正交基，能同时将 $S$ 和 $\sqrt{T^*T}$ 的矩阵都化为对角矩阵这样的好形式——可能 $S$ 需要一个规范正交基，而 $\sqrt{T^*T}$ 需要另一个规范正交基.

不过（仍然假设 $\mathbf{F} = \mathbf{C}$），如果 $T$ 是正规的，那就可以选出一个规范正交基，使得 $S$ 和 $\sqrt{T^*T}$ 关于这一个基都有对角矩阵——见本节习题 31. 逆命题也是对的：如果 $T \in \mathcal{L}(V)$ 且 $T = S\sqrt{T^*T}$，其中 $S \in \mathcal{L}(V)$ 是幺正算子且 $S$ 和 $\sqrt{T^*T}$ 关于同一规范正交基有对角矩阵，那么 $T$ 是正规的. 具体而言，在这一逆命题的条件下，可得 $T$ 关于同一规范正交基有对角矩阵，也就可得 $T$ 是正规的【根据 7.31 中 (c) 和 (a) 的等价关系】，即结论成立.

下面的极分解在实内积空间和复内积空间上都成立，对这些空间上的所有算子都适用.

**7.93 极分解（polar decomposition）**

设 $T \in \mathcal{L}(V)$，那么存在幺正算子 $S \in \mathcal{L}(V)$ 使得

$$T = S\sqrt{T^*T}.$$

**证明** 令 $s_1, \ldots, s_m$ 是 $T$ 的正奇异值，令 $e_1, \ldots, e_m$ 和 $f_1, \ldots, f_m$ 是 $V$ 中的规范正交组，使得

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \tag{7.94}$$

对任一 $v \in V$ 都成立. 将 $e_1, \ldots, e_m$ 和 $f_1, \ldots, f_m$ 扩充为 $V$ 的规范正交基 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$.

定义 $S \in \mathcal{L}(V)$ 为

$$Sv = \langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n$$

对任一 $v \in V$ 都成立. 那么

$$\begin{aligned} \|Sv\|^2 &= \|\langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n\|^2 \\ &= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 \\ &= \|v\|^2. \end{aligned}$$

因此，$S$ 是幺正算子.

将 $T^*$ 作用于式 (7.94) 两侧，再用由式 (7.77) 给出的 $T^*$ 的表达式，可得

$$T^*Tv = s_1^2 \langle v, e_1 \rangle e_1 + \cdots + s_m^2 \langle v, e_m \rangle e_m$$

对任一 $v \in V$ 都成立. 那么，如果 $v \in V$ 就有

$$\sqrt{T^*T}v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_m \langle v, e_m \rangle e_m.$$

这是因为，将 $v$ 映成上式右侧这一表达式的算子为正算子且其平方就等于 $T^*T$. 这样一来，就有

$$\begin{aligned} S\sqrt{T^*T}v &= S(s_1 \langle v, e_1 \rangle e_1 + \cdots + s_m \langle v, e_m \rangle e_m) \\ &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \\ &= Tv, \end{aligned}$$

其中最后一个等式由式 (7.94) 得到. ■

本节习题 27 表明，以上这一证明中得到的幺正算子 $S$，是最接近 $T$ 的幺正算子.

极分解还有其他的证法，其中直接使用了谱定理而回避了奇异值分解. 然而，以上的证明看起来比那些证明更简洁.

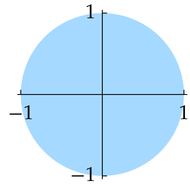
## 作用于椭球和平行体的算子

**7.95 定义：球（ball）、$B$**

$V$ 中半径为 1、以 0 为心的**球（ball）**，记为 $B$，定义为

$$B = \{v \in V : \|v\| < 1\}.$$

如果 $\dim V = 2$，那么我们有时候用的词是**“圆盘”（disk）**而不是**“球”**. 然而，在所有维度中都用**“球”**这个词可以避免混乱. 类似地，如果 $\dim V = 2$，那么我们有时候用的词是**“椭圆”（ellipse）**而不是我们将要定义的**“椭球”**. 同样，在所有维度中都用**“椭球”**这个词可以避免混乱.



$\mathbf{R}^2$ 中的球 $B$.

以下定义的椭球可以视为将球 $B$ 沿每个 $f_k$ 轴伸缩至 $s_k$ 倍得到的.

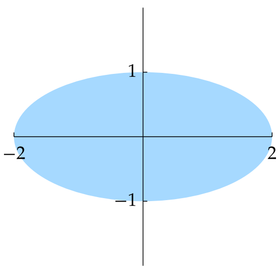
**7.96 定义：椭球（ellipsoid）、$E(s_1 f_1, \ldots, s_n f_n)$，主轴（principal axes）**

设 $f_1, \ldots, f_n$ 是 $V$ 的规范正交基，$s_1, \ldots, s_n$ 是正数. **主轴**为 $s_1 f_1, \ldots, s_n f_n$ 的**椭球** $E(s_1 f_1, \ldots, s_n f_n)$ 定义为

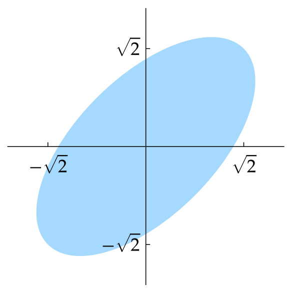
$$E(s_1 f_1, \ldots, s_n f_n) = \left\{ v \in V : \frac{|\langle v, f_1 \rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n \rangle|^2}{s_n^2} < 1 \right\}.$$

椭球的记号 $E(s_1 f_1, \ldots, s_n f_n)$ 并没有显式地包含内积空间 $V$，尽管以上的定义依赖于 $V$. 然而，从上下文以及“$f_1, \ldots, f_n$ 是 $V$ 的规范正交基”这一要求来看，内积空间 $V$ 应该是明确的.

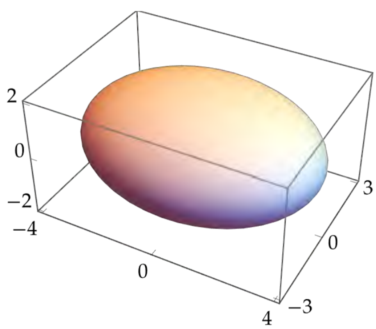
**7.97 例：椭球**



$\mathbf{R}^2$ 中的椭球 $E(2f_1, f_2)$，其中 $f_1, f_2$ 是 $\mathbf{R}^2$ 的标准基.



$\mathbf{R}^2$ 中的椭球 $E(2f_1, f_2)$，其中 $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), f_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.



$\mathbf{R}^3$ 中的椭球 $E(4f_1, 3f_2, 2f_3)$，其中 $f_1, f_2, f_3$ 是 $\mathbf{R}^3$ 的标准基.

对于 $V$ 的任一规范正交基 $f_1, \ldots, f_n$，椭球 $E(f_1, \ldots, f_n)$ 都等于 $V$ 中的球 $B$【根据帕塞瓦尔恒等式 6.30 (b)】.

**7.98 记号：$T(\Omega)$**

对于定义在 $V$ 上的函数 $T$，以及 $\Omega \subseteq V$，定义 $T(\Omega)$ 为

$$T(\Omega) = \{Tv : v \in \Omega\}.$$

那么，如果 $T$ 是定义在 $V$ 上的函数，就有 $T(V) = \operatorname{range} T$.

接下来这条结果指出，每个可逆算子 $T \in \mathcal{L}(V)$ 都将 $V$ 中的球 $B$ 映成 $V$ 中的椭球. 其证明表明这个椭球的主轴就由 $T$ 的奇异值分解得到.

**7.99 可逆算子化球为椭球**

设 $T \in \mathcal{L}(V)$ 可逆，那么 $T$ 将 $V$ 中的球 $B$ 映成 $V$ 中的椭球.

**证明** 设 $T$ 有如下奇异值分解：

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n \tag{7.100}$$

对所有 $v \in V$ 成立，其中 $s_1, \ldots, s_n$ 是 $T$ 的奇异值，而 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$ 都是 $V$ 的规范正交基. 我们将证明 $T(B) = E(s_1 f_1, \ldots, s_n f_n)$.

首先设 $v \in B$. 因为 $T$ 可逆，所以奇异值 $s_1, \ldots, s_n$ 都不等于 0（见 7.68）. 那么，由式 (7.100) 可得

$$\frac{|\langle Tv, f_1 \rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n \rangle|^2}{s_n^2} = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 < 1.$$

因此，$Tv \in E(s_1 f_1, \ldots, s_n f_n)$. 于是，$T(B) \subseteq E(s_1 f_1, \ldots, s_n f_n)$.

为证明另一方向的包含关系，现在设 $w \in E(s_1 f_1, \ldots, s_n f_n)$，令

$$v = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_n \rangle}{s_n} e_n.$$

那么，由 $\|v\| < 1$ 和式 (7.100)，可得 $Tv = \langle w, f_1 \rangle f_1 + \cdots + \langle w, f_n \rangle f_n = w$. 于是，$T(B) \supseteq E(s_1 f_1, \ldots, s_n f_n)$. ■

我们现在用前一条结果证明，可逆算子将所有椭球——不只是半径为 1 的球——都映成椭球.

**7.101 可逆算子化椭球为椭球**

设 $T \in \mathcal{L}(V)$ 可逆，且 $E$ 是 $V$ 中的椭球. 那么 $T(E)$ 是 $V$ 中的椭球.

**证明** 存在 $V$ 的规范正交基 $f_1, \ldots, f_n$ 和正数 $s_1, \ldots, s_n$ 使得 $E = E(s_1 f_1, \ldots, s_n f_n)$. 定义 $S \in \mathcal{L}(V)$ 为

$$S(a_1 f_1 + \cdots + a_n f_n) = a_1 s_1 f_1 + \cdots + a_n s_n f_n.$$

那么，$S$ 将 $V$ 中的球 $B$ 映成 $E$（你可以验证）. 因此，

$$T(E) = T\big(S(B)\big) = (TS)(B).$$

结合上式，并将 7.99 应用于 $TS$，就可得到 $T(E)$ 是 $V$ 中的椭球. ■

回顾一下（见 3.95），如果 $u \in V$，$\Omega \subseteq V$，那么 $u+\Omega$ 定义为

$$u+\Omega=\{u+w : w \in \Omega\}.$$

几何上，集合 $\Omega$ 和 $u+\Omega$ 看起来一样，但是位置不同.

在接下来的定义中，如果 $\dim V=2$，那么我们经常用**“平行四边形”（parallelogram）**一词而不是**“平行体”**.

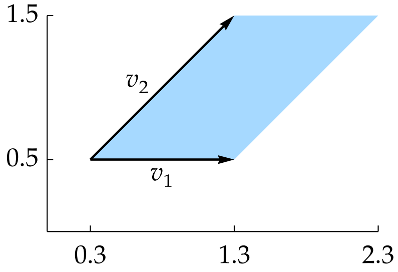
**7.102 定义：$P(v_1,\ldots,v_n)$、平行体（parallelepiped）**

设 $v_1,\ldots,v_n$ 是 $V$ 的基，令

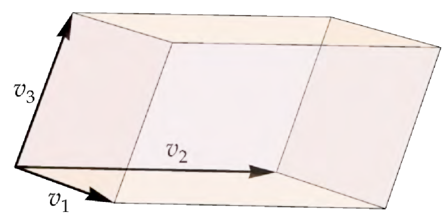
$$P(v_1,\ldots,v_n)=\{a_1v_1+\cdots+a_nv_n : a_1,\ldots,a_n \in (0,1)\}.$$

**平行体**是形如 $u+P(v_1,\ldots,v_n)$ 的集合，其中 $u \in V$. 向量 $v_1,\ldots,v_n$ 称为该平行体的**边（edges）**. ♣

**7.103 例：平行体**




$\mathbf{R}^2$ 中的平行体 $(0.3,0.5)+P\big((1,0),(1,1)\big)$.




$\mathbf{R}^3$ 中的一个平行体.

**7.104 可逆算子化平行体为平行体**

设 $u \in V$ 且 $v_1,\ldots,v_n$ 是 $V$ 的基. 设 $T \in \mathcal{L}(V)$ 可逆，那么

$$T\big(u+P(v_1,\ldots,v_n)\big)=Tu+P(Tv_1,\ldots,Tv_n).$$

♡

**证明** 因为 $T$ 可逆，所以组 $Tv_1,\ldots,Tv_n$ 是 $V$ 的基. $T$ 是线性的，就意味着

$$T(u+a_1v_1+\cdots+a_nv_n)=Tu+a_1Tv_1+\cdots+a_nTv_n$$

对所有 $a_1,\ldots,a_n \in (0,1)$ 成立. 因此，$T\big(u+P(v_1,\ldots,v_n)\big)=Tu+P(Tv_1,\ldots,Tv_n)$. ■

正如我们把矩形从 $\mathbf{R}^2$ 中的平行四边形里特别区分出来一样，我们也给 $V$ 中定义边（defining edges）[6]互相正交的平行体取一个特别的名字.

**7.105 定义：长方体（box）**

$V$ 中的**长方体**[7]是形如

$$u+P(r_1e_1,\ldots,r_ne_n)$$

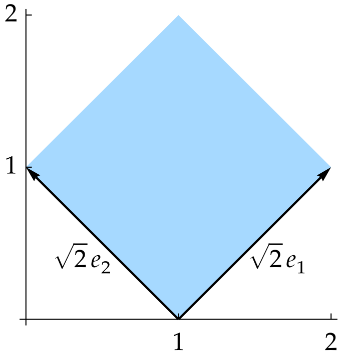
的集合，其中 $u \in V$，$r_1,\ldots,r_n$ 是正数，$e_1,\ldots,e_n$ 是 $V$ 的规范正交基. ♣

[6]即 7.102 中所说的“边”.

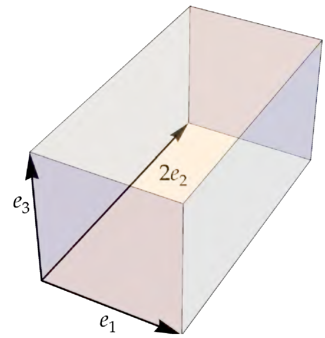
[7]译者更喜欢它的另一个名字——超矩形（hyperrectangle）.

注意到，在 $\mathbf{R}^2$ 这一特殊情形中，长方体就是矩形，但是**“长方体”**这一术语适用于所有维度.

**7.106 例：长方体**




长方体 $(1,0)+P(\sqrt{2}e_1,\sqrt{2}e_2)$，其中 $e_1=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}), e_2=(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$.




长方体 $P(e_1,2e_2,e_3)$，其中 $e_1,e_2,e_3$ 是 $\mathbf{R}^3$ 的标准基.

设 $T\in\mathcal{L}(V)$ 可逆. 那么 $T$ 将 $V$ 中每个平行体都映成 $V$ 中的平行体（根据 7.104）. 特别地，$T$ 将 $V$ 中每个长方体都映成 $V$ 中的平行体. 这就提出一个问题：$T$ 是否把 $V$ 中的某些长方体映成 $V$ 中的长方体？接下来这条结果借助奇异值分解回答了这个问题.

**7.107 每个可逆算子都将某些长方体化成长方体**

设 $T\in\mathcal{L}(V)$ 可逆. 设 $T$ 有奇异值分解

$$Tv=s_1\langle v,e_1\rangle f_1+\cdots+s_n\langle v,e_n\rangle f_n,$$

其中，$s_1,\ldots,s_n$ 是 $T$ 的奇异值，$e_1,\ldots,e_n$ 和 $f_1,\ldots,f_n$ 是 $V$ 的规范正交基，上式对所有 $v\in V$ 成立. 那么，对于所有正数 $r_1,\ldots,r_n$ 和所有 $u\in V$，$T$ 将长方体 $u+P(r_1e_1,\ldots,r_ne_n)$ 映成长方体 $Tu+P(r_1s_1f_1,\ldots,r_ns_nf_n)$.

**证明** 如果 $a_1,\ldots,a_n\in(0,1)$，$r_1,\ldots,r_n$ 是正数，$u\in V$，那么

$$T(u+a_1r_1e_1+\cdots+a_nr_ne_n)=Tu+a_1r_1s_1f_1+\cdots+a_nr_ns_nf_n.$$

因此，$T\big(u+P(r_1e_1,\ldots,r_ne_n)\big)=Tu+P(r_1s_1f_1,\ldots,r_ns_nf_n)$. ■

## 通过奇异值计算体积

本小节我们的目标是，理解算子如何改变其定义域子集的体积. 因为体积的概念属于分析学而非线性代数，所以我们仅仅讨论体积的直觉概念. 而在分析学的体系之下，我们用来处理体积的直观方法可以转化为适当的正确定义、正确陈述和正确证明.

我们关于体积的直觉，在实内积空间上最有效. 因此在本小节的剩余部分，经常会出现 $\mathbf{F}=\mathbf{R}$ 这一假设.

如果 $\dim V = n$，那么我们说的**"体积"**指的是 $n$ 维体积．你应该很熟悉 $\mathbf{R}^3$ 中的这一概念．当 $n=2$ 时，我们通常称之为"面积"而不是"体积"，但是为了保持一致，我们在所有维度上都用"体积"这个词．关于体积最基本的直觉就是，长方体（根据定义，其定义边互相正交）的体积是所有定义边长度的乘积．因此，我们做出以下定义．

**7.108 定义：长方体的体积（volume of a box）**

设 $\mathbf{F}=\mathbf{R}$．如果 $u \in V$，$r_1,\ldots,r_n$ 是正数，$e_1,\ldots,e_n$ 是 $V$ 的规范正交基，那么

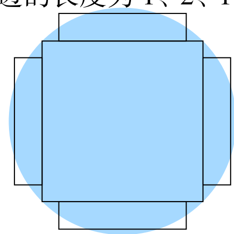
$$\operatorname{volume}\big(u + P(r_1 e_1,\ldots,r_n e_n)\big) = r_1 \times \cdots \times r_n.$$

♣

以上定义跟常见的 $\mathbf{R}^2$ 中矩形面积（这里我们称为体积）公式和 $\mathbf{R}^3$ 中长方体的体积公式是一致的．例如，例 7.106 中第一个长方体具有二维体积（或面积），其值为 2，因为其定义边的长度为 $\sqrt{2}$ 和 $\sqrt{2}$；第二个长方体具有三维体积，其值为 2，因为其定义边的长度为 1、2、1.

为了定义 $V$ 的子集的体积，我们用有限个互不相交的长方体来逼近这一子集，并且将用于逼近的这些长方体的体积加起来．我们对越多的不相交长方体取并集，就能越准确地逼近 $V$ 的一子集，对其体积的逼近也越好．

这些想法应该让你想起了黎曼积分（Riemann integral）是怎样定义的——用若干互不相交的矩形来逼近曲线下的面积．上面的讨论引出了以下不严谨但合乎直觉的定义．



该球的体积 ≈ 五个长方体的体积之和．

**7.109 定义：体积（volume）**

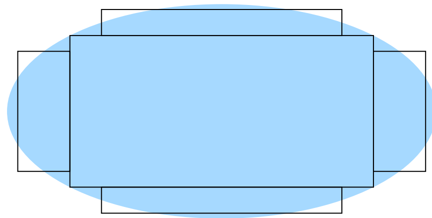
设 $\mathbf{F}=\mathbf{R}$，$\Omega \subseteq V$．那么 $\Omega$ 的**体积**，记为 $\operatorname{volume}\Omega$，约等于逼近 $\Omega$ 的若干个不相交长方体的体积之和．

♣

我们用一种直观的方法处理体积，如此就忽略了许多值得推敲的问题．例如，如果我们用关于一个基的长方体逼近 $\Omega$，那么我们再用关于另一个基的长方体逼近 $\Omega$ 又能否得到相同的体积？如果 $\Omega_1$ 和 $\Omega_2$ 是 $V$ 的不相交子集，那么是否有 $\operatorname{volume}(\Omega_1 \cup \Omega_2) = \operatorname{volume}\Omega_1 + \operatorname{volume}\Omega_2$？假如我们仅考虑 $V$ 的性质相当好的子集，那么借助分析学的手段可以说明，这两个问题都有肯定的回答，和我们对体积的直觉一致．

**7.110 例：线性映射改变体积**

设 $T \in \mathcal{L}(\mathbf{R}^2)$ 定义为 $Tv = 2\langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2$，其中 $e_1, e_2$ 是 $\mathbf{R}^2$ 的标准基．这一线性映射使向量沿 $e_1$ 轴拉伸至 2 倍而沿 $e_2$ 轴不变．上面那个由五个长方体逼近的球，由 $T$ 映射为此处展示的椭球．原始图片中的五个长方体，每一个都被映射为宽度为原来的两倍而高度跟原来相同的长方体．于是每个长方体都被映射为体积（面积）为原来两倍的长方体．这五个新长方体的体积总和就近似为这个椭球的体积．所以，$T$ 将球的体积改变为原来的 2 倍．



此图中每个长方体的宽度都是上一张图中的两倍，而其高度不变．

在以上例子中，$T$ 将关于基 $e_1, e_2$ 的长方体映射为关于同一基的长方体，我们可以从中看出 $T$ 是怎样改变体积的．一般来说，算子将长方体映射为不是长方体的平行体．然而，如果我们选取适当的基（来自奇异值分解的基！），那么关于该基的长方体就能被映射为关于可能不同的一个基的长方体，如 7.107 所示．观察到这一点，就能得出接下来这条结果的自然证明．

**7.111　体积变化倍数是奇异值的乘积**

设 $\mathbf{F} = \mathbf{R}$，$T \in \mathcal{L}(V)$ 可逆，且 $\Omega \subseteq V$．那么

$$\operatorname{volume} T(\Omega) = (T \text{ 的奇异值的乘积})(\operatorname{volume} \Omega).$$

**证明**　设 $T$ 有奇异值分解

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

对所有 $v \in V$ 成立，其中 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$ 是 $V$ 的规范正交基.

用形如 $u + P(r_1 e_1, \ldots, r_n e_n)$ 的长方体逼近 $\Omega$，这些长方体的体积是 $r_1 \times \cdots \times r_n$．算子 $T$ 将每个长方体 $u + P(r_1 e_1, \ldots, r_n e_n)$ 都映成长方体 $Tu + P(r_1 s_1 f_1, \ldots, r_n s_n f_n)$，其体积为 $(s_1 \times \cdots \times s_n)(r_1 \times \cdots \times r_n)$.

算子 $T$ 将用于逼近 $\Omega$ 的那些长方体映成用于逼近 $T(\Omega)$ 的那些长方体．因为 $T$ 改变了用于逼近 $\Omega$ 的每个长方体的体积，对应倍数为 $s_1 \times \cdots \times s_n$，所以线性映射 $T$ 按同样的倍数改变 $\Omega$ 的体积. ■

设 $T \in \mathcal{L}(V)$．等我们讨论到行列式的时候，我们就会看到，$T$ 的奇异值的乘积就等于 $|\det T|$，见 9.60 和 9.61.

## 取决于特征值的算子性质

我们以下面这个表格结束本章．注意，这个表格描述的是有限维复内积空间中的情形．表格的第一列显示的是正规算子在有限维复内积空间上可能具有的性质；表格的第二列显示的是相应的 $\mathbf{C}$ 的子集——当且仅当一算子的所有特征值都属于这一特定子集，该算子才具有左边那列中相应的性质．例如，表格的第一行指出，正规算子可逆当且仅当它的所有特征值非零（这第一行是表格中唯一不需要假设算子正规的）.

你得确保自己可以解释表格里所有结果为什么都成立．例如，表格的最后一行成立，是因为算子的范数等于它的最大奇异值（根据 7.85），而正规算子的奇异值——假定 $\mathbf{F} = \mathbf{C}$——等于特征值的绝对值（根据 7E 节习题 7）.

| **正规算子的性质** | **特征值属于** |
|---|---|
| 可逆 | $\mathbf{C} \setminus \{0\}$ |
| 自伴 | $\mathbf{R}$ |
| 斜 | $\{\lambda \in \mathbf{C} : \operatorname{Re} \lambda = 0\}$ |
| 正交投影 | $\{0, 1\}$ |
| 正 | $[0, \infty)$ |
| 幺正 | $\{\lambda \in \mathbf{C} : \lvert\lambda\rvert = 1\}$ |
| 范数小于 1 | $\{\lambda \in \mathbf{C} : \lvert\lambda\rvert < 1\}$ |

## 习题 7F

**1** 证明：如果 $S, T \in \mathcal{L}(V, W)$，那么 $\big|\|S\| - \|T\|\big| \leq \|S - T\|$.

注 以上不等式被称为**反向三角不等式**.

**2** 设 $T \in \mathcal{L}(V)$ 是自伴的，或者设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ 是正规的. 证明：

$$\|T\| = \max\left\{|\lambda| : \lambda \text{ 是 } T \text{ 的特征值}\right\}.$$

**3** 设 $T \in \mathcal{L}(V, W)$，$v \in V$. 证明：

$$\|Tv\| = \|T\|\,\|v\| \iff T^*Tv = \|T\|^2 v.$$

**4** 设 $T \in \mathcal{L}(V, W)$，$v \in V$ 且 $\|Tv\| = \|T\|\,\|v\|$. 证明：如果 $u \in V$ 且 $\langle u, v\rangle = 0$，那么 $\langle Tu, Tv\rangle = 0$.

**5** 设 $U$ 是有限维内积空间，$T \in \mathcal{L}(V, U)$，$S \in \mathcal{L}(U, W)$. 证明：

$$\|ST\| \leq \|S\|\,\|T\|.$$

**6** 证明或给出一反例：如果 $S, T \in \mathcal{L}(V)$，那么 $\|ST\| = \|TS\|$.

**7** 证明：定义为 $d(S, T) = \|S - T\|$（$S, T \in \mathcal{L}(V, W)$）的 $d$ 是 $\mathcal{L}(V, W)$ 上的度量.

注 本题是为熟悉度量空间的读者准备的.

**8** (a) 证明：如果 $T \in \mathcal{L}(V)$ 且 $\|I - T\| < 1$，那么 $T$ 可逆.

(b) 设 $S \in \mathcal{L}(V)$ 可逆. 证明：如果 $T \in \mathcal{L}(V)$ 且 $\|S - T\| < 1/\big\|S^{-1}\big\|$，那么 $T$ 是可逆的.

注 本题表明，$\mathcal{L}(V)$ 中的可逆算子构成的集合是 $\mathcal{L}(V)$ 的开子集，这里所用的是第 7 题中定义的度量.

**9** 设 $T \in \mathcal{L}(V)$. 证明：对于任意 $\epsilon > 0$，都存在可逆算子 $S \in \mathcal{L}(V)$ 使得 $0 < \|T - S\| < \epsilon$.

**10** 设 $\dim V > 1$，$T \in \mathcal{L}(V)$ 不可逆. 证明：对于任意 $\epsilon > 0$，都存在 $S \in \mathcal{L}(V)$ 使得 $0 < \|T - S\| < \epsilon$ 且 $S$ 不可逆.

**11** 设 $\mathbf{F} = \mathbf{C}$，$T \in \mathcal{L}(V)$. 证明：对于任意 $\epsilon > 0$，都存在可对角化算子 $S \in \mathcal{L}(V)$ 使得 $0 < \|T - S\| < \epsilon$.

**12** 设 $T \in \mathcal{L}(V)$ 是正算子. 证明：$\big\|\sqrt{T}\big\| = \sqrt{\|T\|}$.

**13** 设 $S, T \in \mathcal{L}(V)$ 是正算子. 证明：

$$\|S - T\| \leq \max\{\|S\|, \|T\|\} \leq \|S + T\|.$$

**14** 设 $U$ 和 $W$ 是 $V$ 的子空间且 $\|P_U - P_W\| < 1$. 证明：$\dim U = \dim W$.

**15** 定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 是

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

求出使得 $T = S\sqrt{T^*T}$ 的幺正算子 $S \in \mathcal{L}(\mathbf{F}^3)$ 的显式表达式.

**16** 设 $S \in \mathcal{L}(V)$ 是可逆正算子. 证明：存在 $\delta > 0$ 使得任一满足 $\|S - T\| < \delta$ 的自伴算子 $T \in \mathcal{L}(V)$ 都是正算子.

**17** 证明：如果 $u \in V$ 而 $\varphi_u$ 是 $V$ 上定义为 $\varphi_u(v) = \langle v, u\rangle$ 的线性泛函，那么 $\|\varphi_u\| = \|u\|$.

注 这里我们将标量域 $\mathbf{F}$ 视为内积空间，其内积为 $\langle \alpha, \beta\rangle = \alpha\overline{\beta}$ 对所有 $\alpha, \beta \in \mathbf{F}$ 成立. 因此，$\|\varphi_u\|$ 意即 $\varphi_u$ 作为从 $V$ 到 $\mathbf{F}$ 的线性映射的范数.

**18** 设 $e_1,\ldots,e_n$ 是 $V$ 的规范正交基，$T\in\mathcal{L}(V,W)$.

(a) 证明：$\max\{\|Te_1\|,\ldots,\|Te_n\|\}\le\|T\|\le\left(\|Te_1\|^2+\cdots+\|Te_n\|^2\right)^{1/2}$.

(b) 证明：$\|T\|=\left(\|Te_1\|^2+\cdots+\|Te_n\|^2\right)^{1/2}$，当且仅当 $\dim\operatorname{range}T\le 1$.

注　这里的 $e_1,\ldots,e_n$ 是 $V$ 的任意一个规范正交基，而不一定跟 $T$ 的奇异值分解有关. 如果 $s_1,\ldots,s_n$ 是 $T$ 的奇异值，那么以上不等式的右侧就等于 $\left(s_1^2+\cdots+s_n^2\right)^{1/2}$，如 7E 节习题 11 (a) 所示.

**19** 证明：如果 $T\in\mathcal{L}(V,W)$，那么 $\|T^*T\|=\|T\|^2$.

注　$\|T^*T\|$ 的这个表达式，引出了 $C^*$-代数（$C^*$-algebras）这一重要学科.

**20** 设 $T\in\mathcal{L}(V)$ 是正规的. 证明：对于任一正整数 $k$，都有 $\left\|T^k\right\|=\|T\|^k$ 成立.

**21** 设 $\dim V>1$ 且 $\dim W>1$. 证明：$\mathcal{L}(V,W)$ 上的范数并不来自内积. 换句话说，是证明：不存在 $\mathcal{L}(V,W)$ 上的内积使得

$$\max\left\{\|Tv\|: v\in V \text{ 且 } \|v\|\le 1\right\}=\sqrt{\langle T,T\rangle}$$

对所有 $T\in\mathcal{L}(V,W)$ 成立.

**22** 设 $T\in\mathcal{L}(V,W)$. 令 $n=\dim V$，令 $s_1\ge\cdots\ge s_n$ 表示 $T$ 的奇异值. 证明：如果 $1\le k\le n$，那么

$$\min\left\{\|T|_U\|: U \text{ 是 } V \text{ 的子空间且 } \dim U=k\right\}=s_{n-k+1}.$$

**23** 设 $T\in\mathcal{L}(V,W)$. 证明：关于 $V$ 和 $W$ 上源自其范数的度量（见 6B 节习题 23），$T$ 是一致连续的.

**24** 设 $T\in\mathcal{L}(V)$ 可逆. 证明：

$$\left\|T^{-1}\right\|=\|T\|^{-1}\iff\frac{T}{\|T\|}\text{ 是幺正算子.}$$

**25** 取定 $u,x\in V$，其中 $u\ne 0$. 定义 $T\in\mathcal{L}(V)$ 为：对任一 $v\in V$ 都有 $Tv=\langle v,u\rangle x$. 证明：

$$\sqrt{T^*T}v=\frac{\|x\|}{\|u\|}\langle v,u\rangle u$$

对任一 $v\in V$ 都成立.

**26** 设 $T\in\mathcal{L}(V)$. 证明：$T$ 可逆，当且仅当存在唯一幺正算子 $S\in\mathcal{L}(V)$ 使得 $T=S\sqrt{T^*T}$.

**27** 设 $T\in\mathcal{L}(V)$，$s_1,\ldots,s_n$ 是 $T$ 的奇异值. 令 $e_1,\ldots,e_n$ 和 $f_1,\ldots,f_n$ 都是 $V$ 的规范正交基，使得

$$Tv=s_1\langle v,e_1\rangle f_1+\cdots+s_n\langle v,e_n\rangle f_n$$

对所有 $v\in V$ 成立. 定义 $S\in\mathcal{L}(V)$ 为

$$Sv=\langle v,e_1\rangle f_1+\cdots+\langle v,e_n\rangle f_n.$$

(a) 证明：$S$ 是幺正的且 $\|T-S\|=\max\{|s_1-1|,\ldots,|s_n-1|\}$.

(b) 证明：如果 $E\in\mathcal{L}(V)$ 是幺正的，那么 $\|T-E\|\ge\|T-S\|$.

注　本题所求的是，（在幺正算子当中）尽可能接近给定算子 $T$ 的幺正算子 $S$.

**28** 设 $T\in\mathcal{L}(V)$. 证明：存在幺正算子 $S\in\mathcal{L}(V)$ 使得 $T=\sqrt{TT^*}S$.

**29** 设 $T \in \mathcal{L}(V)$.

(a) 用极分解证明：存在幺正算子 $S \in \mathcal{L}(V)$ 使得 $TT^* = ST^*TS^*$.

(b) 证明：(a) 蕴涵着 $T$ 和 $T^*$ 的奇异值相同.

**30** 设 $T \in \mathcal{L}(V)$，而 $S \in \mathcal{L}(V)$ 是幺正算子，$R \in \mathcal{L}(V)$ 是使得 $T = SR$ 的正算子. 证明：$R = \sqrt{T^*T}$.

注　本题表明，如果我们将 $T$ 写成一个幺正算子和一个正算子的乘积（如 7.93 的极分解所示），那么其中的正算子就等于 $\sqrt{T^*T}$.

**31** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ 是正规的. 证明：存在幺正算子 $S \in \mathcal{L}(V)$ 使得 $T = S\sqrt{T^*T}$，并且 $S$ 和 $\sqrt{T^*T}$ 关于 $V$ 的同一个规范正交基有对角矩阵.

**32** 设 $T \in \mathcal{L}(V, W)$，$T \neq 0$. 令 $s_1, \ldots, s_m$ 表示 $T$ 的正奇异值. 证明：存在 $(\operatorname{null} T)^{\perp}$ 的规范正交基 $e_1, \ldots, e_m$ 使得

$$T\left(E\left(\frac{e_1}{s_1}, \ldots, \frac{e_m}{s_m}\right)\right)$$

等于 $\operatorname{range} T$ 中以 0 为心而半径为 1 的球.

# 第 8 章 复向量空间上的算子

在本章中，我们更深入地研究算子的结构，并将大多注意力放在复向量空间上．本章中的有些结论同时适用于实向量空间和复向量空间，于是我们并不总是假定 $\mathbf{F}=\mathbf{C}$．并且，内积对于本章内容并无用处，所以我们回到有限维向量空间这个一般的假定上来．

即便是在有限维复向量空间上，一个算子也可能不具有足够的特征向量来形成向量空间的基．于是我们将考察一种名为广义特征向量的对象，它与上述问题联系紧密．我们将看到，对于有限维复向量空间上的每个算子，都存在由该算子的广义特征向量构成的该空间的基．届时，利用广义特征空间分解就能很好地描述有限维复向量空间上的任意算子．

幂零算子是自乘若干次后会等于 0 的算子，在上面这些问题的研究中，它们扮演着重要的角色．在我们证明有限维复向量空间上的每个可逆算子都有平方根，以及研究若当型的时候，幂零算子都是关键的工具．

本章结尾处定义了迹，并证明了它的关键性质．

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 代表 $\mathbf{F}$ 上的非零有限维向量空间.



图为都柏林大学（University of Dublin）古图书馆的长阅读室（Long Room of the Old Library）．威廉·哈密尔顿（William Hamilton，1805-1865）曾在此地求学，后又成为这里的一名教员，他在 1853 年证明了我们现在所说的凯莱-哈密尔顿定理（Cayley-Hamilton theorem）的一个特例．

# 8A 广义特征向量和幂零算子

## 算子的幂的零空间

我们从研究算子的幂的零空间开启本章.

**8.1 递增的零空间序列**

设 $T \in \mathcal{L}(V)$. 那么

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^k \subseteq \operatorname{null} T^{k+1} \subseteq \cdots.$$

**证明** 设 $k$ 是非负整数，且 $v \in \operatorname{null} T^k$. 那么 $T^k v = 0$，这表明 $T^{k+1} v = T(T^k v) = T(0) = 0$. 从而 $v \in \operatorname{null} T^{k+1}$. 因此 $\operatorname{null} T^k \subseteq \operatorname{null} T^{k+1}$，原命题得证. ■

下面结论表明，如果上述子空间序列中有两个相邻项相等，那么排在这两项之后的所有项也都相等.

关于递减的值域序列，也有类似的结论，可参看习题 6、7、8.

**8.2 零空间序列中的等式**

设 $T \in \mathcal{L}(V)$，$m$ 是非负整数，满足

$$\operatorname{null} T^m = \operatorname{null} T^{m+1}.$$

那么

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+2} = \operatorname{null} T^{m+3} = \cdots.$$

**证明** 令 $k$ 是正整数. 我们想要证明

$$\operatorname{null} T^{m+k} = \operatorname{null} T^{m+k+1}.$$

由 8.1，我们已知 $\operatorname{null} T^{m+k} \subseteq \operatorname{null} T^{m+k+1}$.

为了证明此包含关系反过来也成立，设 $v \in \operatorname{null} T^{m+k+1}$. 那么

$$T^{m+1}(T^k v) = T^{m+k+1} v = 0.$$

因此

$$T^k v \in \operatorname{null} T^{m+1} = \operatorname{null} T^m.$$

于是 $T^{m+k} v = T^m(T^k v) = 0$，意味着 $v \in \operatorname{null} T^{m+k}$. 这表明 $\operatorname{null} T^{m+k+1} \subseteq \operatorname{null} T^{m+k}$，从而完成了证明. ■

上面结论引出了一个问题：是否存在非负整数 $m$，使得 $\operatorname{null} T^m = \operatorname{null} T^{m+1}$？下面结论就说明，这个等式至少在 $m$ 等于 $T$ 的定义空间的维数时会成立.

**8.3 零空间停止增长**

设 $T \in \mathcal{L}(V)$. 那么

$$\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1} = \operatorname{null} T^{\dim V+2} = \cdots.$$

**证明** 我们只需证明 $\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1}$（由 8.2）. 假设此式不成立. 那么，由 8.1 和 8.2，我们有

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^{\dim V} \subsetneq \operatorname{null} T^{\dim V+1},$$

其中的真包含号 $\subsetneq$ 意为“包含于但不等于”. 在上式的每个真包含号处，维数至少会增加 1. 因此 $\dim \operatorname{null} T^{\dim V+1} \geq \dim V + 1$，这就推出了矛盾——因为 $V$ 的子空间的维数不会大于 $\dim V$. ■

$V = \operatorname{null} T \oplus \operatorname{range} T$ 并不对每个 $T \in \mathcal{L}(V)$ 都成立. 不过，有下面这条好用的替代结论.

**8.4 $V$ 是 $\operatorname{null} T^{\dim V}$ 与 $\operatorname{range} T^{\dim V}$ 的直和**

设 $T \in \mathcal{L}(V)$. 那么

$$V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}.$$

**证明** 令 $n = \dim V$. 首先我们证明

$$(\operatorname{null} T^n) \cap (\operatorname{range} T^n) = \{0\}. \tag{8.5}$$

设 $v \in (\operatorname{null} T^n) \cap (\operatorname{range} T^n)$. 那么 $T^n v = 0$，并且存在 $u \in V$ 使得 $v = T^n u$. 将 $T^n$ 作用于此式两侧，可得 $T^n v = T^{2n} u$. 因此 $T^{2n} u = 0$，这表明 $T^n u = 0$（由 8.3）. 于是 $v = T^n u = 0$，这就证明了式 (8.5).

这样一来，由式 (8.5) 即得 $\operatorname{null} T^n + \operatorname{range} T^n$ 是直和（由 1.46）. 并且，

$$\dim(\operatorname{null} T^n \oplus \operatorname{range} T^n) = \dim \operatorname{null} T^n + \dim \operatorname{range} T^n = \dim V,$$

其中第一个等号源于 3.94，第二个等号来自线性映射基本定理（3.21）. 上述等式就表明 $\operatorname{null} T^n \oplus \operatorname{range} T^n = V$（参看 2.39），即原命题得证. ■

上述结论的加强版本见习题 19.

---

**8.6 例：对于 $T \in \mathcal{L}(\mathbf{F}^3)$，$\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$**

定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 为

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

从而 $\operatorname{null} T = \{(z_1, 0, 0) : z_1 \in \mathbf{F}\}$ 且 $\operatorname{range} T = \{(z_1, 0, z_3) : z_1, z_3 \in \mathbf{F}\}$. 于是 $\operatorname{null} T \cap \operatorname{range} T \neq \{0\}$. 于是 $\operatorname{null} T + \operatorname{range} T$ 不是直和. 同时，注意到 $\operatorname{null} T + \operatorname{range} T \neq \mathbf{F}^3$. 然而，我们有 $T^3(z_1, z_2, z_3) = (0, 0, 125z_3)$. 从而我们可知

$$\operatorname{null} T^3 = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{F}\} \quad 且 \quad \operatorname{range} T^3 = \{(0, 0, z_3) : z_3 \in \mathbf{F}\}.$$

因此正如 8.4 所述，$\mathbf{F}^3 = \operatorname{null} T^3 \oplus \operatorname{range} T^3$.

---

## 广义特征向量

仅凭特征向量不足以很好地描述有些算子. 于是在本小节中，我们引入广义特征向量的概念，它将在描述算子的结构方面发挥重要作用.

为了理解为何需要拓展特征向量，我们来考虑这个问题：通过将一算子的定义空间分解为

若干不变子空间来描述该算子．取定 $T \in \mathcal{L}(V)$．我们想通过找到一个"好"的直和分解

$$V = V_1 \oplus \cdots \oplus V_n$$

（其中各 $V_k$ 都是 $V$ 的在 $T$ 下不变的子空间）来描述 $T$．最简单的非零不变子空间是一维的．而当且仅当 $V$ 具有由 $T$ 的特征向量构成的基时，上述直和分解式中各 $V_k$ 才全是 $V$ 的在 $T$ 下不变的一维子空间（见 5.55）．这又等价于 $V$ 具有特征空间分解

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T), \tag{8.7}$$

其中 $\lambda_1, \ldots, \lambda_m$ 是 $T$ 的所有互异特征值（见 5.55）．

由上章中的谱定理可得，若 $V$ 是内积空间，那么形如式 (8.7) 的分解对于每个自伴算子（若 $\mathbf{F} = \mathbf{R}$）及每个正规算子（若 $\mathbf{F} = \mathbf{C}$）都成立，因为这两类算子的特征向量都足以形成 $V$ 的基（见 7.29 和 7.31）．

然而，对于更一般的算子来说，形如式 (8.7) 的分解不一定成立，即便在复向量空间上也是如此．比如，5.57 中展示的算子的特征向量就不足以使式 (8.7) 成立．我们现在要引入的广义特征向量和广义特征空间就能改善这种局面．

**8.8 定义：广义特征向量（generalized eigenvector）**

设 $T \in \mathcal{L}(V)$，$\lambda$ 是 $T$ 的特征值．称向量 $v \in V$ 是 $T$ 对应于 $\lambda$ 的**广义特征向量**，若 $v \neq 0$ 且对某个正整数 $k$ 有

$$(T - \lambda I)^k v = 0.$$

♣

非零向量 $v \in V$ 是 $T$ 对应于 $\lambda$ 的广义特征向量，当且仅当

$$(T - \lambda I)^{\dim V} v = 0.$$

将 8.1 和 8.3 应用于算子 $T - \lambda I$ 即可得出这点．

我们并未定义广义特征值，因为这样不会得到任何新东西．原因是：若 $(T - \lambda I)^k$ 对于某个正整数 $k$ 不是单射，那么 $T - \lambda I$ 也不是单射，因此 $\lambda$ 就是 $T$ 的特征值．

我们知道，复向量空间上的算子的特征向量可能不足以形成定义空间的基．下面结论说明，这些算子的广义特征向量总能形成定义空间的基．

**8.9 由广义特征向量构成的基**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$．那么存在由 $T$ 的广义特征向量构成的 $V$ 的基．

♡

**证明** 令 $n = \dim V$．我们将对 $n$ 用归纳法．首先，注意到欲证结论在 $n = 1$ 情形成立，因为这样的话 $V$ 中每个非零向量都是 $T$ 的特征向量．

现假设 $n > 1$，并假设欲证结论在 $\dim V$ 为更小的值时都成立．设 $\lambda$ 是 $T$ 的特征值．将 8.4 应用于 $T - \lambda I$，可得

$$V = \operatorname{null}(T - \lambda I)^n \oplus \operatorname{range}(T - \lambda I)^n.$$

正是这一步用上了 $\mathbf{F} = \mathbf{C}$ 这个前提条件，因为如果 $\mathbf{F} = \mathbf{R}$，那么 $T$ 可能不存在特征值．

若 $\operatorname{null}(T - \lambda I)^n = V$，那么 $V$ 中每个非零向量都是 $T$ 的广义特征向量，于是在这种情况下，存在由 $T$ 的广义特征向量构成的 $V$ 的基．因此我们可设 $\operatorname{null}(T - \lambda I)^n \neq V$，这表明 $\operatorname{range}(T - \lambda I)^n \neq \{0\}$．

同时，$\operatorname{null}(T-\lambda I)^n \neq \{0\}$，因为 $\lambda$ 是 $T$ 的一个特征值．于是我们有

$$0 < \dim \operatorname{range}(T-\lambda I)^n < n.$$

此外，$\operatorname{range}(T-\lambda I)^n$ 在 $T$ 下不变【由 5.18，其中取 $p(z)=(z-\lambda)^n$】．令 $S \in \mathcal{L}(\operatorname{range}(T-\lambda I)^n)$ 等于限制在 $\operatorname{range}(T-\lambda I)^n$ 上的算子 $T$．将归纳假设应用于算子 $S$，可得存在由 $S$ 的广义特征向量构成的 $\operatorname{range}(T-\lambda I)^n$ 的基，而这些向量当然也是 $T$ 的广义特征向量．将 $\operatorname{range}(T-\lambda I)^n$ 的这个基与 $\operatorname{null}(T-\lambda I)^n$ 的基合并，就得到了由 $T$ 的广义特征向量构成的 $V$ 的基． ■

若 $\mathbf{F}=\mathbf{R}$ 且 $\dim V>1$，那么 $V$ 上有些算子的广义特征向量可构成 $V$ 的基，其他算子则不具有该性质（与 $\mathbf{F}=\mathbf{C}$ 情形不同）．判定一算子是否具有该性质的充分必要条件见于习题 11.

**8.10 例：$\mathbf{C}^3$ 上一算子的广义特征向量**

定义 $T \in \mathcal{L}(\mathbf{C}^3)$ 为：对每个 $(z_1,z_2,z_3) \in \mathbf{C}^3$，

$$T(z_1,z_2,z_3)=(4z_2,0,5z_3).$$

由特征值的定义即知 $T$ 的特征值是 0 和 5．此外，对应于特征值 0 的特征向量是形如 $(z_1,0,0)$ 的非零向量，对应于特征值 5 的特征向量是形如 $(0,0,z_3)$ 的非零向量．因此该算子的特征向量不足以张成该算子的定义空间 $\mathbf{C}^3$．

我们算出 $T^3(z_1,z_2,z_3)=(0,0,125z_3)$．于是 8.1 和 8.3 表明，$T$ 对应于特征值 0 的广义特征向量是形如 $(z_1,z_2,0)$ 的非零向量．

同时，我们也有 $(T-5I)^3(z_1,z_2,z_3)=(-125z_1+300z_2,-125z_2,0)$．因此，$T$ 对应于特征值 5 的广义特征向量是形如 $(0,0,z_3)$ 的非零向量．

由上面几段的论述，可知 $\mathbf{C}^3$ 的标准基中每个向量都是 $T$ 的广义特征向量．于是正如 8.9 所言，$\mathbf{C}^3$ 的确具有由 $T$ 的广义特征向量构成的基.

若 $v$ 是 $T \in \mathcal{L}(V)$ 的特征向量，那么其对应的特征值 $\lambda$ 由方程 $Tv=\lambda v$ 唯一确定，且该方程只对一个 $\lambda \in \mathbf{F}$ 成立（因 $v \neq 0$）．然而，若 $v$ 是 $T$ 的广义特征向量，那么我们并不显然可知是否仅有一个 $\lambda \in \mathbf{F}$ 能满足方程 $(T-\lambda I)^{\dim V}v=0$．好在下面结论告诉我们，该问题的回答是肯定的.

**8.11 广义特征向量对应于唯一的特征值**

设 $T \in \mathcal{L}(V)$．那么 $T$ 的每个广义特征向量都仅对应于 $T$ 的一个特征值.

**证明** 设 $v \in V$ 是 $T$ 的同时对应于 $T$ 的两个特征值 $\alpha$ 和 $\lambda$ 的广义特征向量．令 $m$ 是满足 $(T-\alpha I)^m v=0$ 的最小正整数．令 $n=\dim V$．那么

$$\begin{aligned} 0 &= (T-\lambda I)^n v \\ &= \big((T-\alpha I)+(\alpha-\lambda)I\big)^n v \\ &= \sum_{k=0}^{n} b_k(\alpha-\lambda)^{n-k}(T-\alpha I)^k v, \end{aligned}$$

其中 $b_0=1$，其余各二项式系数 $b_k$ 的值无关紧要．将算子 $(T-\alpha I)^{m-1}$ 作用于上式两边，可得

$$0=(\alpha-\lambda)^n(T-\alpha I)^{m-1}v.$$

因为 $(T-\alpha I)^{m-1}v \neq 0$，所以上式表明 $\alpha=\lambda$，则原命题得证. ■

我们之前看到，对应于不同特征值的特征向量是线性无关的（5.11）. 现在我们证明适用于广义特征向量的类似结论，所用证法大致遵循了之前那个结论的证明思路.

**8.12 线性无关的广义特征向量**

设 $T\in\mathcal{L}(V)$. 那么由对应于 $T$ 的互异特征值的广义特征向量构成的每个向量组都是线性无关的.

**证明** 设欲证结论不成立. 那么存在最小正整数 $m$，使得对应于 $T$ 的互异特征值 $\lambda_1,\ldots,\lambda_m$ 的广义特征向量 $v_1,\ldots,v_m$ 构成线性相关向量组（注意 $m\geq 2$，因为由定义可知广义特征向量是非零的）. 于是，存在全不为 0（因为 $m$ 最小）的数 $a_1,\ldots,a_m\in\mathbf{F}$ 使得

$$a_1v_1+\cdots+a_mv_m=0.$$

令 $n=\dim V$. 将 $(T-\lambda_mI)^n$ 作用于上式两侧，可得

$$a_1(T-\lambda_mI)^nv_1+\cdots+a_{m-1}(T-\lambda_mI)^nv_{m-1}=0. \tag{8.13}$$

设 $k\in\{1,\ldots,m-1\}$. 那么

$$(T-\lambda_mI)^nv_k\neq 0.$$

否则 $v_k$ 就会成为 $T$ 的同时对应于不同特征值 $\lambda_k$ 和 $\lambda_m$ 的广义特征向量，这会与 8.11 相矛盾. 又有

$$(T-\lambda_kI)^n\big((T-\lambda_mI)^nv_k\big)=(T-\lambda_mI)^n\big((T-\lambda_kI)^nv_k\big)=0.$$

从而，综合上述两式得 $(T-\lambda_mI)^nv_k$ 是 $T$ 对应于特征值 $\lambda_k$ 的广义特征向量. 因此

$$(T-\lambda_mI)^nv_1,\ldots,(T-\lambda_mI)^nv_{m-1}$$

是由 $m-1$ 个对应于互异特征值的广义特征向量所构成的线性相关组【由式 (8.13)】，这就与 $m$ 最小相矛盾. 由该矛盾即可证明原命题. ■

## 幂零算子

**8.14 定义：幂零（nilpotent）**

称一个算子是**幂零的**，如果它的某个幂等于 0.

于是，算子 $T\in\mathcal{L}(V)$ 是幂零的，若 $V$ 中每个非零向量都是 $T$ 对应于特征值 0 的广义特征向量.

**8.15 例：幂零算子**

(a) 定义为

$$T(z_1,z_2,z_3,z_4)=(0,0,z_1,z_2)$$

的算子 $T\in\mathcal{L}(\mathbf{F}^4)$ 是幂零的，因为 $T^2=0$.

转下页

(b) 关于标准基的矩阵为

$$\begin{pmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{pmatrix}$$

的 $\mathbf{F}^3$ 上的算子是幂零的，通过求该矩阵的立方得到零矩阵即可证明.

(c) $\mathcal{P}_m(\mathbf{R})$ 上的微分算子是幂零的，因为每个不超过 $m$ 次的多项式的 $m+1$ 阶导数都等于 0. 注意，在这个维数是 $m+1$ 的向量空间上，我们需求该算子的 $m+1$ 次幂才能得到算子 0.

---

下面结论说明，当我们求一个幂零算子的幂时，不用考虑幂次高于其定义空间的维数的幂. 一个稍微强一些的结论见习题 18.

拉丁文词语 **"nil"** 意为"无"或"零"，**"potens"** 意为"幂". 于是 **"nilpotent"** 的字面意思就是"幂零".

**8.16 $n$ 维空间上幂零算子的 $n$ 次幂等于 0**

设 $T \in \mathcal{L}(V)$ 是幂零的. 那么 $T^{\dim V} = 0$.

**证明** 因为 $T$ 是幂零的，所以存在正整数 $k$ 使得 $T^k = 0$. 于是 $\operatorname{null} T^k = V$. 这样，由 8.1 和 8.3 即知 $\operatorname{null} T^{\dim V} = V$. 于是 $T^{\dim V} = 0$. ∎

**8.17 幂零算子的特征值**

设 $T \in \mathcal{L}(V)$.

(a) 如果 $T$ 是幂零的，那么 0 是 $T$ 的特征值，并且 $T$ 没有其他的特征值.

(b) 若 $\mathbf{F} = \mathbf{C}$，且 0 是 $T$ 的唯一特征值，那么 $T$ 是幂零的.

**证明**

(a) 为证明 (a)，设 $T$ 是幂零的. 因此存在正整数 $m$ 使得 $T^m = 0$. 这就表明 $T$ 不是单射. 于是 0 是 $T$ 的特征值.

为证明 $T$ 没有其他特征值，设 $\lambda$ 是 $T$ 的特征值. 那么存在非零向量 $v \in V$ 使得

$$\lambda v = Tv.$$

将 $T$ 反复作用于上式两端，可得

$$\lambda^m v = T^m v = 0.$$

于是 $\lambda = 0$，原命题得证.

(b) 设 $\mathbf{F} = \mathbf{C}$ 且 0 是 $T$ 的唯一特征值. 根据 5.27 (b)，$T$ 的最小多项式等于 $z^m$（$m$ 为正整数）. 于是 $T^m = 0$. 因此 $T$ 是幂零的. ∎

习题 23 说明了上述结论 (b) 中 $\mathbf{F} = \mathbf{C}$ 这个前提条件不可删去.

给定 $V$ 上的一个算子，我们想要求出 $V$ 的一个基，使得该算子关于该基的矩阵尽可能简单，意即矩阵中包含尽可能多的 0. 下面结论说明，如果 $T$ 是幂零的，那么我们可以选取 $V$ 的一个基，使得 $T$ 关于该基的矩阵有超过一半的元素等于 0. 在本章后面部分我们还将得到更好的结论.

**8.18 幂零算子的最小多项式和上三角矩阵**

设 $T \in \mathcal{L}(V)$. 那么下面各命题等价.

(a) $T$ 是幂零的.

(b) $T$ 的最小多项式等于 $z^m$（$m$ 为正整数）.

(c) 存在 $V$ 的一个基，使得 $T$ 关于该基的矩阵形如

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

其中对角线及对角线下方各元素均等于 0.

**证明** 假设 (a) 成立，那么 $T$ 是幂零的. 于是存在正整数 $n$ 使得 $T^n = 0$. 那么由 5.29 可得 $z^n$ 是 $T$ 的最小多项式的多项式倍. 于是 $T$ 的最小多项式就是 $z^m$（$m$ 为正整数）. 这证明了 (a) 蕴涵 (b).

现设 (b) 成立，那么 $T$ 的最小多项式是 $z^m$（其中 $m$ 为正整数）. 由此，根据 5.27 (a) 可得，0（$z^m$ 的唯一零点）是 $T$ 的唯一特征值；根据 5.44 还有，存在 $V$ 的一个基，使得 $T$ 关于该基具有上三角矩阵. 进而，由 5.41 可知，该矩阵中对角线上的所有元素都等于 0. 这就证明了 (b) 蕴涵 (c).

现设 (c) 成立. 那么由 5.40 可得 $T^{\dim V} = 0$. 于是 $T$ 是幂零的，这就证明了 (c) 蕴涵 (a). ■

## 习题 8A

**1** 设 $T \in \mathcal{L}(V)$. 证明：若 $\dim \operatorname{null} T^4 = 8$ 且 $\dim \operatorname{null} T^6 = 9$，那么对于所有整数 $m \geq 5$，有 $\dim \operatorname{null} T^m = 9$.

**2** 设 $T \in \mathcal{L}(V)$，$m$ 是正整数，$v \in V$，且 $T^{m-1}v \neq 0$ 但 $T^m v = 0$. 证明：$v, Tv, T^2v, \ldots, T^{m-1}v$ 是线性无关的.

**注** 本习题的结果将被用于 8.45 的证明中.

**3** 设 $T \in \mathcal{L}(V)$. 证明：

$$V = \operatorname{null} T \oplus \operatorname{range} T \iff \operatorname{null} T^2 = \operatorname{null} T.$$

**4** 设 $T \in \mathcal{L}(V)$，$\lambda \in \mathbf{F}$，$m$ 为正整数，使得 $T$ 的最小多项式是 $(z-\lambda)^m$ 的多项式倍. 证明：

$$\dim \operatorname{null}(T - \lambda I)^m \geq m.$$

**5** 设 $T \in \mathcal{L}(V)$，$m$ 是正整数. 证明：

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

**提示** 3B 节习题 21 可能有用.

**6** 设 $T \in \mathcal{L}(V)$. 证明

$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots \supseteq \operatorname{range} T^k \supseteq \operatorname{range} T^{k+1} \supseteq \cdots.$$

**7** 设 $T \in \mathcal{L}(V)$，$m$ 是非负整数，使得

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

证明：对于所有 $k > m$，有 $\operatorname{range} T^k = \operatorname{range} T^m$.

**8** 设 $T \in \mathcal{L}(V)$. 证明

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V+1} = \operatorname{range} T^{\dim V+2} = \cdots.$$

**9** 设 $T \in \mathcal{L}(V)$，$m$ 是非负整数. 证明

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \iff \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

**10** 定义 $T \in \mathcal{L}(\mathbf{C}^2)$ 为 $T(w,z) = (z,0)$. 求出 $T$ 的所有广义特征向量.

**11** 设 $T \in \mathcal{L}(V)$. 证明：$V$ 有由 $T$ 的广义特征向量构成的基，当且仅当 $T$ 的最小多项式等于 $(z-\lambda_1)\cdots(z-\lambda_m)$（其中 $\lambda_1,\ldots,\lambda_m \in \mathbf{F}$）.

注　假设 $\mathbf{F} = \mathbf{R}$，因为 $\mathbf{F} = \mathbf{C}$ 情形由 5.27 (b) 和 8.9 可证.

这道习题表明，“$V$ 有由 $T$ 的广义特征向量构成的基”的条件，与“$V$ 有一个基使得 $T$ 关于该基有上三角矩阵”的条件（见 5.44）是一样的.

**注意**：如果 $T$ 关于 $V$ 的基 $v_1,\ldots,v_n$ 有上三角矩阵，那么 $v_1$ 是 $T$ 的特征向量，但 $v_2,\ldots,v_n$ 不一定是 $T$ 的广义特征向量.

**12** 设 $T \in \mathcal{L}(V)$ 满足 $V$ 中的每个向量都是 $T$ 的广义特征向量. 证明：存在 $\lambda \in \mathbf{F}$，使得 $T - \lambda I$ 是幂零的.

**13** 设 $S,T \in \mathcal{L}(V)$ 且 $ST$ 是幂零的. 证明：$TS$ 是幂零的.

**14** 设 $T \in \mathcal{L}(V)$ 是幂零的且 $T \neq 0$. 证明：$T$ 不可对角化.

**15** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明：$T$ 可对角化当且仅当 $T$ 的每个广义特征向量都是 $T$ 的特征向量.

注　对于 $\mathbf{F} = \mathbf{C}$，本习题给 5.55 中的可对角化等价条件列表又增加了一个等价条件.

**16** (a) 举出一例：相同向量空间上的幂零算子 $S,T$，使得无论 $S+T$ 还是 $ST$ 都不是幂零的.

(b) 设 $S,T \in \mathcal{L}(V)$ 是幂零的且 $ST = TS$. 证明 $S+T$ 和 $ST$ 是幂零的.

**17** 设 $T \in \mathcal{L}(V)$ 是幂零的，$m$ 是正整数，满足 $T^m = 0$.

(a) 证明：$I - T$ 是可逆的，且 $(I-T)^{-1} = I + T + \cdots + T^{m-1}$.

(b) 解释你会如何猜想出上述公式.

**18** 设 $T \in \mathcal{L}(V)$ 是幂零的. 证明：$T^{1+\dim \operatorname{range} T} = 0$.

注　若 $\dim \operatorname{range} T < \dim V - 1$，那么本习题加强了 8.16 的结论.

**19** 设 $T \in \mathcal{L}(V)$ 不是幂零的. 证明：$V = \operatorname{null} T^{\dim V-1} \oplus \operatorname{range} T^{\dim V-1}$.

注　对于不是幂零的算子，本习题加强了 8.4 的结论.

**20** 设 $V$ 是内积空间，$T \in \mathcal{L}(V)$ 是正规的且是幂零的. 证明 $T = 0$.

**21** 设 $T \in \mathcal{L}(V)$ 满足 $\operatorname{null} T^{\dim V-1} \neq \operatorname{null} T^{\dim V}$. 证明：$T$ 是幂零的，且对于每个满足 $0 \le k \le \dim V$ 的整数 $k$，有 $\dim \operatorname{null} T^k = k$.

**22** 设 $T \in \mathcal{L}(\mathbf{C}^5)$ 满足 $\operatorname{range} T^4 \neq \operatorname{range} T^5$. 证明 $T$ 是幂零的.

**23** 举出一例：$T$ 是有限维实向量空间上的算子且 0 是 $T$ 的唯一特征值，但 $T$ 不是幂零的.

注 本习题表明，去掉 $\mathbf{F}=\mathbf{C}$ 的前提后，8.17 中的蕴涵关系 (b)$\Rightarrow$(a) 不成立.

**24** 对例 8.15 中的每条示例，求出算子的定义空间的一个基，使得该幂零算子关于该基的矩阵具有 8.18 (c) 所示的上三角形式.

**25** 设 $V$ 是内积空间，$T \in \mathcal{L}(V)$ 是幂零的. 证明：存在 $V$ 的一个规范正交基，使得 $T$ 关于该基的矩阵具有 8.18 (c) 所示的上三角形式.

# 8B 广义特征空间分解

## 广义特征空间

**8.19 定义：广义特征空间（generalized eigenspace）、$G(\lambda, T)$**

设 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$. $T$ 对应于 $\lambda$ 的**广义特征空间**，记作 $G(\lambda, T)$，定义为

$$G(\lambda, T) = \{v \in V : (T - \lambda I)^k v = 0, k \text{ 为某正整数}\}.$$

于是，$G(\lambda, T)$ 是由 $T$ 对应于 $\lambda$ 的广义特征向量以及向量 0 所构成的集合.

因为 $T$ 的每个特征向量都是 $T$ 的广义特征向量（在广义特征向量的定义中取 $k = 1$ 即可），所以每个特征空间都包含于相对应的广义特征空间. 换言之，若 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$，那么 $E(\lambda, T) \subseteq G(\lambda, T)$.

下面结论表明，如果 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$，那么广义特征空间 $G(\lambda, T)$ 是 $V$ 的一个子空间（因为 $V$ 上每个线性映射的零空间都是 $V$ 的子空间）.

**8.20 广义特征空间的描述**

设 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$. 那么 $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$.

**证明** 设 $v \in \operatorname{null}(T-\lambda I)^{\dim V}$. 由广义特征空间的定义知 $v \in G(\lambda, T)$. 则 $G(\lambda, T) \supseteq \operatorname{null}(T-\lambda I)^{\dim V}$.

反之，设 $v \in G(\lambda, T)$. 于是存在正整数 $k$ 使得 $v \in \operatorname{null}(T - \lambda I)^k$. 由 8.1 和 8.3（将其中 $T$ 用 $T - \lambda I$ 替代），我们可得 $v \in \operatorname{null}(T - \lambda I)^{\dim V}$. 于是 $G(\lambda, T) \subseteq \operatorname{null}(T - \lambda I)^{\dim V}$，这就完成了证明. ■

---

**8.21 例：$\mathbf{C}^3$ 上一算子的广义特征空间**

定义 $T \in \mathcal{L}(\mathbf{C}^3)$ 为

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3).$$

在例 8.10 中，我们得知 $T$ 的特征值是 0 和 5，且求出了与之对应的广义特征向量所构成的集合. 将这些集合分别与 $\{0\}$ 取并集，我们有

$$G(0, T) = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{C}\} \quad \text{及} \quad G(5, T) = \{(0, 0, z_3) : z_3 \in \mathbf{C}\}.$$

注意，$\mathbf{C}^3 = G(0, T) \oplus G(5, T)$.

---

在例 8.21 中，定义空间 $\mathbf{C}^3$ 是该例中算子 $T$ 的广义特征空间的直和. 接下来的结论就表明，这个性质是普遍成立的. 具体而言，下面这个关键结论说明了，如果 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$，那么 $V$ 是 $T$ 的广义特征空间的直和，该直和中各项都在 $T$ 下不变，并且作用于其上的 $T$ 等于一个幂零算子加上恒等算子的标量倍. 于是，借助下面结论，我们得以完成我们的目标：将 $V$ 分解成不变子空间，且 $T$ 在这些子空间上的性质为我们所知.

我们将看到，以下证明过程综合运用了目前所学有关广义特征空间的知识，随后利用了这个结论：对于每个算子 $T \in \mathcal{L}(V)$，都存在由 $T$ 的广义特征向量构成的 $V$ 的基.

**8.22 广义特征空间分解**

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$. 令 $\lambda_1,\dots,\lambda_m$ 是 $T$ 的所有互异特征值. 那么

(a) 对每个 $k=1,\dots,m$，$G(\lambda_k,T)$ 在 $T$ 下是不变的；

(b) 对每个 $k=1,\dots,m$，$(T-\lambda_k I)|_{G(\lambda_k,T)}$ 是幂零的；

(c) $V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T)$.

♡

**证明**

(a) 设 $k\in\{1,\dots,m\}$. 那么 8.20 表明

$$G(\lambda_k,T)=\operatorname{null}(T-\lambda_k I)^{\dim V}.$$

于是由 5.18【其中取 $p(z)=(z-\lambda_k)^{\dim V}$】，得 $G(\lambda_k,T)$ 在 $T$ 下是不变的，(a) 得证.

(b) 设 $k\in\{1,\dots,m\}$. 若 $v\in G(\lambda_k,T)$，那么 $(T-\lambda_k I)^{\dim V}v=0$（由 8.20）. 于是 $\big((T-\lambda_k I)|_{G(\lambda_k,T)}\big)^{\dim V}=0$，因此 $(T-\lambda_k I)|_{G(\lambda_k,T)}$ 是幂零的，(b) 得证.

(c) 为了说明 $G(\lambda_1,T)+\cdots+G(\lambda_m,T)$ 是直和，设

$$v_1+\cdots+v_m=0,$$

其中各 $v_k$ 属于 $G(\lambda_k,T)$. 因为 $T$ 对应于互异特征值的广义特征向量线性无关（由 8.12），所以此式中各 $v_k$ 等于 0. 从而 $G(\lambda_1,T)+\cdots+G(\lambda_m,T)$ 是直和（由 1.45）.

最后，$V$ 中每个向量都可被写成 $T$ 的广义特征向量的有限和的形式（由 8.9）. 于是

$$V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T),$$

(c) 得证. ■

$\mathbf{F}=\mathbf{R}$ 时的类似结论见于习题 8.

## 特征值的重数

若 $V$ 是复向量空间且 $T\in\mathcal{L}(V)$，那么由广义特征空间分解（8.22）给出的 $V$ 的分解式是个有力的工具. 该分解式中各子空间的维数很重要，因而在下面定义中被赋予特殊的名称.

**8.23 定义：重数（multiplicity）**

- 设 $T\in\mathcal{L}(V)$. 定义 $T$ 的特征值 $\lambda$ 的**重数**为其对应的广义特征空间 $G(\lambda,T)$ 的维数.
- 换言之，$T$ 的特征值 $\lambda$ 的**重数**等于

$$\dim\operatorname{null}(T-\lambda I)^{\dim V}.$$

♣

上述第二点成立是因为 $G(\lambda,T)=\operatorname{null}(T-\lambda I)^{\dim V}$（见 8.20）.

**8.24 例：一算子的各特征值的重数**

定义 $T\in\mathcal{L}(\mathbf{C}^3)$ 为

$$T(z_1,z_2,z_3)=(6z_1+3z_2+4z_3,\,6z_2+2z_3,\,7z_3).$$

转下页

$T$ 关于标准基的矩阵为

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

由 5.41 可得，$T$ 的特征值是矩阵对角线上的元素 6 和 7. 你可自行验证，$T$ 的广义特征空间是

$$G(6,T) = \operatorname{span}\big((1,0,0),(0,1,0)\big) \quad 和 \quad G(7,T) = \operatorname{span}\big((10,2,1)\big).$$

于是，特征值 6 的重数是 2，特征值 7 的重数是 1. 由 8.22 所述广义特征空间分解可写出直和 $\mathbf{C}^3 = G(6,T) \oplus G(7,T)$. 如 8.9 所言，$T$ 的广义特征向量 $(1,0,0),(0,1,0),(10,2,1)$ 构成 $\mathbf{C}^3$ 的一个基. $\mathbf{C}^3$ 中不存在由该算子的特征向量构成的基.

---

在上例中，$T$ 的特征值的重数之和等于 3，这正是 $T$ 的定义空间的维数. 下面结论说明，这个性质对于有限维复向量空间上的每个算子都成立.

在本例中，每个特征值的重数，都等于该特征值在算子的上三角矩阵对角线上出现的次数. 我们将在 8.31 中看到，这个性质总是成立的.

**8.25 重数之和等于 $\dim V$**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 那么 $T$ 的所有特征值的重数之和等于 $\dim V$.

**证明** 由广义特征空间分解（8.22）与直和的维数公式（见 3.94）即可得欲证结论. ■

有些书中会使用**代数重数**（**algebraic multiplicity**）和**几何重数**（**geometric multiplicity**）这两个术语. 如果碰到这两个术语，你应明白代数重数与此处定义的重数是一样的，而几何重数是相应的特征空间的维数. 换言之，若 $T \in \mathcal{L}(V)$ 且 $\lambda$ 是 $T$ 的一个特征值，那么

$$\lambda \text{ 的代数重数} = \dim \operatorname{null}(T-\lambda I)^{\dim V} = \dim G(\lambda,T),$$
$$\lambda \text{ 的几何重数} = \dim \operatorname{null}(T-\lambda I) = \dim E(\lambda,T).$$

注意，按照上述定义，代数重数作为某个零空间的维数，同样有几何意义. 此处给出的重数定义比涉及行列式的传统定义更加简洁，9.62 将说明这两种定义是等价的.

若 $V$ 是内积空间，$T \in \mathcal{L}(V)$ 是正规的，并且 $\lambda$ 是 $T$ 的一个特征值，那么将 7A 节习题 27 应用于正规算子 $T-\lambda I$ 上，即可见 $\lambda$ 的代数重数等于 $\lambda$ 的几何重数. 由此顺便可得，$S \in \mathcal{L}(V,W)$（此处 $V$ 和 $W$ 都是有限维内积空间）的奇异值依赖于自伴算子 $S^*S$ 的特征值的重数（代数重数或几何重数）.

下面定义将有限维复向量空间上的每个算子都与一个首一多项式联系起来.

**8.26 定义：特征多项式（characteristic polynomial）**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 令 $\lambda_1,\ldots,\lambda_m$ 表示 $T$ 的所有互异特征值，且其重数分别为 $d_1,\ldots,d_m$. 称多项式

$$(z-\lambda_1)^{d_1}\cdots(z-\lambda_m)^{d_m}$$

为 $T$ 的**特征多项式**.

**8.27 例：一个算子的特征多项式**

设 $T \in \mathcal{L}(\mathbf{C}^3)$ 定义如例 8.24. 因为 $T$ 的特征值 6 的重数是 2，特征值 7 的重数是 1，所以我们可得 $T$ 的特征多项式是 $(z-6)^2(z-7)$.

**8.28 特征多项式的次数和零点**

设 $\mathbf{F}=\mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 那么

(a) $T$ 的特征多项式的次数是 $\dim V$；

(b) $T$ 的特征多项式的零点就是 $T$ 的特征值.

**证明** 由关于重数之和的结论（8.25）可得 (a). 由特征多项式的定义可得 (b). ■

多数教材利用行列式定义特征多项式（由 9.62，它和此处的定义等价）. 此处采用的处理方法要简洁许多，并可引出凯莱-哈密尔顿定理的下面这个漂亮的证明.

**8.29 凯莱-哈密尔顿定理（Cayley-Hamilton theorem）**

设 $\mathbf{F}=\mathbf{C}$，$T \in \mathcal{L}(V)$，且 $q$ 是 $T$ 的特征多项式. 那么 $q(T)=0$.

**证明** 设 $\lambda_1,\ldots,\lambda_m$ 是 $T$ 的所有互异特征值，并令 $d_k=\dim G(\lambda_k,T)$. 我们知道，对每个 $k \in \{1,\ldots,m\}$，$(T-\lambda_k I)|_{G(\lambda_k,T)}$ 都是幂零的. 于是由 8.16，我们有：对每个 $k \in \{1,\ldots,m\}$，

$$(T-\lambda_k I)^{d_k}|_{G(\lambda_k,T)}=0.$$

广义特征空间分解（8.22）指出，$V$ 中每个向量都是 $G(\lambda_1,T),\ldots,G(\lambda_m,T)$ 中向量的和. 于是，为证明 $q(T)=0$，我们只需证明对各 $k$ 均有 $q(T)|_{G(\lambda_k,T)}=0$.

亚瑟・凯莱（Arthur Cayley，1821-1895）在取得学士学位之前就发表了三篇数学论文.

固定 $k \in \{1,\ldots,m\}$. 我们有

$$q(T)=(T-\lambda_1 I)^{d_1}\cdots(T-\lambda_m I)^{d_m}.$$

上式右侧的算子都是可交换的，因此我们可将乘积项 $(T-\lambda_k I)^{d_k}$ 移至右侧表达式的最后一项. 因为 $(T-\lambda_k I)^{d_k}|_{G(\lambda_k,T)}$ 等于 0，所以我们有 $q(T)|_{G(\lambda_k,T)}=0$，原命题得证. ■

下面结论表明，如果一算子 $T \in \mathcal{L}(V)$ 的最小多项式的次数是 $\dim V$（大多数情况下都是如此——见 5.24 之后的几段话），那么 $T$ 的特征多项式就等于 $T$ 的最小多项式.

**8.30 特征多项式是最小多项式的多项式倍**

设 $\mathbf{F}=\mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 那么 $T$ 的特征多项式是 $T$ 的最小多项式的多项式倍.

**证明** 由凯莱-哈密尔顿定理（8.29）和 5.29 立刻可得上述结论. ■

现在，我们可以证明，例 8.24 体现出的结论对于有限维复向量空间上的所有算子都成立.

**8.31 特征值的重数等于其在对角线上出现的次数**

设 $\mathbf{F}=\mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 设 $v_1,\ldots,v_n$ 是 $V$ 的一个基且使得 $\mathcal{M}(T,(v_1,\ldots,v_n))$ 为上三角矩阵. 那么 $T$ 的每个特征值 $\lambda$ 在 $\mathcal{M}(T,(v_1,\ldots,v_n))$ 对角线上出现的次数，就等于 $\lambda$ 作为 $T$ 的特征值的重数.

**证明** 令 $A = \mathcal{M}(T,(v_1,\ldots,v_n))$. 于是 $A$ 是上三角矩阵. 令 $\lambda_1,\ldots,\lambda_n$ 表示 $A$ 对角线上的各元素. 于是，对每个 $k \in \{1,\ldots,n\}$，我们有

$$Tv_k = u_k + \lambda_k v_k,$$

其中 $u_k \in \operatorname{span}(v_1,\ldots,v_{k-1})$. 因此若 $k \in \{1,\ldots,n\}$ 且 $\lambda_k \neq 0$，那么 $Tv_k$ 不是 $Tv_1,\ldots,Tv_{k-1}$ 的线性组合. 现由线性相关性引理（2.19）可得由这些 $Tv_k$（满足 $\lambda_k \neq 0$）所构成的组是线性无关的.

令 $d$ 表示各下标 $k \in \{1,\ldots,n\}$ 中使得 $\lambda_k = 0$ 的个数. 由上段结论可知

$$\dim \operatorname{range} T \geq n - d.$$

因为 $n = \dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$，所以上述不等式表明

$$\dim \operatorname{null} T \leq d. \tag{8.32}$$

算子 $T^n$ 关于基 $v_1,\ldots,v_n$ 的矩阵是上三角矩阵 $A^n$，且其对角线上的元素是 $\lambda_1^n,\ldots,\lambda_n^n$【见 5C 节习题 2 (b)】. 因为 $\lambda_k^n = 0$ 当且仅当 $\lambda_k = 0$，所以 $A^n$ 的对角线上出现 0 的次数就等于 $d$. 于是利用式 (8.32)（将其中的 $T$ 替换为 $T^n$），我们就有

$$\dim \operatorname{null} T^n \leq d. \tag{8.33}$$

对于 $T$ 的一个特征值 $\lambda$，令 $m_\lambda$ 表示 $\lambda$ 作为 $T$ 的特征值的重数，并令 $d_\lambda$ 表示 $\lambda$ 出现在 $A$ 的对角线上的次数. 将式 (8.33) 中的 $T$ 替换成 $T - \lambda I$，我们就会发现

$$m_\lambda \leq d_\lambda \tag{8.34}$$

对 $T$ 的每个特征值 $\lambda$ 都成立. $T$ 的所有特征值 $\lambda$ 的重数 $m_\lambda$ 之和等于 $n$，也就是 $V$ 的维数（由 8.25）. $T$ 的所有特征值 $\lambda$ 的出现次数 $d_\lambda$ 之和也等于 $n$，因为 $A$ 的对角线上共有 $n$ 个数.

因此，将式 (8.34) 的两端对 $T$ 的所有特征值 $\lambda$ 求和会得到一个等式. 因此，式 (8.34) 对于 $T$ 的每个特征值 $\lambda$ 也一定是个等式. 于是 $\lambda$ 作为 $T$ 的特征值的重数，等于 $\lambda$ 在 $A$ 的对角线上出现的次数，即原命题得证. ■

## 分块对角矩阵

为了以矩阵形式解读我们的结论，我们通过推广对角矩阵的概念，作出如下定义. 如果下述定义中各矩阵 $A_k$ 都是 $1 \times 1$ 矩阵，那么我们其实就得到了对角矩阵.

> 通常，将一矩阵看成由多个小矩阵组合而成，可使我们更好地理解该矩阵.

**8.35 定义：分块对角矩阵（block diagonal matrix）**

一个**分块对角矩阵**是形如

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

的方阵，其中 $A_1,\ldots,A_m$ 是排列在对角线上的方阵，且矩阵其他各元素都等于 0. ♣

**8.36 例：一个分块对角矩阵**

$5 \times 5$ 矩阵

$$A = \begin{pmatrix} (4) & 0 & 0 & 0 & 0 \\ 0 & \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} & & 0 & 0 \\ 0 & & & 0 & 0 \\ 0 & 0 & 0 & \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix} & \\ 0 & 0 & 0 & & \end{pmatrix}$$

是形如

$$A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & A_3 \end{pmatrix},$$

的分块对角矩阵，其中

$$A_1 = \begin{pmatrix} 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}.$$

这里，我们将 $5 \times 5$ 矩阵分解出一块一块的内部矩阵，以便说明我们是怎么将其看成分块对角矩阵的.

注意到，在上例中，$A_1, A_2, A_3$ 都是上三角矩阵，且各自对角线上元素都相等. 下面结论表明，有限维复向量空间中每个算子关于适当的基都具有该形式的矩阵. 注意，这个结论给出的矩阵比一般的上三角矩阵具有更多的零.

**8.37 由上三角块构成的分块对角矩阵**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 令 $\lambda_1, \ldots, \lambda_m$ 是 $T$ 的所有互异特征值，它们的重数分别为 $d_1, \ldots, d_m$. 那么存在 $V$ 的一个基，使得 $T$ 关于该基具有形如

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

的分块对角矩阵，其中各 $A_k$ 是形如

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

的 $d_k \times d_k$ 上三角矩阵.

**证明** 每个 $(T - \lambda_k I)|_{G(\lambda_k, T)}$ 都是幂零的（见 8.22）. 对各 $k$，选取 $G(\lambda_k, T)$（它是维数为 $d_k$ 的向量空间）的一个基，使得 $(T - \lambda_k I)|_{G(\lambda_k, T)}$ 具有形如 8.18 (c) 的矩阵. 而 $T|_{G(\lambda_k, T)} = (T - \lambda_k I)|_{G(\lambda_k, T)} + \lambda_k I|_{G(\lambda_k, T)}$，于是 $T|_{G(\lambda_k, T)}$ 关于该基的矩阵就具有上面所示 $A_k$ 的形式.

广义特征空间分解（8.22）表明，将上面选取的各 $G(\lambda_k, T)$ 的基合并起来，就得到 $V$ 的一个基. $T$ 关于该基的矩阵就具有我们期望的形式. ■

**8.38 例：由广义特征向量得出分块对角矩阵**

定义 $T \in \mathcal{L}(\mathbf{C}^3)$ 为 $T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3)$. $T$（关于标准基）的矩阵是

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix},$$

这是上三角矩阵，但不是 8.37 所给出的形式.

在例 8.24 中我们看到，$T$ 的特征值是 6 和 7，且有

$$G(6, T) = \operatorname{span}\big((1,0,0),(0,1,0)\big) \quad 及 \quad G(7, T) = \operatorname{span}\big((10,2,1)\big).$$

我们还知道由 $T$ 的广义特征向量所构成的 $\mathbf{C}^3$ 的基是

$$(1,0,0),(0,1,0),(10,2,1).$$

$T$ 关于该基的矩阵是

$$\begin{pmatrix} \begin{pmatrix} 6 & 3 \\ 0 & 6 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 7 \end{pmatrix} \end{pmatrix},$$

该矩阵具有 8.37 所给出的分块对角形式.

## 习题 8B

**1** 定义 $T \in \mathcal{L}(\mathbf{C}^2)$ 为 $T(w, z) = (-z, w)$. 求出对应于 $T$ 的所有互异特征值的广义特征空间.

**2** 设 $T \in \mathcal{L}(V)$ 是可逆的. 证明：对每个 $\lambda \in \mathbf{F}$ 且 $\lambda \neq 0$，$G(\lambda, T) = G\left(\frac{1}{\lambda}, T^{-1}\right)$.

**3** 设 $T \in \mathcal{L}(V)$. 设 $S \in \mathcal{L}(V)$ 是可逆的. 证明：$T$ 和 $S^{-1}TS$ 具有相同重数的相同特征值.

**4** 设 $\dim V \geq 2$ 且 $T \in \mathcal{L}(V)$ 满足 $\operatorname{null} T^{\dim V - 2} \neq \operatorname{null} T^{\dim V - 1}$. 证明：$T$ 至多有两个互异的特征值.

**5** 设 $T \in \mathcal{L}(V)$ 且 3 和 8 是 $T$ 的特征值. 令 $n = \dim V$. 证明：$V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

**6** 设 $T \in \mathcal{L}(V)$ 且 $\lambda$ 是 $T$ 的特征值. 解释：$T$ 的最小多项式的因式分解中，$z - \lambda$ 的指数为何是满足 $(T - \lambda I)^m|_{G(\lambda, T)} = 0$ 的最小正整数 $m$.

**7** 设 $T \in \mathcal{L}(V)$，$\lambda$ 是 $T$ 的特征值且重数为 $d$. 证明：$G(\lambda, T) = \operatorname{null}(T - \lambda I)^d$.

注　若 $d < \dim V$，那么本题加强了 8.20 的结论.

**8** 设 $T \in \mathcal{L}(V)$，$\lambda_1, \ldots, \lambda_m$ 是 $T$ 的所有互异特征值. 证明

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

当且仅当 $T$ 的最小多项式等于 $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$（$k_1, \ldots, k_m$ 为正整数）.

注　$\mathbf{F} = \mathbf{C}$ 情形由 5.27 (b) 和广义特征空间分解（8.22）立即可证. 所以这道题只有在 $\mathbf{F} = \mathbf{R}$ 时才有些意思.

**9** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 证明存在 $D, N \in \mathcal{L}(V)$ 使得 $T = D + N$ 成立，其中算子 $D$ 可对角化，$N$ 是幂零的，且 $DN = ND$.

**10** 设 $V$ 是复内积空间，$e_1,\ldots,e_n$ 是 $T$ 的规范正交基，且 $T\in\mathcal{L}(V)$. 令 $\lambda_1,\ldots,\lambda_n$ 是 $T$ 的特征值（其中各特征值出现次数等于其重数）. 证明：

$$|\lambda_1|^2+\cdots+|\lambda_n|^2\le\|Te_1\|^2+\cdots+\|Te_n\|^2.$$

注　见 7A 节习题 5 后的评注.

**11** 举出一个 $\mathbf{C}^4$ 上的算子，其特征多项式等于 $(z-7)^2(z-8)^2$.

**12** 举出一个 $\mathbf{C}^4$ 上的算子，其特征多项式等于 $(z-1)(z-5)^3$ 且最小多项式等于 $(z-1)(z-5)^2$.

**13** 举出一个 $\mathbf{C}^4$ 上的算子，其特征多项式和最小多项式都等于 $z(z-1)^2(z-3)$.

**14** 举出一个 $\mathbf{C}^4$ 上的算子，其特征多项式等于 $z(z-1)^2(z-3)$ 且最小多项式等于 $z(z-1)(z-3)$.

**15** 令 $T$ 是 $\mathbf{C}^4$ 上的算子，定义为 $T(z_1,z_2,z_3,z_4)=(0,z_1,z_2,z_3)$. 求出 $T$ 的特征多项式和最小多项式.

**16** 令 $T$ 是 $\mathbf{C}^6$ 上的算子，定义为

$$T(z_1,z_2,z_3,z_4,z_5,z_6)=(0,z_1,z_2,0,z_4,0).$$

求出 $T$ 的特征多项式和最小多项式.

**17** 设 $\mathbf{F}=\mathbf{C}$, 且 $P\in\mathcal{L}(V)$ 满足 $P^2=P$. 证明: $P$ 的特征多项式是 $z^m(z-1)^n$, 其中 $m=\dim\operatorname{null}P$ 且 $n=\dim\operatorname{range}P$.

**18** 设 $T\in\mathcal{L}(V)$，$\lambda$ 是 $T$ 的特征值. 解释为什么下面这四个数都相等.

(a) $T$ 的最小多项式的因式分解中，$z-\lambda$ 的指数.

(b) 满足 $(T-\lambda I)^m|_{G(\lambda,T)}=0$ 的最小正整数 $m$.

(c) 满足

$$\operatorname{null}(T-\lambda I)^m=\operatorname{null}(T-\lambda I)^{m+1}$$

的最小正整数 $m$.

(d) 满足

$$\operatorname{range}(T-\lambda I)^m=\operatorname{range}(T-\lambda I)^{m+1}$$

的最小正整数 $m$.

**19** 设 $\mathbf{F}=\mathbf{C}$ 且 $S\in\mathcal{L}(V)$ 是幺正算子. 证明：$S$ 的特征多项式中常数项的绝对值为 1.

**20** 设 $\mathbf{F}=\mathbf{C}$ 且 $V_1,\ldots,V_m$ 是 $V$ 的非零子空间且满足

$$V=V_1\oplus\cdots\oplus V_m.$$

设 $T\in\mathcal{L}(V)$ 且各 $V_k$ 在 $T$ 下不变. 对每个 $k$，令 $p_k$ 表示 $T|_{V_k}$ 的特征多项式. 证明：$T$ 的特征多项式等于 $p_1\cdots p_m$.

**21** 设 $p,q\in\mathcal{P}(\mathbf{C})$ 是具有相同零点的首一多项式且 $q$ 是 $p$ 的多项式倍. 证明: 存在 $T\in\mathcal{L}(\mathbf{C}^{\deg q})$ 使得 $T$ 的特征多项式是 $q$ 且 $T$ 的最小多项式是 $p$.

注　本习题表明，每个首一多项式都是某个算子的特征多项式.

**22** 设 $A$ 和 $B$ 是形如

$$A=\begin{pmatrix}A_1 & & 0\\ & \ddots & \\ 0 & & A_m\end{pmatrix},\quad B=\begin{pmatrix}B_1 & & 0\\ & \ddots & \\ 0 & & B_m\end{pmatrix}$$

的分块对角矩阵，其中各 $A_k$ 和 $B_k$（$k=1,\ldots,m$）是具有相同大小的方阵．证明，$AB$ 是形如

$$AB=\begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

的分块对角矩阵.

**23** 设 $\mathbf{F}=\mathbf{R}$．$T\in\mathcal{L}(V)$，且 $\lambda\in\mathbf{C}$.

(a) 证明 $u+\mathrm{i}v\in G(\lambda,T_{\mathbf{C}})$ 当且仅当 $u-\mathrm{i}v\in G(\overline{\lambda},T_{\mathbf{C}})$.

(b) 证明 $\lambda$ 作为 $T_{\mathbf{C}}$ 的特征值的重数等于 $\overline{\lambda}$ 作为 $T_{\mathbf{C}}$ 的特征值的重数.

(c) 利用 (b) 和有关重数之和的结果（8.25）证明如果 $\dim V$ 是奇数，则 $T_{\mathbf{C}}$ 有实特征值.

(d) 利用 (c) 和有关 $T_{\mathbf{C}}$ 的实特征值的结果（5A 节习题 17）证明如果 $\dim V$ 是奇数，则 $T$ 有特征值（从而得出了 5.34 的替代证法）.

注　复化 $T_{\mathbf{C}}$ 的定义见 3B 节习题 33.

# 8C 广义特征空间分解的推论

## 算子的平方根

回忆一下，算子 $T \in \mathcal{L}(V)$ 的平方根是满足 $R^2 = T$ 的算子 $R \in \mathcal{L}(V)$（见 7.36）. 每个复数都有平方根，但复向量空间上并非每个算子都有平方根. 例如，$\mathbf{C}^3$ 上定义为 $T(z_1, z_2, z_3) = (z_2, z_3, 0)$ 的算子没有平方根，本节习题 1 就要求你证明这一点. 很快我们将看到，这个无平方根的算子不可逆并非偶然. 我们首先证明恒等算子加上每个幂零算子都有平方根.

**8.39 恒等算子加上幂零算子有平方根**

设 $T \in \mathcal{L}(V)$ 是幂零的. 那么 $I + T$ 有平方根.

**证明** 考虑函数 $\sqrt{1+x}$ 的泰勒级数

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots. \tag{8.40}$$

我们并不需要求出系数的确切表达式，也无需担心无限和是否收敛，因为我们仅仅是将该式用作引子.

因为 $a_1 = \frac{1}{2}$，所以上述公式表明，当 $x$ 很小时，$1 + \frac{x}{2}$ 可很好地近似 $\sqrt{1+x}$.

因为 $T$ 是幂零的，所以对某正整数 $m$ 有 $T^m = 0$. 假设我们将式 (8.40) 中的 $x$ 换成 $T$、1 换成 $I$. 那么式子右端的无限和就变为有限和（因为对所有 $k \geq m$ 有 $T^k = 0$）. 于是我们猜想，$I + T$ 具有形如

$$I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}$$

的平方根. 作出猜想后，我们尝试找出 $a_1, a_2, \ldots, a_{m-1}$ 使得上述算子的平方等于 $I + T$. 则

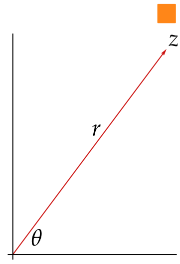
$$\begin{aligned}
&(I + a_1 T + a_2 T^2 + a_3 T^3 + \cdots + a_{m-1} T^{m-1})^2 \\
&= I + 2a_1 T + (2a_2 + a_1^2) T^2 + (2a_3 + 2a_1 a_2) T^3 + \cdots \\
&\quad + (2a_{m-1} + \text{含 } a_1, \ldots, a_{m-2} \text{ 的项}) T^{m-1}.
\end{aligned}$$

我们想让上式右侧等于 $I + T$. 因此取 $a_1$ 使得 $2a_1 = 1$（于是 $a_1 = \frac{1}{2}$）. 接下来，取 $a_2$ 使得 $2a_2 + a_1^2 = 0$（于是 $a_2 = -\frac{1}{8}$）. 然后选取 $a_3$ 使得上式中 $T^3$ 前系数为 0（于是 $a_3 = \frac{1}{16}$）. 对每个 $k = 4, \ldots, m-1$ 都按此方式求解下去，即在每步都求出使得上式右侧 $T^k$ 项的系数等于 0 的 $a_k$. 事实上我们不关心各 $a_k$ 的确切值，而只需要知道存在一组 $a_k$ 能使 $I + a_1 T + \cdots + a_{m-1} T^{m-1}$ 等于 $I + T$ 的平方根. ■

上述引理在实向量空间和复向量空间上都成立. 然而，下面的结论则只在复向量空间上成立. 例如，一维实向量空间 $\mathbf{R}$ 上"与 $-1$ 相乘"这个算子就没有平方根.

为证明下一条结果，我们需要知道每个复数 $z \in \mathbf{C}$ 都有属于 $\mathbf{C}$ 的平方根. 为此，将 $z$ 写成

$$z = r(\cos\theta + \mathrm{i}\sin\theta),$$



用极坐标来表达一个复数.

其中 $r$ 是复平面上从原点到 $z$ 点的线段长度，$\theta$ 是该线段与横轴正方向的夹角．那么

$$\sqrt{r}\left(\cos\frac{\theta}{2}+\mathrm{i}\sin\frac{\theta}{2}\right)$$

就是 $z$ 的平方根．你可通过证明上面这个复数的平方等于 $z$ 来验证这点．

**8.41　C 上，可逆算子具有平方根**

设 $V$ 是复向量空间，$T \in \mathcal{L}(V)$ 是可逆的．那么 $T$ 有平方根．

**证明**　令 $\lambda_1,\ldots,\lambda_m$ 是 $T$ 的所有互异特征值．对每个 $k$ 都存在幂零算子 $T_k \in \mathcal{L}\big(G(\lambda_k,T)\big)$ 使得 $T|_{G(\lambda_k,T)} = \lambda_k I + T_k$【见 8.22 (b)】．因为 $T$ 是可逆的，所以各 $\lambda_k$ 均不等于 0，因此我们可对每个 $k$ 写出

$$T|_{G(\lambda_k,T)} = \lambda_k\left(I+\frac{T_k}{\lambda_k}\right).$$

因为 $T_k/\lambda_k$ 是幂零的，所以 $I+T_k/\lambda_k$ 具有平方根（由 8.39）．将复数 $\lambda_k$ 的平方根和 $I+T_k/\lambda_k$ 的平方根相乘，我们就可得 $T|_{G(\lambda_k,T)}$ 的平方根 $R_k$．

由广义特征空间分解（8.22），向量 $v \in V$ 可被唯一写成下面形式：

$$v = u_1 + \cdots + u_m,$$

其中各 $u_k$ 属于 $G(\lambda_k,T)$．利用这个分解式，定义算子 $R \in \mathcal{L}(V)$ 为

$$Rv = R_1u_1 + \cdots + R_mu_m.$$

你应自行验证，这个算子 $R$ 是 $T$ 的平方根．证毕．■

你可通过效仿本小节中使用的技巧，来证明如果 $V$ 是复向量空间且 $T \in \mathcal{L}(V)$ 可逆，那么 $T$ 具有任意正整数次方根．

## 若当型

我们知道，如果 $V$ 是复向量空间，那么对每个 $T \in \mathcal{L}(V)$，都存在 $V$ 的一个基，使得 $T$ 关于该基有上三角矩阵（见 8.37），这已经是个很好的结论．在本小节中我们会得到更好的结论——存在 $V$ 的一个基，使得 $T$ 关于该基的矩阵除了对角线及紧挨在对角线正上方的元素之外，其他各元素都是 0．

我们先来看两个有关幂零算子的例子．

**8.42　例：具有很好的矩阵的幂零算子**

设 $T$ 是 $\mathbf{C}^4$ 上定义为

$$T(z_1,z_2,z_3,z_4) = (0,z_1,z_2,z_3)$$

的算子．那么 $T^4=0$，从而 $T$ 是幂零的．若 $v=(1,0,0,0)$，那么 $T^3v, T^2v, Tv, v$ 是 $\mathbf{C}^4$ 的基．$T$ 关于该基的矩阵是

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

幂零算子的下个例子比上例复杂些.

**8.43 例：具有稍复杂点的矩阵的幂零算子**

设 $T$ 是 $\mathbf{C}^6$ 上定义为

$$T(z_1,z_2,z_3,z_4,z_5,z_6)=(0,z_1,z_2,0,z_4,0)$$

的算子. 那么 $T^3=0$，从而 $T$ 是幂零的. 该算子不像上例中的幂零算子那样容易处理，因为并不存在向量 $v\in\mathbf{C}^6$ 使得 $T^5v,T^4v,T^3v,T^2v,Tv,v$ 是 $\mathbf{C}^6$ 的一个基. 然而，如果我们取 $v_1=(1,0,0,0,0,0),v_2=(0,0,0,1,0,0),v_3=(0,0,0,0,0,1)$，那么 $T^2v_1,Tv_1,v_1,Tv_2,v_2,v_3$ 是 $\mathbf{C}^6$ 的一个基. $T$ 关于该基的矩阵是

$$\left(\begin{array}{cccccc}\left(\begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix}\right. & & \left.\right) & 0 & 0 & 0\\ & & & 0&0&0\\ & & & 0&0&0\\ 0&0&0&\left(\begin{matrix}0&1\\0&0\end{matrix}\right) & & 0\\ 0&0&0& & &0\\ 0&0&0&0&0&(0)\end{array}\right).$$

我们将上面这个 $6\times6$ 矩阵分解出一块一块的内部矩阵，以便说明我们是怎么将其看成分块对角矩阵的：它包含一个 $3\times3$ 的块（其中紧挨在对角线正上方的元素是 1，其余元素均为 0），一个 $2\times2$ 的块（其中紧挨在对角线正上方的元素是 1，其余元素均为 0），以及一个包含 0 的 $1\times1$ 的块.

我们接下来的目标，就是证明每个幂零算子 $T\in\mathcal{L}(V)$ 都和上例中的算子具有类似性质. 具体而言，存在有限多个向量 $v_1,\ldots,v_n\in V$，使得 $V$ 有由形如 $T^jv_k$ 的向量构成的基，其中各 $k$ 从 1 取至 $n$ 而 $j$ 从使得 $T^{m_k}v_k\neq0$ 的最大非负整数 $m_k$ 取至 0. $T$ 关于该基的矩阵就类似于上例中的矩阵. 更具体地说，$T$ 关于该基有分块对角矩阵，其中每个块除了紧挨在对角线正上方的元素，其余各元素都等于 0.

在下面定义的各 $A_k$ 中，对角线上的元素都是 $T$ 的特征值 $\lambda_k$，紧挨在对角线正上方的元素都是 1，其余各元素都是 0（要理解为何每个 $\lambda_k$ 都是 $T$ 的特征值，见 5.41）. 各 $\lambda_k$ 不一定互异；同时，$A_k$ 也可能是仅包含 $T$ 的特征值的 $1\times1$ 矩阵（$\lambda_k$）. 若各 $\lambda_k$ 都是 0，那么下面定义就与上段所述幂零算子的性质相合（回忆一下，若 $T$ 是幂零的，那么 0 是 $T$ 唯一的特征值）.

**8.44 定义：若当基（Jordan basis）**

设 $T\in\mathcal{L}(V)$. 称 $V$ 的一个基是**若当基**，如果 $T$ 关于该基具有分块对角矩阵

$$\begin{pmatrix}A_1 & & 0\\ & \ddots & \\ 0 & & A_p\end{pmatrix},$$

转下页

其中各 $A_k$ 是形如

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}$$

的上三角矩阵.

证明“有限维复向量空间上每个算子都有若当基”的大多数过程，都包含在下面对于幂零算子这个特殊情形的证明中. 这个特殊情形在实向量空间和复向量空间上都成立.

**8.45 每个幂零算子都有若当基**

设 $T \in \mathcal{L}(V)$ 是幂零的. 那么 $V$ 中有一个基是 $T$ 的若当基.

**证明** 我们将对 $\dim V$ 用归纳法来证明这个结论. 首先，注意到 $\dim V = 1$ 时欲证结论成立（因为在此情形下，唯一的幂零算子是 0 算子）. 现在假设 $\dim V > 1$ 且欲证结论对于所有维数更小的向量空间都成立.

令 $m$ 是使得 $T^m = 0$ 的最小正整数. 于是存在 $u \in V$ 使得 $T^{m-1}u \neq 0$. 令

$$U = \operatorname{span}(u, Tu, \ldots, T^{m-1}u).$$

$u, Tu, \ldots, T^{m-1}u$ 是线性无关的（见 8A 节习题 2）. 若 $U = V$，那么将该组反过来写，就得到了 $T$ 的一个若当基，证明完成. 于是我们可设 $U \neq V$.

注意到 $U$ 在 $T$ 下不变. 由归纳假设知 $U$ 有一个基是 $T|_U$ 的若当基. 我们的证明策略是，找到 $V$ 的子空间 $W$，它同样在 $T$ 下不变且满足 $V = U \oplus W$. 再利用归纳假设可知，$W$ 有一个基是 $T|_W$ 的若当基. 将 $T|_U$ 和 $T|_W$ 的若当基合并就得到了 $T$ 的若当基.

令 $\varphi \in V'$ 满足 $\varphi(T^{m-1}u) \neq 0$. 令

$$W = \left\{ v \in V : \varphi(T^k v) = 0 \text{ 对每个} k = 0, \ldots, m-1 \text{都成立} \right\}.$$

那么 $W$ 是 $V$ 的在 $T$ 下不变的子空间【这是因为如果 $v \in W$，那么对 $k = 0, \ldots, m-1$ 有 $\varphi\left(T^k(Tv)\right) = 0$，其中 $k = m-1$ 的情形成立是由于 $T^m = 0$】. 下面我们证明 $V = U \oplus W$——根据上段的论述，这样即可完成证明.

为了说明 $U + W$ 是直和，设 $v \in U \cap W$ 且 $v \neq 0$. 因为 $v \in U$，所以存在 $c_0, \ldots, c_{m-1} \in \mathbf{F}$ 使得

$$v = c_0 u + c_1 Tu + \cdots + c_{m-1} T^{m-1} u.$$

令 $j$ 是使得 $c_j \neq 0$ 的最小下标. 将 $T^{m-j-1}$ 作用于上式两侧，可得

$$T^{m-j-1} v = c_j T^{m-1} u,$$

这里我们利用了等式 $T^m = 0$. 现将 $\varphi$ 作用于上式两端，可得

$$\varphi(T^{m-j-1} v) = c_j \varphi(T^{m-1} u) \neq 0.$$

上式表明 $v \notin W$. 因此我们证明了 $U \cap W = \{0\}$，这就表明 $U + W$ 是直和（见 1.46）.

为了证明 $U \oplus W = V$，定义 $S: V \to \mathbf{F}^m$ 为

$$Sv = \big(\varphi(v), \varphi(Tv), \ldots, \varphi(T^{m-1}v)\big).$$

于是 $\operatorname{null} S = W$．因此

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m,$$

其中第二个等号源于线性映射基本定理（3.21）．利用上述不等式，可得

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

从而 $U \oplus W = V$（由 2.39），这就完成了证明． ■

现在，利用广义特征空间分解可将上述结论推广至非幂零的算子．该推广只有在复向量空间上才可进行．

卡米耶・若当（Camille Jordan，1838-1922）于 1870 年发表了 8.46 的证明．

**8.46 若当型（Jordan form）**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$．那么 $V$ 有一个基是 $T$ 的若当基．

**证明** 令 $\lambda_1, \ldots, \lambda_m$ 是 $T$ 的所有互异特征值．由广义特征空间分解得

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

且各 $(T - \lambda_k I)|_{G(\lambda_k,T)}$ 都是幂零的（见 8.22）．于是 8.45 就表明，每个 $G(\lambda_k, T)$ 中都有某基是 $(T - \lambda_k I)|_{G(\lambda_k,T)}$ 的若当基．将这些基合并，即可得到 $V$ 的一个基，且它是 $T$ 的若当基． ■

## 习题 8C

**1** 设 $T \in \mathcal{L}(\mathbf{C}^3)$ 是定义为 $T(z_1, z_2, z_3) = (z_2, z_3, 0)$ 的算子．证明：$T$ 没有平方根．

**2** 定义 $T \in \mathcal{L}(\mathbf{F}^5)$ 为 $T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0)$．

(a) 证明：$T$ 是幂零的．

(b) 求出 $I + T$ 的一个平方根．

**3** 设 $V$ 是复向量空间．证明 $V$ 上的每个可逆算子都有立方根．

**4** 设 $V$ 是实向量空间．证明：当且仅当 $\dim V$ 为偶数，$V$ 上的算子 $-I$ 才有平方根．

**5** 设 $T \in \mathcal{L}(\mathbf{C}^2)$ 是定义为 $T(w, z) = (-w - z, 9w + 5z)$ 的算子．求出 $T$ 的一个若当基．

**6** 求出 $\mathcal{P}_4(\mathbf{R})$ 的一个基，它是 $\mathcal{P}_4(\mathbf{R})$ 上定义为 $Dp = p'$ 的微分算子的若当基．

**7** 设 $T \in \mathcal{L}(V)$ 是幂零的，$v_1, \ldots, v_n$ 是 $T$ 的若当基．证明：$T$ 的最小多项式是 $z^{m+1}$，其中 $m$ 是 $T$ 关于 $v_1, \ldots, v_n$ 的矩阵中，紧挨在对角线上方的这条线上连续出现的 1 的最大数目．

**8** 设 $T \in \mathcal{L}(V)$，$V$ 的基 $v_1, \ldots, v_n$ 是 $T$ 的若当基．描述 $T^2$ 关于该基的矩阵．

**9** 设 $T \in \mathcal{L}(V)$ 是幂零的．解释为何存在 $v_1, \ldots, v_n \in V$ 和非负整数 $m_1, \ldots, m_n$ 使得下列 (a) 和 (b) 都成立．

(a) $T^{m_1}v_1, \ldots, Tv_1, v_1, \ldots, T^{m_n}v_n, \ldots, Tv_n, v_n$ 是 $V$ 的一个基．

(b) $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$．

**10** 设 $T \in \mathcal{L}(V)$，$V$ 的基 $v_1, \ldots, v_n$ 是 $T$ 的若当基．描述 $T$ 关于基 $v_n, \ldots, v_1$（通过颠倒各 $v$ 的顺序得到）的矩阵．

**11** 设 $T \in \mathcal{L}(V)$. 解释为何 $T$ 的每个若当基中的每个向量都是 $T$ 的广义特征向量.

**12** 设 $T \in \mathcal{L}(V)$ 可对角化. 证明：关于 $T$ 的每个若当基的 $\mathcal{M}(T)$ 都是对角矩阵.

**13** 设 $T \in \mathcal{L}(V)$ 是幂零的. 证明：若 $v_1, \dots, v_n$ 是 $V$ 中向量，$m_1, \dots, m_n$ 是非负整数且使得

$$T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n \text{ 是 } V \text{ 的一个基}$$

以及

$$T^{m_1+1}v_1 + \cdots + T^{m_n+1}v_n = 0$$

成立，那么 $T^{m_1}v_1, \dots, T^{m_n}v_n$ 是 $\operatorname{null} T$ 的一个基.

注　本习题表明 $n = \dim \operatorname{null} T$. 于是上面出现的正整数 $n$ 仅取决于 $T$，而不依赖于 $T$ 的若当基具体如何选取.

**14** 设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$，证明：$V$ 不能分解为两个在 $T$ 下不变的子空间的直和，当且仅当 $T$ 的最小多项式形如 $(z - \lambda)^{\dim V}$（其中 $\lambda \in \mathbf{C}$）.

# 8D 联系矩阵与算子的桥梁——迹

我们从定义方阵的迹开启本节，并先讨论它的一些性质，然后再利用方阵的迹这一概念来定义算子的迹.

**8.47 定义：矩阵的迹（trace of a matrix）**

设 $A$ 是各元素均属于 $\mathbf{F}$ 的方阵. $A$ 的**迹**，记为 $\operatorname{tr} A$，定义为 $A$ 对角线上各元素之和. ♣

**8.48 例：一个 $3\times 3$ 矩阵的迹**

设

$$A=\begin{pmatrix}3 & -1 & -2\\ 3 & 2 & -3\\ 1 & 2 & 0\end{pmatrix}.$$

$A$ 对角线上的元素是 3、2 和 0，以红色标出. 于是 $\operatorname{tr} A=3+2+0=5$.

矩阵乘法不满足交换律. 但是，下面结论表明，交换乘积项的顺序不影响乘积的迹.

**8.49 $AB$ 的迹等于 $BA$ 的迹**

设 $A$ 是 $m\times n$ 矩阵且 $B$ 是 $n\times m$ 矩阵. 那么

$$\operatorname{tr}(AB)=\operatorname{tr}(BA).$$

♡

**证明** 设

$$A=\begin{pmatrix}A_{1,1} & \cdots & A_{1,n}\\ \vdots & & \vdots\\ A_{m,1} & \cdots & A_{m,n}\end{pmatrix},\quad B=\begin{pmatrix}B_{1,1} & \cdots & B_{1,m}\\ \vdots & & \vdots\\ B_{n,1} & \cdots & B_{n,m}\end{pmatrix}.$$

$m\times m$ 矩阵 $AB$ 对角线上的第 $j$ 个元素等于 $\sum_{k=1}^{n}A_{j,k}B_{k,j}$. 于是

$$\begin{aligned}\operatorname{tr}(AB)&=\sum_{j=1}^{m}\sum_{k=1}^{n}A_{j,k}B_{k,j}\\ &=\sum_{k=1}^{n}\sum_{j=1}^{m}B_{k,j}A_{j,k}\\ &=\sum_{k=1}^{n}\bigl(n\times n \text{ 矩阵 } BA \text{ 对角线上的第 } k \text{ 个元素}\bigr)\\ &=\operatorname{tr}(BA),\end{aligned}$$

命题得证. ■

我们想要定义算子 $T\in\mathcal{L}(V)$ 的迹为 $T$ 关于 $V$ 的某个基的矩阵的迹. 然而，这定义不应依赖于基的选取. 下面的结论说明这是能做到的.

**8.50 算子的矩阵的迹不依赖于基的选取**

设 $T \in \mathcal{L}(V)$. 设 $u_1, \ldots, u_n$ 和 $v_1, \ldots, v_n$ 是 $V$ 的基. 那么

$$\operatorname{tr} \mathcal{M}\big(T, (u_1, \ldots, u_n)\big) = \operatorname{tr} \mathcal{M}\big(T, (v_1, \ldots, v_n)\big).$$

♡

**证明** 令 $A = \mathcal{M}\big(T, (u_1, \ldots, u_n)\big)$, $B = \mathcal{M}\big(T, (v_1, \ldots, v_n)\big)$. 换基公式告诉我们，存在可逆的 $n \times n$ 矩阵 $C$ 使得 $A = C^{-1}BC$（见 3.84）. 于是

$$\begin{aligned}
\operatorname{tr} A &= \operatorname{tr}\big((C^{-1}B)C\big) \\
&= \operatorname{tr}\big(C(C^{-1}B)\big) \\
&= \operatorname{tr}\big((CC^{-1})B\big) \\
&= \operatorname{tr} B,
\end{aligned}$$

其中第二行源于 8.49. ■

因为 8.50，所以下面定义是合理的.

**8.51 定义：算子的迹（trace of an operator）**

设 $T \in \mathcal{L}(V)$. $T$ 的**迹**，记作 $\operatorname{tr} T$，定义为

$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}\big(T, (v_1, \ldots, v_n)\big),$$

其中 $v_1, \ldots, v_n$ 是 $V$ 的任意一个基.

♣

设 $T \in \mathcal{L}(V)$ 且 $\lambda$ 是 $T$ 的一个特征值. 回忆一下，我们定义 $\lambda$ 的重数为广义特征空间 $G(\lambda, T)$ 的维数（见 8.23）；我们证明了该重数等于 $\dim \operatorname{null}(T - \lambda I)^{\dim V}$（见 8.20）；此外，若 $V$ 是复向量空间，那么 $T$ 的所有特征值的重数之和等于 $\dim V$（见 8.25）.

在下面结论中，“各特征值出现次数等于其重数”的特征值之和，意为如果 $\lambda_1, \ldots, \lambda_m$ 是 $T$ 的所有互异特征值，且重数分别为 $d_1, \ldots, d_m$，那么求和式为

$$d_1\lambda_1 + \cdots + d_m\lambda_m.$$

如果你更喜欢将所有特征值按一组未必互异的数（其中各特征值出现的次数等于其重数）来处理，那么可将所有特征值记为 $\lambda_1, \ldots, \lambda_n$（其中 $n$ 等于 $\dim V$），这样求和式可写为

$$\lambda_1 + \cdots + \lambda_n.$$

**8.52 在复向量空间上，迹等于特征值之和**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 那么 $\operatorname{tr} T$ 等于 $T$ 的特征值之和，其中各特征值出现次数等于其重数.

♡

**证明** 存在 $V$ 的一个基使得 $T$ 关于该基有上三角矩阵，且对角线上各元素都是 $T$ 的特征值，并且各特征值出现次数等于其重数（见 8.37）. 于是由算子的迹的定义以及 8.50（该结论让我们得以自由选取基），可得 $\operatorname{tr} T$ 等于 $T$ 的特征值之和，其中各特征值出现次数等于其重数. ■

**8.53 例：$\mathbf{C}^3$ 上一个算子的迹**

定义 $T \in \mathcal{L}(\mathbf{C}^3)$ 为

$$T(z_1, z_2, z_3) = (3z_1 - z_2 - 2z_3, 3z_1 + 2z_2 - 3z_3, z_1 + 2z_2).$$

$T$ 关于 $\mathbf{C}^3$ 的标准基的矩阵是

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

求该矩阵对角线上各元素的和，我们可见 $\operatorname{tr} T = 5$.

$T$ 的特征值是 $1, 2+3\mathrm{i}, 2-3\mathrm{i}$，它们的重数都是 1（你可自行验证）. 这些特征值之和（其中各特征值出现次数等于其重数）是 $1 + (2+3\mathrm{i}) + (2-3\mathrm{i})$，等于 5，正如 8.52 所言.

迹和特征多项式具有紧密的联系. 假设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$，且 $\lambda_1, \ldots, \lambda_n$ 是 $T$ 的特征值（各特征值出现的次数等于其重数）. 那么由定义（见 8.26），$T$ 的特征多项式等于

$$(z - \lambda_1) \cdots (z - \lambda_n).$$

将上述多项式展开，我们就可将 $T$ 的特征多项式写成这个形式：

$$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n).$$

由上述表达式立刻可得出下面结论. 该结论也见于 9.65，但 9.65 无需假设 $\mathbf{F} = \mathbf{C}$.

**8.54 迹与特征多项式**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 令 $n = \dim V$. 那么 $\operatorname{tr} T$ 等于 $T$ 的特征多项式中 $z^{n-1}$ 项的系数的相反数.

下面结论给出了内积空间上算子的迹的一个很漂亮的公式.

**8.55 内积空间上的迹**

设 $V$ 是内积空间，$T \in \mathcal{L}(V)$，$e_1, \ldots, e_n$ 是 $V$ 的规范正交基. 则

$$\operatorname{tr} T = \langle T e_1, e_1 \rangle + \cdots + \langle T e_n, e_n \rangle.$$

**证明** 注意到 $\mathcal{M}(T, (e_1, \ldots, e_n))$ 第 $k$ 行第 $k$ 列中的元素等于 $\langle T e_k, e_k \rangle$【利用 6.30 (a)，取 $v = T e_k$】，即可得欲证的公式. ■

方阵的迹的代数性质可迁移至算子的迹，如下面结论所示.

**8.56 迹是线性的**

函数 $\operatorname{tr} : \mathcal{L}(V) \to \mathbf{F}$ 是 $\mathcal{L}(V)$ 上的线性泛函，且使

$$\operatorname{tr}(ST) = \operatorname{tr}(TS)$$

对所有 $S, T \in \mathcal{L}(V)$ 都成立.

**证明** 取定 $V$ 的一个基. 本证明中所有算子的矩阵都是关于该基的. 设 $S, T \in \mathcal{L}(V)$.

若 $\lambda \in \mathbf{F}$，那么

$$\operatorname{tr}(\lambda T) = \operatorname{tr} \mathcal{M}(\lambda T) = \operatorname{tr}\big(\lambda \mathcal{M}(T)\big) = \lambda \operatorname{tr} \mathcal{M}(T) = \lambda \operatorname{tr} T,$$

其中第一个和最后一个等号源于算子的迹的定义，第二个等号来自 3.38，而第三个等号由方阵的迹的定义可得.

同时，

$$\operatorname{tr}(S+T) = \operatorname{tr} \mathcal{M}(S+T) = \operatorname{tr}\big(\mathcal{M}(S) + \mathcal{M}(T)\big) = \operatorname{tr} \mathcal{M}(S) + \operatorname{tr} \mathcal{M}(T) = \operatorname{tr} S + \operatorname{tr} T,$$

其中第一个和最后一个等号源于算子的迹的定义，第二个等号来自 3.35，而第三个等号由方阵的迹的定义可得. 上述两段说明了 $\operatorname{tr}: \mathcal{L}(V) \to \mathbf{F}$ 是 $\mathcal{L}(V)$ 上的线性泛函.

此外，

$$\operatorname{tr}(ST) = \operatorname{tr} \mathcal{M}(ST) = \operatorname{tr}\big(\mathcal{M}(S)\mathcal{M}(T)\big) = \operatorname{tr}\big(\mathcal{M}(T)\mathcal{M}(S)\big) = \operatorname{tr} \mathcal{M}(TS) = \operatorname{tr}(TS),$$

其中第二个和第四个等号源于 3.43，关键的第三个等号则来自 8.49. ■

等式 $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ 和 $\operatorname{tr} I = \dim V$ 在 $\mathcal{L}(V)$ 上的线性泛函中唯一刻画了迹——见习题 10.

等式 $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ 引出了下面结论，该结论在无限维向量空间上不成立（见本节习题 13）. 然而，对 $S, T$ 和 $V$ 附加限制条件，即可将下面结论推广至无限维向量空间，推广后的结论在量子理论中有重要应用.

> 下面结论的表述中并不涉及迹，但是这个简短的证明用到了迹. 数学中，发生这种情况时，往往是有个好的定义做好了铺垫.

**8.57 恒等算子不等于 $ST$ 与 $TS$ 之差**

不存在使得 $ST - TS = I$ 成立的算子 $S, T \in \mathcal{L}(V)$. ♡

**证明** 设 $S, T \in \mathcal{L}(V)$. 那么

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

其中两个等号都源自 8.56. 而 $I$ 的迹等于 $\dim V$，肯定不是 0. 因为 $ST - TS$ 和 $I$ 有不同的迹，所以它们不可能相等. ■

## 习题 8D

**1** 设 $V$ 是内积空间且 $v, w \in V$. 定义算子 $T \in \mathcal{L}(V)$ 为 $Tu = \langle u, v\rangle w$. 求出 $\operatorname{tr} T$ 的表达式.

**2** 设 $P \in \mathcal{L}(V)$ 满足 $P^2 = P$. 证明

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

**3** 设 $T \in \mathcal{L}(V)$ 且 $T^5 = T$. 证明：$\operatorname{tr} T$ 的实部和虚部都是整数.

**4** 设 $V$ 是内积空间且 $T \in \mathcal{L}(V)$. 证明：

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

**5** 设 $V$ 是内积空间. 设 $T \in \mathcal{L}(V)$ 是正算子，且 $\operatorname{tr} T = 0$. 证明 $T = 0$.

**6** 设 $V$ 是内积空间且 $P, Q \in \mathcal{L}(V)$ 是正交投影. 证明 $\operatorname{tr}(PQ) \geq 0$.

**7** 设算子 $T \in \mathcal{L}(\mathbf{C}^3)$ 的矩阵为

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

某人（准确无误地）告诉你 $-48$ 和 $24$ 是 $T$ 的特征值．不用计算机也不动笔，求出 $T$ 的第三个特征值.

**8** 证明或给出反例：若 $S, T \in \mathcal{L}(V)$，那么 $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

**9** 设 $T \in \mathcal{L}(V)$ 使得 $\operatorname{tr}(ST) = 0$ 对所有 $S \in \mathcal{L}(V)$ 成立．证明 $T = 0$.

**10** 证明：迹是唯一使得

$$\tau(ST) = \tau(TS)$$

对所有 $S, T \in \mathcal{L}(V)$ 成立且满足 $\tau(I) = \dim V$ 的线性泛函 $\tau : \mathcal{L}(V) \to \mathbf{F}$.

提示 设 $v_1, \dots, v_n$ 是 $V$ 的基.对 $j, k \in \{1, \dots, n\}$，定义 $P_{j,k} \in \mathcal{L}(V)$ 为 $P_{j,k}(a_1 v_1 + \cdots + a_n v_n) = a_k v_j$．试证明

$$\tau(P_{j,k}) = \begin{cases} 1 & 若 j = k, \\ 0 & 若 j \neq k. \end{cases}$$

那么对于 $T \in \mathcal{L}(V)$，利用等式 $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ 可证明 $\tau(T) = \operatorname{tr} T$.

**11** 设 $V$ 和 $W$ 是内积空间，$T \in \mathcal{L}(V, W)$．证明：若 $e_1, \dots, e_n$ 是 $V$ 的规范正交基且 $f_1, \dots, f_m$ 是 $W$ 的规范正交基，那么

$$\operatorname{tr}(T^*T) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle Te_k, f_j \rangle|^2.$$

注 数 $\langle Te_k, f_j \rangle$ 是 $T$ 关于规范正交基 $e_1, \dots, e_n$ 和 $f_1, \dots, f_m$ 的矩阵中的各元素．这些数依赖于基的选取，但 $\operatorname{tr}(T^*T)$ 与基的选取无关．于是本题说明，那个矩阵各元素绝对值的平方和与规范正交基的取法无关.

**12** 设 $V$ 和 $W$ 是有限维内积空间.

(a) 证明：$\langle S, T \rangle = \operatorname{tr}(T^*S)$ 定义了 $\mathcal{L}(V, W)$ 上的一个内积.

(b) 设 $e_1, \dots, e_n$ 是 $V$ 的规范正交基且 $f_1, \dots, f_m$ 是 $W$ 的规范正交基．证明：(a) 中所得 $\mathcal{L}(V, W)$ 上的内积与 $\mathbf{F}^{mn}$ 上的标准内积相同，此处我们将 $\mathcal{L}(V, W)$ 中各元素与其（关于刚才提到的基的）矩阵等同起来看，进而与 $\mathbf{F}^{mn}$ 的元素等同起来看.

注 注意：对于线性映射 $T \in \mathcal{L}(V, W)$，定义如 7.86 的范数与源于上面 (a) 中内积的范数不一样．除非另行说明，总假定 $\|T\|$ 指代的是定义如 7.86 的范数．源于上面 (a) 中内积的范数被称为**弗罗贝尼乌斯范数**（**Frobenius norm**）或**希尔伯特-施密特范数**（**Hilbert-Schmidt norm**）.

**13** 求出算子 $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ 使得 $ST - TS = I$ 成立.

提示 在例 3.9 中的算子上作出适当的修改.

注 本题说明，为将 8.57 推广至无限维向量空间，需要对 $S$ 和 $T$ 附加前提条件.

# 第 9 章 多重线性代数和行列式

我们从研究向量空间上的双线性型和二次型开启本章. 接着我们转向对多重线性型的研究. 我们将证明，在 $n$ 维向量空间上，交错 $n$ 重线性型所构成的向量空间的维数为 1. 该结论将使我们得以给出简洁且不依赖于基的算子行列式定义.

借助交错多重线性型研究行列式，我们可直截了当地证明行列式的关键性质. 例如，我们将看到行列式是可乘的，意即对于同一向量空间上的所有算子 $S$ 和 $T$，有 $\det(ST) = (\det S)(\det T)$. 我们还将看到，当且仅当 $\det T \neq 0$ 时，$T$ 是可逆的. 还有一个重要结论指出，复向量空间上的算子的行列式等于该算子的特征值之积，其中各特征值出现次数等于其重数.

本章结尾处介绍了张量积.

**以下假设在本章中总是成立：**

- $\mathbf{F}$ 代表 $\mathbf{R}$ 或 $\mathbf{C}$.
- $V$ 和 $W$ 代表 $\mathbf{F}$ 上的非零有限维向量空间.

♣



图为哥廷根大学（University of Göttingen）的数学研究所（Mathematical Institute）. 这座建筑于 1930 年开放，当时埃米·诺特（Emmy Noether，1882-1935）已经在这所大学担任数学研究员和教师长达 15 年（其中前 8 年没有薪水）. 1933 年，诺特被纳粹政府解雇. 那时，诺特及其合作者已经创建了现代代数的不少基础理论，包括对线性代数发展作出贡献的抽象代数观点.

# 9A 双线性型和二次型

## 双线性型

$V$ 上的一个双线性型是一个从 $V \times V$ 到 $\mathbf{F}$ 的函数 $\beta(\cdot,\cdot)$，它在两个位置上分别是线性的，意即如果我们固定任意一个位置中的向量，那么就会得到另一个位置上的线性函数. 正式的定义如下.

**9.1 定义：双线性型（bilinear form）**

$V$ 上的一个**双线性型**是一个函数 $\beta : V \times V \to \mathbf{F}$，该函数满足，对于所有 $u \in V$，

$$v \mapsto \beta(v,u) \quad 与 \quad v \mapsto \beta(u,v)$$

都是 $V$ 上的线性泛函.

例如，若 $V$ 是实内积空间，那么将有序对 $(u,v) \in V \times V$ 对应到 $\langle u,v \rangle$ 的函数就是 $V$ 上的双线性型. 若 $V$ 是非零复内积空间，那么这个函数就不是双线性型，因为内积在第二个位置上不是线性的（第二个位置上的复数被提出时会变为其复共轭）.

回忆一下，上述定义中使用的术语**“线性泛函”**，意为映射至标量域 $\mathbf{F}$ 的线性函数. 因而，术语**“双线性泛函”**比**“双线性型”**与之更具有一致性. 不幸的是，后者已经成为标准用法了.

若 $\mathbf{F} = \mathbf{R}$，那么双线性型与内积的不同之处在于，内积需要具备对称性【意即对所有 $v, w \in V$，$\beta(v,w) = \beta(w,v)$】和正定性【意即对所有 $v \in V \backslash \{0\}$，$\beta(v,v) > 0$】，但是双线性型不需要具备这些性质.

**9.2 例：双线性型**

- 定义为
$$\beta\big((x_1,x_2,x_3),(y_1,y_2,y_3)\big) = x_1y_2 - 5x_2y_3 + 2x_3y_1$$
的函数 $\beta : \mathbf{F}^3 \times \mathbf{F}^3 \to \mathbf{F}$ 是 $\mathbf{F}^3$ 上的双线性型.
- 设 $A$ 是 $n \times n$ 矩阵，其中第 $j$ 行第 $k$ 列元素 $A_{j,k} \in \mathbf{F}$. 定义 $\mathbf{F}^n$ 上的双线性型 $\beta_A$ 为
$$\beta_A\big((x_1,\ldots,x_n),(y_1,\ldots,y_n)\big) = \sum_{k=1}^{n}\sum_{j=1}^{n} A_{j,k}x_jy_k.$$
第一个例子就是本例取 $n = 3$ 和
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{pmatrix}$$
时的特殊情况.
- 设 $V$ 是实内积空间且 $T \in \mathcal{L}(V)$. 那么定义为
$$\beta(u,v) = \langle u, Tv \rangle$$
的函数 $\beta : V \times V \to \mathbf{R}$ 是 $V$ 上的双线性型.

转下页

- 若 $n$ 是正整数，那么定义为
$$\beta(p,q)=p(2)\cdot q'(3)$$
的函数 $\beta:\mathcal{P}_n(\mathbf{R})\times\mathcal{P}_n(\mathbf{R})\to\mathbf{R}$ 是 $\mathcal{P}_n(\mathbf{R})$ 上的双线性型.
- 设 $\varphi,\tau\in V'$. 那么定义为
$$\beta(u,v)=\varphi(u)\cdot\tau(v)$$
的函数 $\beta:V\times V\to\mathbf{F}$ 是 $V$ 上的双线性型.
- 更一般地，设 $\varphi_1,\ldots,\varphi_n,\tau_1,\ldots,\tau_n\in V'$. 那么定义为
$$\beta(u,v)=\varphi_1(u)\cdot\tau_1(v)+\cdots+\varphi_n(u)\cdot\tau_n(v)$$
的函数 $\beta:V\times V\to\mathbf{F}$ 是 $V$ 上的双线性型.

$V$ 上的双线性型是从 $V\times V$ 到 $\mathbf{F}$ 的函数，$V\times V$ 是向量空间，由此生出一个问题：一个双线性型能否也是从 $V\times V$ 到 $\mathbf{F}$ 的线性映射？注意到，9.2 中的各双线性型，除了在某些特殊情况下等于零映射外，其余时候都不是线性映射. 习题 3 表明，$V$ 上的双线性型 $\beta$ 是 $V\times V$ 上的线性映射仅当 $\beta=0$.

**9.3 定义：$V^{(2)}$**

$V$ 上的双线性型构成的集合记为 $V^{(2)}$. ♣

在通常的函数加法与标量乘法运算下，$V^{(2)}$ 是向量空间.

对于 $n$ 维向量空间 $V$ 上的算子 $T$ 和 $V$ 的基 $e_1,\ldots,e_n$，我们用一个 $n\times n$ 矩阵来给出 $T$ 的信息. 现在我们对 $V$ 上的双线性型也采取同样做法.

**9.4 定义：双线性型的矩阵（matrix of a bilinear form）、$\mathcal{M}(\beta)$**

设 $\beta$ 是 $V$ 上的双线性型，$e_1,\ldots,e_n$ 是 $V$ 的基. $\beta$ 关于该基的**矩阵**是 $n\times n$ 矩阵 $\mathcal{M}(\beta)$，其中第 $j$ 行第 $k$ 列中的元素 $\mathcal{M}(\beta)_{j,k}$ 由下式给出：
$$\mathcal{M}(\beta)_{j,k}=\beta(e_j,e_k).$$
如果从上下文不能明确基 $e_1,\ldots,e_n$ 的选取，就用 $\mathcal{M}(\beta,(e_1,\ldots,e_n))$ 这个记号. ♣

回忆一下，$\mathbf{F}^{n,n}$ 代表由元素属于 $\mathbf{F}$ 的 $n\times n$ 矩阵构成的向量空间，且有 $\dim\mathbf{F}^{n,n}=n^2$（见 3.39 和 3.40）.

**9.5 $\dim V^{(2)}=(\dim V)^2$**

设 $e_1,\ldots,e_n$ 是 $V$ 的基. 那么映射 $\beta\mapsto\mathcal{M}(\beta)$ 是将 $V^{(2)}$ 映成 $\mathbf{F}^{n,n}$ 的同构. 此外，$\dim V^{(2)}=(\dim V)^2$. ♡

**证明** 映射 $\beta\mapsto\mathcal{M}(\beta)$ 显然是从 $V^{(2)}$ 到 $\mathbf{F}^{n,n}$ 的线性映射.

对于 $A\in\mathbf{F}^{n,n}$，定义 $V$ 上的双线性型 $\beta_A$ 为：对 $x_1,\ldots,x_n,y_1,\ldots,y_n\in\mathbf{F}$，
$$\beta_A(x_1e_1+\cdots+x_ne_n,y_1e_1+\cdots+y_ne_n)=\sum_{k=1}^{n}\sum_{j=1}^{n}A_{j,k}x_jy_k$$

（若 $V=\mathbf{F}^n$ 且 $e_1,\ldots,e_n$ 是 $\mathbf{F}^n$ 的标准基，那么该 $\beta_A$ 就与例 9.2 中第二例所给出的双线性型 $\beta_A$ 相同）.

从 $V^{(2)}$ 到 $\mathbf{F}^{n,n}$ 的线性映射 $\beta\mapsto\mathcal{M}(\beta)$ 和从 $\mathbf{F}^{n,n}$ 到 $V^{(2)}$ 的线性映射 $A\mapsto\beta_A$ 是彼此的逆，这是因为，对所有 $\beta\in V^{(2)}$ 皆有 $\beta_{\mathcal{M}(\beta)}=\beta$，且对所有 $A\in\mathbf{F}^{n,n}$ 皆有 $\mathcal{M}(\beta_A)=A$（你应自行验证）.

于是这两个映射都是同构，它们所联系的两个空间就有相同的维数. 因此 $\dim V^{(2)}=\dim\mathbf{F}^{n,n}=n^2=(\dim V)^2$. ■

回忆一下，$C^{\mathrm{t}}$ 表示矩阵 $C$ 的转置. 将 $C$ 的行与列互换可得到矩阵 $C^{\mathrm{t}}$.

**9.6 双线性型与算子的复合**

设 $\beta$ 是 $V$ 上的双线性型，$T\in\mathcal{L}(V)$. 定义 $V$ 上的双线性型 $\alpha$ 和 $\rho$ 为

$$\alpha(u,v)=\beta(u,Tv)\quad\text{与}\quad\rho(u,v)=\beta(Tu,v).$$

令 $e_1,\ldots,e_n$ 是 $V$ 的基. 那么

$$\mathcal{M}(\alpha)=\mathcal{M}(\beta)\mathcal{M}(T)\quad\text{且}\quad\mathcal{M}(\rho)=\mathcal{M}(T)^{\mathrm{t}}\mathcal{M}(\beta).$$

**证明** 若 $j,k\in\{1,\ldots,n\}$，则

$$\begin{aligned}\mathcal{M}(\alpha)_{j,k}&=\alpha(e_j,e_k)\\&=\beta(e_j,Te_k)\\&=\beta\Big(e_j,\sum_{m=1}^{n}\mathcal{M}(T)_{m,k}e_m\Big)\\&=\sum_{m=1}^{n}\beta(e_j,e_m)\mathcal{M}(T)_{m,k}\\&=\big(\mathcal{M}(\beta)\mathcal{M}(T)\big)_{j,k}.\end{aligned}$$

于是 $\mathcal{M}(\alpha)=\mathcal{M}(\beta)\mathcal{M}(T)$. 类似可证 $\mathcal{M}(\rho)=\mathcal{M}(T)^{\mathrm{t}}\mathcal{M}(\beta)$. ■

下面结论展示了当基改变时，双线性型的矩阵如何变化. 下面结论中的公式应与算子的矩阵的换基公式（见 3.84）比照看待. 两个公式是类似的，区别在于下面公式中含转置 $C^{\mathrm{t}}$，而算子的矩阵的换基公式中含逆 $C^{-1}$.

**9.7 换基公式（change-of-basis formula）**

设 $\beta\in V^{(2)}$. 设 $e_1,\ldots,e_n$ 和 $f_1,\ldots,f_n$ 是 $V$ 的基. 令

$$A=\mathcal{M}\big(\beta,(e_1,\ldots,e_n)\big)\quad\text{且}\quad B=\mathcal{M}\big(\beta,(f_1,\ldots,f_n)\big)$$

以及 $C=\mathcal{M}\big(I,(e_1,\ldots,e_n),(f_1,\ldots,f_n)\big)$. 那么

$$A=C^{\mathrm{t}}BC.$$

**证明** 线性映射引理（3.4）告诉我们，存在算子 $T\in\mathcal{L}(V)$ 使得 $Tf_k=e_k$ 对每个 $k=1,\ldots,n$ 成立. 由算子关于一基的矩阵的定义可得

$$\mathcal{M}\big(T,(f_1,\ldots,f_n)\big)=C.$$

定义 $V$ 上的双线性型 $\alpha,\rho$ 为

$$\alpha(u,v)=\beta(u,Tv) \quad 与 \quad \rho(u,v)=\alpha(Tu,v)=\beta(Tu,Tv).$$

那么对所有 $j,k\in\{1,\ldots,n\}$ 均有 $\beta(e_j,e_k)=\beta(Tf_j,Tf_k)=\rho(f_j,f_k)$. 于是

$$\begin{aligned}A&=\mathcal{M}\big(\rho,(f_1,\ldots,f_n)\big)\\&=C^{\mathrm{t}}\mathcal{M}\big(\alpha,(f_1,\ldots,f_n)\big)\\&=C^{\mathrm{t}}BC,\end{aligned}$$

其中第二、三行都源自 9.6. ■

---

**9.8 例：$\mathcal{P}_2(\mathbf{R})$ 上的一个双线性型的矩阵**

定义 $\mathcal{P}_2(\mathbf{R})$ 上的双线性型 $\beta$ 为 $\beta(p,q)=p(2)\cdot q'(3)$. 令

$$A=\mathcal{M}\big(\beta,(1,x-2,(x-3)^2)\big) \quad 及 \quad B=\mathcal{M}\big(\beta,(1,x,x^2)\big)$$

以及 $C=\mathcal{M}\big(I,(1,x-2,(x-3)^2),(1,x,x^2)\big)$. 那么

$$A=\begin{pmatrix}0&1&0\\0&0&0\\0&1&0\end{pmatrix} \quad 且 \quad B=\begin{pmatrix}0&1&6\\0&2&12\\0&4&24\end{pmatrix} \quad 且 \quad C=\begin{pmatrix}1&-2&9\\0&1&-6\\0&0&1\end{pmatrix}.$$

现由换基公式 9.7 可断言 $A=C^{\mathrm{t}}BC$，你可通过用上面这些矩阵作乘法来验证.

---

## 对称双线性型

**9.9 定义：对称双线性型（symmetric bilinear form）、$V^{(2)}_{\mathrm{sym}}$**

称双线性型 $\rho\in V^{(2)}$ **是对称的**，若

$$\rho(u,w)=\rho(w,u)$$

对所有 $u,w\in V$ 都成立. $V$ 上对称双线性型构成的集合记作 $V^{(2)}_{\mathrm{sym}}$. ♣

---

**9.10 例：对称双线性型**

- 若 $V$ 是实内积空间且 $\rho\in V^{(2)}$ 定义为

  $$\rho(u,w)=\langle u,w\rangle,$$

  那么 $\rho$ 是 $V$ 上的对称双线性型.
- 设 $V$ 是实内积空间且 $T\in\mathcal{L}(V)$. 定义 $\rho\in V^{(2)}$ 为

  $$\rho(u,w)=\langle u,Tw\rangle,$$

  那么当且仅当 $T$ 是自伴算子时，$\rho$ 是 $V$ 上的对称双线性型（上个例子就是 $T=I$ 时的特殊情况）.
- 定义 $\rho:\mathcal{L}(V)\times\mathcal{L}(V)\to\mathbf{F}$ 为

  $$\rho(S,T)=\operatorname{tr}(ST).$$

转下页

那么 $\rho$ 是 $\mathcal{L}(V)$ 上的对称双线性型，这是因为迹是 $\mathcal{L}(V)$ 上的线性泛函，且 $\operatorname{tr}(ST)=\operatorname{tr}(TS)$ 对所有 $S,T\in\mathcal{L}(V)$ 成立（参见 8.56）.

---

**9.11 定义：对称矩阵（symmetric matrix）**

若方阵 $A$ 与其转置相等，则称 $A$ 是**对称的**. ♣

$V$ 上的算子可能关于 $V$ 的某些基（但不是所有基）具有对称矩阵. 相比之下，下面结论表明，$V$ 上的双线性型要么关于 $V$ 的所有基都具有对称矩阵，要么关于 $V$ 的所有基都不具有对称矩阵.

**9.12 对称双线性型是可对角化的**

设 $\rho\in V^{(2)}$. 那么下面各命题等价.

(a) $\rho$ 是 $V$ 上的对称双线性型.

(b) $\mathcal{M}\big(\rho,(e_1,\dots,e_n)\big)$ 对 $V$ 的每个基 $e_1,\dots,e_n$ 都是对称矩阵.

(c) $\mathcal{M}\big(\rho,(e_1,\dots,e_n)\big)$ 对 $V$ 的某个基 $e_1,\dots,e_n$ 是对称矩阵.

(d) $\mathcal{M}\big(\rho,(e_1,\dots,e_n)\big)$ 对 $V$ 的某个基 $e_1,\dots,e_n$ 是对角矩阵. ♡

**证明** 先设 (a) 成立，则 $\rho$ 是对称双线性型. 设 $e_1,\dots,e_n$ 是 $V$ 的一个基且 $j,k\in\{1,\dots,n\}$. 那么因为 $\rho$ 是对称的，所以 $\rho(e_j,e_k)=\rho(e_k,e_j)$. 于是 $\mathcal{M}\big(\rho,(e_1,\dots,e_n)\big)$ 是对称矩阵，这说明 (a) 蕴涵 (b).

显然 (b) 蕴涵 (c).

现设 (c) 成立，并设 $e_1,\dots,e_n$ 是 $V$ 的一个基，使得 $\mathcal{M}\big(\rho,(e_1,\dots,e_n)\big)$ 是对称矩阵. 设 $u,w\in V$. 存在 $a_1,\dots,a_n,b_1,\dots,b_n\in\mathbf{F}$ 使得 $u=a_1e_1+\cdots+a_ne_n$ 与 $w=b_1e_1+\cdots+b_ne_n$ 成立. 则

$$\begin{aligned}\rho(u,w)&=\rho\Big(\sum_{j=1}^n a_je_j,\sum_{k=1}^n b_ke_k\Big)\\&=\sum_{j=1}^n\sum_{k=1}^n a_jb_k\rho(e_j,e_k)\\&=\sum_{j=1}^n\sum_{k=1}^n a_jb_k\rho(e_k,e_j)\\&=\rho\Big(\sum_{k=1}^n b_ke_k,\sum_{j=1}^n a_je_j\Big)\\&=\rho(w,u),\end{aligned}$$

其中，第三行成立是因为 $\mathcal{M}(\rho)$ 是对称矩阵. 上式表明 $\rho$ 是对称双线性型，证明了 (c) 蕴涵 (a).

现在，我们已证明了 (a)、(b)、(c) 等价. 因为每个对角矩阵都是对称的，所以 (d) 蕴涵 (c). 为完成证明，我们通过对 $n=\dim V$ 作归纳法来说明 (a) 蕴涵 (d).

若 $n=1$，那么因为每个 $1\times1$ 矩阵都是对角矩阵，所以 (a) 蕴涵 (d). 现设 $n>1$ 且蕴涵关系 (a) $\Longrightarrow$ (d) 对维数少 1 的情形成立. 设 (a) 成立，则 $\rho$ 是对称双线性型. 如果 $\rho=0$，那

么 $\rho$ 关于 $V$ 的每个基的矩阵都是零矩阵，即为对角矩阵．因此我们可设 $\rho \neq 0$，这意味着存在 $u, w \in V$ 使得 $\rho(u, w) \neq 0$．我们有

$$2\rho(u, w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w).$$

因为上式左侧是非零的，所以右侧三项不会都等于 0．因此存在 $v \in V$ 使得 $\rho(v, v) \neq 0$．

令 $U = \{u \in V : \rho(u, v) = 0\}$．于是 $U$ 是 $V$ 上的线性泛函 $u \mapsto \rho(u, v)$ 的零空间．因为 $v \notin U$，所以这个线性泛函不是零线性泛函．于是 $\dim U = n - 1$．由归纳假设，存在 $U$ 的基 $e_1, \ldots, e_{n-1}$ 使得对称双线性型 $\rho|_{U \times U}$ 关于该基有对角矩阵．

因为 $v \notin U$，所以 $e_1, \ldots, e_{n-1}, v$ 是 $V$ 的一个基．设 $k \in \{1, \ldots, n-1\}$．那么由 $U$ 的定义可知 $\rho(e_k, v) = 0$．因为 $\rho$ 是对称的，所以我们也有 $\rho(v, e_k) = 0$．于是 $\rho$ 关于 $e_1, \ldots, e_{n-1}, v$ 的矩阵是对角矩阵，这就完成了 (a) 蕴涵 (d) 的证明． ■

上述结论说明每个对称双线性型都关于某个基有对角矩阵．如果我们讨论的向量空间恰好又是实内积空间，那么下面结论说明每个对称双线性型都关于某个**规范正交**基有对角矩阵．注意，此处内积与双线性型无关．

**9.13 用规范正交基将对称双线性型对角化**

设 $V$ 是实内积空间且 $\rho$ 是 $V$ 上的对称双线性型．那么 $\rho$ 关于 $V$ 的某个规范正交基有对角矩阵．

**证明** 令 $f_1, \ldots, f_n$ 是 $V$ 的规范正交基．令 $B = \mathcal{M}(\rho, (f_1, \ldots, f_n))$．那么 $B$ 是对称矩阵（由 9.12）．令 $T \in \mathcal{L}(V)$ 是使得 $\mathcal{M}(T, (f_1, \ldots, f_n)) = B$ 的算子．于是 $T$ 是自伴的．

实谱定理（7.29）表明 $T$ 关于 $V$ 的某个规范正交基 $e_1, \ldots, e_n$ 具有对角矩阵．令 $C = \mathcal{M}(I, (e_1, \ldots, e_n), (f_1, \ldots, f_n))$．那么 $C^{-1}BC$ 是 $T$ 关于 $e_1, \ldots, e_n$ 的矩阵（由 3.84）．因此 $C^{-1}BC$ 是对角矩阵．则

$$\mathcal{M}(\rho, (e_1, \ldots, e_n)) = C^{\mathrm{t}}BC = C^{-1}BC,$$

其中第一个等号成立是由 9.7 所得，第二个等号成立是因为 $C$ 是各元素均为实数的幺正矩阵（这意味着 $C^{-1} = C^{\mathrm{t}}$，见 7.57）． ■

现在我们将注意力转向交错双线性型．在本章靠后部分，我们学习行列式时，交错双线性型会起重要作用．

**9.14 定义：交错双线性型（alternating bilinear form）、$V^{(2)}_{\mathrm{alt}}$**

称双线性型 $\alpha \in V^{(2)}$ 是**交错的**，若对于所有 $v \in V$ 有

$$\alpha(v, v) = 0.$$

$V$ 上交错双线性型所构成的集合记为 $V^{(2)}_{\mathrm{alt}}$．

**9.15 例：交错双线性型**

- 设 $n \geq 3$，$\alpha : \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ 定义为

$$\alpha((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1.$$

转下页

那么 $\alpha$ 是 $\mathbf{F}^n$ 上的交错双线性型.

- 设 $\varphi, \tau \in V'$，那么定义为

$$\alpha(u, w) = \varphi(u)\tau(w) - \varphi(w)\tau(u)$$

的 $V$ 上的双线性型 $\alpha$ 是交错的.

---

下面结论表明，一个双线性型是交错的，当且仅当交换两个输入量会使输出量乘以 $-1$.

**9.16 交错双线性型的刻画**

$V$ 上的双线性型 $\alpha$ 是交错的，当且仅当

$$\alpha(u, w) = -\alpha(w, u)$$

对所有 $u, w \in V$ 都成立.

**证明** 首先假设 $\alpha$ 是交错的. 若 $u, w \in V$，那么

$$\begin{aligned}0 &= \alpha(u+w, u+w) \\ &= \alpha(u, u) + \alpha(u, w) + \alpha(w, u) + \alpha(w, w) \\ &= \alpha(u, w) + \alpha(w, u).\end{aligned}$$

因此，$\alpha(u, w) = -\alpha(w, u)$.

为了证明另一方向的蕴涵关系，设 $\alpha(u, w) = -\alpha(w, u)$ 对所有 $u, w \in V$ 都成立. 那么 $\alpha(v, v) = -\alpha(v, v)$ 对所有 $v \in V$ 都成立，这表明 $\alpha(v, v) = 0$ 对所有 $v \in V$ 都成立. 因此 $\alpha$ 是交错的. ■

现在我们证明，$V$ 上双线性型所构成的向量空间，等于 $V$ 上对称双线性型和 $V$ 上交错双线性型各自构成的向量空间的直和.

**9.17 $V^{(2)} = V^{(2)}_{\mathrm{sym}} \oplus V^{(2)}_{\mathrm{alt}}$**

集合 $V^{(2)}_{\mathrm{sym}}$ 和 $V^{(2)}_{\mathrm{alt}}$ 都是 $V^{(2)}$ 的子空间，且有

$$V^{(2)} = V^{(2)}_{\mathrm{sym}} \oplus V^{(2)}_{\mathrm{alt}}.$$

**证明** 由对称双线性型的定义可得，$V$ 上任意两个对称双线性型的和仍是 $V$ 上的对称双线性型，并且 $V$ 上任意对称双线性型乘以任意标量仍得到 $V$ 上的对称双线性型. 于是 $V^{(2)}_{\mathrm{sym}}$ 是 $V^{(2)}$ 的子空间. 类似地，容易验证 $V^{(2)}_{\mathrm{alt}}$ 是 $V^{(2)}$ 的子空间.

接下来我们想证明 $V^{(2)} = V^{(2)}_{\mathrm{sym}} + V^{(2)}_{\mathrm{alt}}$. 为此，设 $\beta \in V^{(2)}$. 定义 $\rho, \alpha \in V^{(2)}$ 为

$$\rho(u, w) = \frac{\beta(u, w) + \beta(w, u)}{2} \quad 与 \quad \alpha(u, w) = \frac{\beta(u, w) - \beta(w, u)}{2}.$$

那么 $\rho \in V^{(2)}_{\mathrm{sym}}$ 且 $\alpha \in V^{(2)}_{\mathrm{alt}}$，且 $\beta = \rho + \alpha$. 于是 $V^{(2)} = V^{(2)}_{\mathrm{sym}} + V^{(2)}_{\mathrm{alt}}$.

最后，为证明我们讨论的两个子空间的交集是 $\{0\}$，设 $\beta \in V^{(2)}_{\mathrm{sym}} \cap V^{(2)}_{\mathrm{alt}}$. 则由 9.16，

$$\beta(u, w) = -\beta(w, u) = -\beta(u, w)$$

对所有 $u, w \in V$ 成立，这表明 $\beta = 0$. 于是由 1.46，$V^{(2)} = V^{(2)}_{\mathrm{sym}} \oplus V^{(2)}_{\mathrm{alt}}$. ■

## 二次型

**9.18 定义：关联于双线性型的二次型（quadratic form associated with a bilinear form）、$q_\beta$**

对于 $V$ 上的双线性型 $\beta$，定义函数 $q_\beta : V \to \mathbf{F}$ 为 $q_\beta(v) = \beta(v, v)$. 称函数 $q : V \to \mathbf{F}$ 是 $V$ 上的**二次型**，如果存在 $V$ 上的双线性型 $\beta$ 使得 $q = q_\beta$.

注意，如果 $\beta$ 是双线性型，那么 $q_\beta = 0$ 当且仅当 $\beta$ 是交错的.

**9.19 例：二次型**

设 $\beta$ 是 $\mathbf{R}^3$ 上的双线性型，定义为

$$\beta\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) = x_1 y_1 - 4x_1 y_2 + 8x_1 y_3 - 3x_3 y_3.$$

那么 $\mathbf{R}^3$ 上的二次型 $q_\beta$ 由下式给出：

$$q_\beta(x_1, x_2, x_3) = x_1^2 - 4x_1 x_2 + 8x_1 x_3 - 3x_3^2.$$

上例中的二次型具有 $\mathbf{F}^n$ 上的二次型的典型特点，如下面结论所示.

**9.20 $\mathbf{F}^n$ 上的二次型**

设 $n$ 是正整数，$q$ 是 $\mathbf{F}^n$ 到 $\mathbf{F}$ 的函数. 那么 $q$ 是 $\mathbf{F}^n$ 上的二次型，当且仅当存在数 $A_{j,k} \in \mathbf{F}$（$j, k \in \{1, \dots, n\}$）使得

$$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

对所有 $(x_1, \dots, x_n) \in \mathbf{F}^n$ 成立.

**证明** 先设 $q$ 是 $\mathbf{F}^n$ 上的二次型. 于是存在 $\mathbf{F}^n$ 上的双线性型 $\beta$ 使得 $q = q_\beta$. 令 $A$ 是 $\beta$ 关于 $\mathbf{F}^n$ 的标准基的矩阵. 那么对于所有 $(x_1, \dots, x_n) \in \mathbf{F}^n$，有

$$q(x_1, \dots, x_n) = \beta\big((x_1, \dots, x_n), (x_1, \dots, x_n)\big) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k.$$

此即我们欲得的等式.

反之，设存在数 $A_{j,k} \in \mathbf{F}$（$j, k \in \{1, \dots, n\}$）使得

$$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

对所有 $(x_1, \dots, x_n) \in \mathbf{F}^n$ 成立. 定义 $\mathbf{F}^n$ 上的双线性型 $\beta$ 为

$$\beta\big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k.$$

那么 $q = q_\beta$，故原命题得证. ■

尽管二次型的定义中用的是一个任意的双线性型，但下面结果中 (a) 和 (b) 的等价关系说明，每个二次型总可与一个**对称**双线性型相对应.

**9.21 二次型的刻画**

设 $q : V \to \mathbf{F}$ 是一个函数. 下面各命题等价.

(a) $q$ 是一个二次型.

(b) $V$ 上存在唯一的对称双线性型 $\rho$ 使得 $q = q_\rho$ 成立.

(c) $q(\lambda v) = \lambda^2 q(v)$ 对所有 $\lambda \in \mathbf{F}$ 和 $v \in V$ 成立，并且函数

$$(u, w) \mapsto q(u+w) - q(u) - q(w)$$

是 $V$ 上的对称双线性型.

(d) $q(2v) = 4q(v)$ 对所有 $v \in V$ 成立，并且函数

$$(u, w) \mapsto q(u+w) - q(u) - q(w)$$

是 $V$ 上的对称双线性型.

**证明** 首先假设 (a) 成立，则 $q$ 是二次型. 因此存在双线性型 $\beta$ 满足 $q = q_\beta$. 由 9.17，存在 $V$ 上的对称双线性型 $\rho$ 和 $V$ 上的交错双线性型 $\alpha$ 满足 $\beta = \rho + \alpha$. 则

$$q = q_\beta = q_\rho + q_\alpha = q_\rho.$$

若 $\rho' \in V^{(2)}_{\mathrm{sym}}$ 也满足 $q_{\rho'} = q$，那么 $q_{\rho'-\rho} = 0$，从而 $\rho' - \rho \in V^{(2)}_{\mathrm{sym}} \cap V^{(2)}_{\mathrm{alt}}$，这表明 $\rho' = \rho$（由 9.17）. 这就完成了 (a) 蕴涵 (b) 的证明.

现假设 (b) 成立，故存在 $V$ 上的对称双线性型 $\rho$ 使得 $q = q_\rho$. 若 $\lambda \in \mathbf{F}$ 且 $v \in V$，那么

$$q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda\rho(v, \lambda v) = \lambda^2\rho(v, v) = \lambda^2 q(v),$$

这说明 (c) 的前半部分成立.

若 $u, w \in V$，那么

$$q(u+w) - q(u) - q(w) = \rho(u+w, u+w) - \rho(u, u) - \rho(w, w) = 2\rho(u, w).$$

于是函数 $(u, w) \mapsto q(u+w) - q(u) - q(w)$ 等于 $2\rho$，即为 $V$ 上的对称双线性型，这就完成了 (b) 蕴涵 (c) 的证明.

显然 (c) 蕴涵 (d).

现假设 (d) 成立. 令 $\rho$ 是 $V$ 上的对称双线性型，定义为

$$\rho(u, w) = \frac{q(u+w) - q(u) - q(w)}{2}.$$

若 $v \in V$，那么

$$\rho(v, v) = \frac{q(2v) - q(v) - q(v)}{2} = \frac{4q(v) - 2q(v)}{2} = q(v).$$

从而 $q = q_\rho$，这就完成了 (d) 蕴涵 (a) 的证明. ■

**9.22 例：与二次型相关联的对称双线性型**

设 $q$ 是 $\mathbf{R}^3$ 上的二次型，由此式给出：

$$q(x_1,x_2,x_3)=x_1^2-4x_1x_2+8x_1x_3-3x_3^2.$$

例 9.19 给出了一个满足 $q=q_\beta$ 的 $\mathbf{R}^3$ 上的双线性型 $\beta$，但这个双线性型不是对称的. 然而，定义为

$$\rho\big((x_1,x_2,x_3),(y_1,y_2,y_3)\big)=x_1y_1-2x_1y_2-2x_2y_1+4x_1y_3+4x_3y_1-3x_3y_3$$

的 $\mathbf{R}^3$ 上的双线性型 $\rho$ 是对称的，且满足 $q=q_\rho$，这与 9.21 (b) 相吻合.

下面结果指出，对每个二次型，我们都可选取一个基，使得该二次型看起来像是各坐标平方的加权和，意即没有形如 $x_jx_k$（$j\neq k$）的交叉项.

**9.23 二次型的对角化**

设 $q$ 是 $V$ 上的二次型.

(a) 存在 $V$ 的基 $e_1,\ldots,e_n$ 和 $\lambda_1,\ldots,\lambda_n\in\mathbf{F}$，使得

$$q(x_1e_1+\cdots+x_ne_n)=\lambda_1x_1^2+\cdots+\lambda_nx_n^2$$

对所有 $x_1,\ldots,x_n\in\mathbf{F}$ 成立.

(b) 若 $\mathbf{F}=\mathbf{R}$ 且 $V$ 是内积空间，那么 (a) 中的基可取为 $V$ 的规范正交基.

**证明**

(a) 存在 $V$ 上的对称双线性型 $\rho$ 使得 $q=q_\rho$ 成立（由 9.21）. 则存在 $V$ 的基 $e_1,\ldots,e_n$ 使得 $\mathcal{M}(\rho,(e_1,\ldots,e_n))$ 是对角矩阵（由 9.12）. 令 $\lambda_1,\ldots,\lambda_n$ 代表该矩阵对角线上各元素. 那么

$$\rho(e_j,e_k)=\begin{cases}\lambda_j & 若\, j=k,\\ 0 & 若\, j\neq k\end{cases}$$

对于所有 $j,k\in\{1,\ldots,n\}$ 成立. 若 $x_1,\ldots,x_n\in\mathbf{F}$，那么

$$\begin{aligned}q(x_1e_1+\cdots+x_ne_n)&=\rho(x_1e_1+\cdots+x_ne_n,x_1e_1+\cdots+x_ne_n)\\&=\sum_{k=1}^{n}\sum_{j=1}^{n}x_jx_k\rho(e_j,e_k)\\&=\lambda_1x_1^2+\cdots+\lambda_nx_n^2,\end{aligned}$$

则原命题得证.

(b) 设 $\mathbf{F}=\mathbf{R}$ 且 $V$ 是内积空间. 那么 9.13 告诉我们，(a) 中的基可取为 $V$ 的规范正交基. ■

## 习题 9A

**1** 证明：如果 $\beta$ 是 $\mathbf{F}$ 上的双线性型，那么存在 $c\in\mathbf{F}$ 使得

$$\beta(x,y)=cxy$$

对所有 $x,y\in\mathbf{F}$ 成立.

**2** 令 $n = \dim V$. 设 $\beta$ 是 $V$ 上的双线性型. 证明：存在 $\varphi_1, \ldots, \varphi_n, \tau_1, \ldots, \tau_n \in V'$ 使得

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

对所有 $u, v \in V$ 成立.

注 本题表明：如果 $n = \dim V$，那么 $V$ 上的每个双线性型，都具有例 9.2 中最后一点给出的形式.

**3** 设 $\beta : V \times V \to \mathbf{F}$ 既是 $V$ 上的双线性型也是 $V \times V$ 上的线性泛函. 证明：$\beta = 0$.

**4** 设 $V$ 是实内积空间，$\beta$ 是 $V$ 上的双线性型. 证明：存在唯一的算子 $T \in \mathcal{L}(V)$ 使得

$$\beta(u, v) = \langle u, Tv \rangle$$

对所有 $u, v \in V$ 成立.

注 本题指出，如果 $V$ 是实内积空间，那么 $V$ 上的每个双线性型，都具有 9.2 中第三点给出的形式.

**5** 设 $\beta$ 是实内积空间 $V$ 上的双线性型，$T$ 是 $V$ 上唯一使得 $\beta(u, v) = \langle u, Tv \rangle$ 对所有 $u, v \in V$ 成立的算子（见第 4 题）. 证明：$\beta$ 是 $V$ 上的内积，当且仅当 $T$ 是 $V$ 上的可逆正算子.

**6** 证明或给出一反例：如果 $\rho$ 是 $V$ 上的对称双线性型，那么

$$\{v \in V : \rho(v, v) = 0\}$$

是 $V$ 的子空间.

**7** 解释为什么：如果去掉 $\mathbf{F} = \mathbf{R}$ 这一假设，9.13（在实内积空间上用规范正交基将对称双线性型对角化）的证明就不成立了.

**8** 求 $\dim V^{(2)}_{\mathrm{sym}}$ 和 $\dim V^{(2)}_{\mathrm{alt}}$ 的表达式（用 $\dim V$ 表示）.

**9** 设 $n$ 是正整数，$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. 定义 $\alpha : V \times V \to \mathbf{R}$ 为

$$\alpha(p, q) = \int_0^1 pq'.$$

证明：$\alpha$ 是 $V$ 上的交错双线性型.

**10** 设 $n$ 是正整数，且

$$V = \big\{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ 且 } p'(0) = p'(1)\big\}.$$

定义 $\rho : V \times V \to \mathbf{R}$ 为

$$\rho(p, q) = \int_0^1 pq''.$$

证明：$\rho$ 是 $V$ 上的对称双线性型.

# 9B 交错多重线性型

## 多重线性型

**9.24 定义：$V^m$**

对于正整数 $m$，定义 $V^m$ 为

$$V^m = \underbrace{V \times \cdots \times V}_{m\text{个}V}.$$

现在我们可将上节中所讨论的双线性型推广得到 $m$ 重线性型的定义.

**9.25 定义：$m$ 重线性型（$m$-linear form）、$V^{(m)}$、多重线性型（multilinear form）**

- 对于正整数 $m$，$V$ 上的 $m$ **重线性型**是一个函数 $\beta : V^m \to \mathbf{F}$，它在每个位置都是线性的（当其他位置的值固定时）. 这意味着，对每个 $k \in \{1, \dots, m\}$ 和所有 $u_1, \dots, u_m \in V$，函数

$$v \mapsto \beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$$

  是 $V$ 到 $\mathbf{F}$ 的线性映射.
- $V$ 上 $m$ 重线性型所构成的集合记作 $V^{(m)}$.
- 若函数 $\beta$ 是 $V$ 上的 $m$ 重线性型（$m$ 为正整数），则称该函数为一个**多重线性型**.

在上述定义中，表达式 $\beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$ 在 $k = 1$ 时表示 $\beta(v, u_2, \dots, u_m)$，在 $k = m$ 时表示 $\beta(u_1, \dots, u_{m-1}, v)$.

$V$ 上的 1 重线性型是 $V$ 上的线性泛函. $V$ 上的 2 重线性型是 $V$ 上的双线性型. 你可验证，带有通常的函数加法和标量乘法运算的 $V^{(m)}$ 是向量空间.

**9.26 例：$m$ 重线性型**

- 设 $\alpha, \rho \in V^{(2)}$. 定义函数 $\beta : V^4 \to \mathbf{F}$ 为

$$\beta(v_1, v_2, v_3, v_4) = \alpha(v_1, v_2)\rho(v_3, v_4).$$

  那么 $\beta \in V^{(4)}$.
- 定义函数 $\beta : (\mathcal{L}(V))^m \to \mathbf{F}$ 为

$$\beta(T_1, \dots, T_m) = \operatorname{tr}(T_1 \cdots T_m).$$

  那么 $\beta$ 是 $\mathcal{L}(V)$ 上的 $m$ 重线性型.

现在，我们定义交错多重线性型. 在我们向行列式定义进发之路上，它将发挥重要作用.

**9.27 定义：交错型（alternating forms）、$V_{\mathrm{alt}}^{(m)}$**

设 $m$ 是正整数.

- 对于 $V$ 上的 $m$ 重线性型 $\alpha$，如果只要 $V$ 中向量组 $v_1,\ldots,v_m$ 满足对于某两个不同的 $j,k\in\{1,\ldots,m\}$ 有 $v_j=v_k$，就有 $\alpha(v_1,\ldots,v_m)=0$，则称 $\alpha$ 是**交错的**.
- $V_{\mathrm{alt}}^{(m)}=\{\alpha\in V^{(m)}:\alpha \text{ 是 } V \text{ 上的交错 } m \text{ 重线性型}\}$.

你应自行验证 $V_{\mathrm{alt}}^{(m)}$ 是 $V^{(m)}$ 的子空间. 交错 2 重线性型的例子见例 9.15. 交错 3 重线性型的例子见习题 2.

下面结论告诉我们，如果将一个线性相关组作为交错多重线性型的输入，那么所得输出等于 0.

**9.28 交错多重线性型和线性相关性**

设 $m$ 是正整数，$\alpha$ 是 $V$ 上的交错 $m$ 重线性型. 若 $v_1,\ldots,v_m$ 是 $V$ 中的线性相关组，那么

$$\alpha(v_1,\ldots,v_m)=0.$$

**证明** 设 $v_1,\ldots,v_m$ 是 $V$ 中的线性相关组. 由线性相关性引理（2.19），某个 $v_k$ 是 $v_1,\ldots,v_{k-1}$ 的线性组合. 于是存在 $b_1,\ldots,b_{k-1}$ 使得 $v_k=b_1v_1+\cdots+b_{k-1}v_{k-1}$. 则

$$\begin{aligned}\alpha(v_1,\ldots,v_m)&=\alpha\Big(v_1,\ldots,v_{k-1},\sum_{j=1}^{k-1}b_jv_j,v_{k+1},\ldots,v_m\Big)\\&=\sum_{j=1}^{k-1}b_j\alpha(v_1,\ldots,v_{k-1},v_j,v_{k+1},\ldots,v_m)\\&=0.\end{aligned}$$
■

下面结果是说，如果 $m>\dim V$，那么除定义在 $V^m$ 上的常值函数 0 以外，$V$ 上不存在交错 $m$ 重线性型.

**9.29 对 $m>\dim V$，不存在非零交错 $m$ 重线性型**

设 $m>\dim V$. 那么 0 是 $V$ 上唯一的交错 $m$ 重线性型.

**证明** 设 $\alpha$ 是 $V$ 上的交错 $m$ 重线性型，且 $v_1,\ldots,v_m\in V$. 因为 $m>\dim V$，所以该组不是线性无关的（由 2.22）. 于是 9.28 就表明 $\alpha(v_1,\ldots,v_m)=0$. 因此 $\alpha$ 是从 $V^m$ 到 $\mathbf{F}$ 的零函数. ■

## 交错多重线性型和排列

**9.30 交换交错多重线性型的输入向量**

设 $m$ 是正整数，$\alpha$ 是 $V$ 上的交错 $m$ 重线性型，且 $v_1,\ldots,v_m$ 是 $V$ 中的向量组. 那么交换 $\alpha(v_1,\ldots,v_m)$ 中任意两个位置上的向量会使 $\alpha$ 的值变为原来的 $-1$ 倍.

**证明** 在前两个位置上都取 $v_1 + v_2$ 可得

$$0 = \alpha(v_1 + v_2, v_1 + v_2, v_3, \ldots, v_m).$$

利用 $\alpha$ 的多重线性，可如 9.16 的证明一般将上式右侧展开，得到

$$\alpha(v_2, v_1, v_3, \ldots, v_m) = -\alpha(v_1, v_2, v_3, \ldots, v_m).$$

类似地，交换 $\alpha(v_1, \ldots, v_m)$ 中任意两个位置上的向量都使 $\alpha$ 的值变为原来的 $-1$ 倍. ■

下面观察多次交换各输入向量会发生什么. 设 $\alpha$ 是 $V$ 上的交错 3 重线性型且 $v_1, v_2, v_3 \in V$. 为用 $\alpha(v_1, v_2, v_3)$ 表示 $\alpha(v_3, v_1, v_2)$，先将 $\alpha(v_3, v_1, v_2)$ 第一个位置和第三个位置中的项互换，得到 $\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3)$，再将此式右侧的第一个位置和第二个位置中的项互换，可得

$$\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3) = \alpha(v_1, v_2, v_3).$$

更一般地，我们可发现如果我们做奇数次交换，那么 $\alpha$ 的值变为原来的 $-1$ 倍；而如果我们做偶数次交换，那么 $\alpha$ 的值不变.

为了处理任意多次的交换，我们需要一些有关排列的知识.

**9.31 定义：排列（permutation）、perm $m$**

设 $m$ 是正整数.

- $(1, \ldots, m)$ 的一个**排列**是不重不漏地包含 $1, \ldots, m$ 的组 $(j_1, \ldots, j_m)$.
- $(1, \ldots, m)$ 的所有排列所构成的集合记为 perm $m$.

♣

例如，$(2, 3, 4, 5, 1) \in \operatorname{perm} 5$. 你应将 perm $m$ 的一个元素看成对于头 $m$ 个正整数的重排.

将一个排列 $(j_1, \ldots, j_m)$ 变回标准排列 $(1, \ldots, m)$ 所用交换的次数与交换的方法是有关的. 下面定义的好处在于，给所有排列都赋予了一个定义完善的符号.

**9.32 定义：排列的符号（sign of a permutation）**

排列 $(j_1, \ldots, j_m)$ 的**符号**定义为

$$\operatorname{sign}(j_1, \ldots, j_m) = (-1)^N,$$

其中 $N$ 是所有整数对 $(k, \ell)$（$1 \le k < \ell \le m$）中，满足 $k$ 在组 $(j_1, \ldots, j_m)$ 中排在 $\ell$ 之后的数目. [1]

♣

因此，若一排列中有偶数个不合自然顺序之处，则其符号等于 1；而若一排列中有奇数个不合自然顺序之处，则其符号等于 $-1$.

**9.33 例：符号**

- $(1, \ldots, m)$（保持自然顺序不变）的符号是 1.
- 在 $(2, 1, 3, 4)$ 这个组中，唯一满足 $k$ 排在 $\ell$ 之后的整数对 $(k, \ell)$（$k < \ell$）是 $(1, 2)$. 因此排列 $(2, 1, 3, 4)$ 的符号是 $-1$.
- 在排列 $(2, 3, \ldots, m, 1)$ 中，满足 $k$ 排在 $\ell$ 之后的整数对 $(k, \ell)$（$k < \ell$）仅有 $(1, 2), (1, 3), \ldots, (1, m)$. 因为这些整数对共有 $m - 1$ 个，所以该排列的符号等于 $(-1)^{m-1}$.

[1]此处的 $N$ 也称为逆序数.

**9.34 交换排列中的两项**

交换排列中的两项会将排列的符号乘以 $-1$.

**证明** 假设我们有两个排列，其中第二个排列是通过交换第一个排列中的两项得到的. 若被交换的两项在第一个排列中按其自然顺序排列，那么它们在第二个排列中就不按其自然顺序排列；反之亦然. 于是到目前为止，不符合自然顺序的数对个数净变化量为 1 或 $-1$（都是奇数）.

考虑两个被交换的项之间的各项. 若某个中间项最初与两个被交换的项都符合自然顺序，那么它现在与两个被交换的项就都不符合自然顺序. 类似地，若某个中间项最初与两个被交换的项都不符合自然顺序，那么它现在与两个被交换的项就都符合自然顺序. 若某个中间项最初恰与两个被交换的项之一符合自然顺序，那么现在仍是如此. 于是，对于所有仅包含一个特定中间项的数对，不符合自然顺序的数对个数净变化量为 2、0 或 $-2$（都是偶数）.[2]

对于其他所有数对，它们是否符合自然顺序的情况都不改变. 于是，不符合自然顺序的数对的个数变化总量是奇数. 因此第二个排列的符号就等于第一个排列的符号乘以 $-1$. ■

**9.35 排列和交错多重线性型**

设 $m$ 是正整数，且 $\alpha \in V_{\mathrm{alt}}^{(m)}$. 那么

$$\alpha(v_{j_1}, \dots, v_{j_m}) = \big(\operatorname{sign}(j_1, \dots, j_m)\big)\alpha(v_1, \dots, v_m)$$

对 $V$ 中每个向量组 $v_1, \dots, v_m$ 以及所有 $(j_1, \dots, j_m) \in \operatorname{perm} m$ 成立.

**证明** 设 $v_1, \dots, v_m \in V$ 且 $(j_1, \dots, j_m) \in \operatorname{perm} m$. 我们可通过对 $\alpha$ 不同位置上的向量进行一系列交换操作，使其下标构成的排列由 $(j_1, \dots, j_m)$ 变为 $(1, \dots, m)$. 每次交换都将 $\alpha$ 的值乘以 $-1$（由 9.30），也会将所得排列的符号乘以 $-1$（由 9.34）. 在一定数量的交换之后，我们得到了排列 $1, \dots, m$，它的符号是 1. 于是若 $\operatorname{sign}(j_1, \dots, j_m) = 1$，那么 $\alpha$ 的值改变符号偶数次；若 $\operatorname{sign}(j_1, \dots, j_m) = -1$，那么 $\alpha$ 的值改变符号奇数次，此即为我们欲证的结论. ■

这样一来，利用排列，我们就可以很自然地得到下面这个漂亮的公式，它用于表达 $n$ 维向量空间上的 $n$ 重线性型.

**9.36 $V$ 上交错 $(\dim V)$ 重线性型的公式**

令 $n = \dim V$. 设 $e_1, \dots, e_n$ 是 $V$ 的一个基且 $v_1, \dots, v_n \in V$. 对每个 $k \in \{1, \dots, n\}$，令 $b_{1,k}, \dots, b_{n,k} \in \mathbf{F}$ 满足

$$v_k = \sum_{j=1}^{n} b_{j,k} e_j.$$

那么

$$\alpha(v_1, \dots, v_n) = \alpha(e_1, \dots, e_n) \sum_{(j_1, \dots, j_n) \in \operatorname{perm} n} \big(\operatorname{sign}(j_1, \dots, j_n)\big) b_{j_1,1} \cdots b_{j_n,n}$$

对于 $V$ 上每个交错 $n$ 重线性型都成立.

[2] 如 9.33 的最后一例中，若交换排列中的 2 和 $m$，则仅包含一个中间项 3 的数对为 $(1,3), (2,3), (3,m)$，原先其中不符合自然顺序的是 $(1,3)$，交换后不符合自然顺序的数对是 $(1,3), (2,3), (3,m)$，净增加量为 2. 对于其他中间项 $4, \dots, m-1$，也是同样道理.

**证明** 设 $\alpha$ 是 $V$ 上的交错 $n$ 重线性型. 那么

$$\begin{aligned}\alpha(v_1,\dots,v_n) &= \alpha\left(\sum_{j_1=1}^{n} b_{j_1,1}e_{j_1},\dots,\sum_{j_n=1}^{n} b_{j_n,n}e_{j_n}\right)\\ &= \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n} b_{j_1,1}\cdots b_{j_n,n}\alpha(e_{j_1},\dots,e_{j_n})\\ &= \sum_{(j_1,\dots,j_n)\in\operatorname{perm} n} b_{j_1,1}\cdots b_{j_n,n}\alpha(e_{j_1},\dots,e_{j_n})\\ &= \alpha(e_1,\dots,e_n)\sum_{(j_1,\dots,j_n)\in\operatorname{perm} n}\big(\operatorname{sign}(j_1,\dots,j_n)\big)b_{j_1,1}\dots b_{j_n,n},\end{aligned}$$

其中第三行成立是因为如果 $j_1,\dots,j_n$ 不是互异的整数，那么 $\alpha(e_{j_1},\dots,e_{j_n})=0$；最后一行成立是由于 9.35. ■

下面结论是引出下一节中行列式的定义的关键.

**9.37 $\dim V_{\mathrm{alt}}^{(\dim V)}=1$**

向量空间 $V_{\mathrm{alt}}^{(\dim V)}$ 的维数是 1.

**证明** 令 $n=\dim V$. 设 $\alpha$ 和 $\alpha'$ 是 $V$ 上的交错 $n$ 重线性型且 $\alpha\neq 0$. 令 $e_1,\dots,e_n$ 满足 $\alpha(e_1,\dots,e_n)\neq 0$. 存在 $c\in\mathbf{F}$ 使得

$$\alpha'(e_1,\dots,e_n)=c\alpha(e_1,\dots,e_n).$$

此外，9.28 表明，$e_1,\dots,e_n$ 是线性无关的，因而它是 $V$ 的一个基.

设 $v_1,\dots,v_n\in V$. 令 $b_{j,k}$（$j,k=1,\dots,n$）定义如 9.36. 那么

$$\begin{aligned}\alpha'(v_1,\dots,v_n) &= \alpha'(e_1,\dots,e_n)\sum_{(j_1,\dots,j_n)\in\operatorname{perm} n}\big(\operatorname{sign}(j_1,\dots,j_n)\big)b_{j_1,1}\cdots b_{j_n,n}\\ &= c\alpha(e_1,\dots,e_n)\sum_{(j_1,\dots,j_n)\in\operatorname{perm} n}\big(\operatorname{sign}(j_1,\dots,j_n)\big)b_{j_1,1}\cdots b_{j_n,n}\\ &= c\alpha(v_1,\dots,v_n),\end{aligned}$$

其中第一行和最后一行源于 9.36. 上式表明 $\alpha'=c\alpha$. 于是 $\alpha',\alpha$ 不是线性无关组，这就表明 $\dim V_{\mathrm{alt}}^{(n)}\leq 1$.

为了完成证明，我们只需证明 $V$ 上总存在非零的交错 $n$ 重线性型 $\alpha$（由此排除 $\dim V_{\mathrm{alt}}^{(n)}=0$ 的可能性）. 为此，令 $e_1,\dots,e_n$ 是 $V$ 的任意一个基，并令 $\varphi_1,\dots,\varphi_n\in V'$ 是 $V$ 上的线性泛函，且它们可用以将 $V$ 的每个元素都表达为 $e_1,\dots,e_n$ 的线性组合，也就是说

$$v=\sum_{j=1}^{n}\varphi_j(v)e_j$$

对每个 $v\in V$ 成立（见 3.114）. 现在对 $v_1,\dots,v_n\in V$，定义

$$\alpha(v_1,\dots,v_n)=\sum_{(j_1,\dots,j_n)\in\operatorname{perm} n}\big(\operatorname{sign}(j_1,\dots,j_n)\big)\varphi_{j_1}(v_1)\cdots\varphi_{j_n}(v_n). \tag{9.38}$$

容易验证 $\alpha$ 是 $V$ 上的 $n$ 重线性型.

为了说明 $\alpha$ 是交错的，设 $v_1,\ldots,v_n\in V$ 满足 $v_1=v_2$．对于每个 $(j_1,\ldots,j_n)\in\operatorname{perm}n$，排列 $(j_2,j_1,j_3,\ldots,j_n)$ 都有与之相反的符号．因为 $v_1=v_2$，所以式 (9.38) 的求和中有关这两个排列的两项彼此抵消．因此 $\alpha(v_1,v_1,v_3,\ldots,v_n)=0$．类似地，如果组 $v_1,\ldots,v_n$ 中任意两个向量相等，那么 $\alpha(v_1,\ldots,v_n)=0$．于是 $\alpha$ 是交错的．

最后，在式 (9.38) 中取各 $v_k=e_k$．因为 $\varphi_j(e_k)$ 等于 0（如果 $j\neq k$）或 1（如果 $j=k$），所以求和式右端只有关于排列 $(1,\ldots,n)$ 的一项是非零的，于是可得 $\alpha(e_1,\ldots,e_n)=1$．于是我们就构造出了 $V$ 上的非零交错 $n$ 重线性型，则原命题得证．■

上个证明中用于构造非零交错 $n$ 重线性型的式 (9.38) 来源于 9.36 中的公式，而后者可很自然地由交错多重线性型的性质得到．

前面我们证明了将交错多重线性型作用于线性相关组会得到 0（见 9.28）．下面结果给出了 9.28 在 $n=\dim V$ 情形下的逆命题．在下面结果中，"$\alpha$ 非零"意为 $\alpha$ 不是 $V^n$ 上的常值函数 0（与它用于函数时的常见含义相同）．

**9.39 交错 $(\dim V)$ 重线性型与线性无关性**

令 $n=\dim V$．设 $\alpha$ 是 $V$ 上的非零交错 $n$ 重线性型，且 $e_1,\ldots,e_n$ 是 $V$ 中的向量组．那么

$$\alpha(e_1,\ldots,e_n)\neq 0$$

当且仅当 $e_1,\ldots,e_n$ 是线性无关的．

**证明** 首先设 $\alpha(e_1,\ldots,e_n)\neq 0$．那么由 9.28，$e_1,\ldots,e_n$ 是线性无关的．

为了证明另一方向的蕴涵关系，现设 $e_1,\ldots,e_n$ 是线性无关的．因为 $n=\dim V$，所以 $e_1,\ldots,e_n$ 是 $V$ 的一个基（见 2.38）．

因为 $\alpha$ 是非零 $n$ 重线性型，所以存在 $v_1,\ldots,v_n\in V$ 满足 $\alpha(v_1,\ldots,v_n)\neq 0$．现由 9.36 可得 $\alpha(e_1,\ldots,e_n)\neq 0$．■

## 习题 9B

1 设 $m$ 是正整数．证明：$\dim V^{(m)}=(\dim V)^m$．

2 设 $n\geq 3$ 且 $\alpha:\mathbf{F}^n\times\mathbf{F}^n\times\mathbf{F}^n\to\mathbf{F}$ 定义为

$$\begin{aligned}&\alpha\big((x_1,\ldots,x_n),(y_1,\ldots,y_n),(z_1,\ldots,z_n)\big)\\&\quad=x_1y_2z_3-x_2y_1z_3-x_3y_2z_1-x_1y_3z_2+x_3y_1z_2+x_2y_3z_1.\end{aligned}$$

证明：$\alpha$ 是 $\mathbf{F}^n$ 上的交错 3 重线性型．

3 设 $m$ 是正整数，$\alpha$ 是 $V$ 上的 $m$ 重线性型，以及：只要 $v_1,\ldots,v_m$ 是 $V$ 中的一组向量，且 $v_j=v_{j+1}$ 对于某个 $j\in\{1,\ldots,m-1\}$ 成立，就有 $\alpha(v_1,\ldots,v_m)=0$．证明：$\alpha$ 是 $V$ 上的交错 $m$ 重线性型．

4 证明或给出一反例：如果 $\alpha\in V^{(4)}_{\mathrm{alt}}$，那么

$$\{(v_1,v_2,v_3,v_4)\in V^4:\alpha(v_1,v_2,v_3,v_4)=0\}$$

是 $V^4$ 的子空间．

**5** 设 $m$ 是正整数，$\beta$ 是 $V$ 上的 $m$ 重线性型．定义 $V$ 上的 $m$ 重线性型 $\alpha$ 为

$$\alpha(v_1,\ldots,v_m)=\sum_{(j_1,\ldots,j_m)\in\operatorname{perm}m}\big(\operatorname{sign}(j_1,\ldots,j_m)\big)\beta(v_{j_1},\ldots,v_{j_m}).$$

其中，$v_1,\ldots,v_m\in V$．解释为什么 $\alpha\in V_{\mathrm{alt}}^{(m)}$．

**6** 设 $m$ 是正整数，$\beta$ 是 $V$ 上的 $m$ 重线性型．定义 $V$ 上的 $m$ 重线性型 $\alpha$ 为

$$\alpha(v_1,\ldots,v_m)=\sum_{(j_1,\ldots,j_m)\in\operatorname{perm}m}\beta(v_{j_1},\ldots,v_{j_m}).$$

其中，$v_1,\ldots,v_m\in V$．解释为什么

$$\alpha(v_{k_1},\ldots,v_{k_m})=\alpha(v_1,\ldots,v_m)$$

对所有 $v_1,\ldots,v_m\in V$ 和所有 $(k_1,\ldots,k_m)\in\operatorname{perm}m$ 成立．

**7** 给出一例：$\mathbf{R}^3$ 上的非零交错 2 重线性型 $\alpha$ 与 $\mathbf{R}^3$ 中一线性无关组 $v_1,v_2$，使得 $\alpha(v_1,v_2)=0$.

注　本题表明，如果删去 $n=\dim V$ 这一假设，9.39 可能就不成立了．

# 9C 行列式

## 定义行列式

下面定义将把我们引向算子的行列式的一个简洁、美妙且不依赖于基的定义.

**9.40 定义：$\alpha_T$**

设 $m$ 是正整数, $T \in \mathcal{L}(V)$. 对 $\alpha \in V_{\text{alt}}^{(m)}$, 定义 $\alpha_T \in V_{\text{alt}}^{(m)}$ 为: 对 $V$ 中每个向量组 $v_1, \ldots, v_m$,

$$\alpha_T(v_1, \ldots, v_m) = \alpha(Tv_1, \ldots, Tv_m).$$

♣

设 $T \in \mathcal{L}(V)$. 如果 $\alpha \in V_{\text{alt}}^{(m)}$, $v_1, \ldots, v_m$ 是 $V$ 中的向量组且对某个 $j \neq k$ 存在 $v_j = v_k$, 那么 $Tv_j = Tv_k$, 这表明 $\alpha_T(v_1, \ldots, v_m) = \alpha(Tv_1, \ldots, Tv_m) = 0$. 于是函数 $\alpha \mapsto \alpha_T$ 是从 $V_{\text{alt}}^{(m)}$ 到其自身的线性映射.

我们知道, $\dim V_{\text{alt}}^{(\dim V)} = 1$ (见 9.37). 每个从一维向量空间到该空间本身的线性映射, 就是将向量与某个特定的标量相乘. 我们现在就将 $\det T$ 定义为线性映射 $\alpha \mapsto \alpha_T$ 中乘在向量之前的那个标量.

**9.41 定义：算子的行列式（determinant of an operator）、$\det T$**

设 $T \in \mathcal{L}(V)$. $T$ 的**行列式**, 记作 $\det T$, 定义为 $\mathbf{F}$ 中唯一使得

$$\alpha_T = (\det T)\alpha$$

对所有 $\alpha \in V_{\text{alt}}^{(\dim V)}$ 都成立的数.

♣

**9.42 例：算子的行列式**

令 $n = \dim V$.

- 若 $I$ 是 $V$ 上的恒等算子, 那么对所有 $\alpha \in V_{\text{alt}}^{(n)}$, $\alpha_I = \alpha$. 于是 $\det I = 1$.
- 更一般地, 如果 $\lambda \in \mathbf{F}$, 那么对所有 $\alpha \in V_{\text{alt}}^{(n)}$, $\alpha_{\lambda I} = \lambda^n \alpha$. 于是 $\det(\lambda I) = \lambda^n$.
- 进一步地, 若 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$, 那么对所有 $\alpha \in V_{\text{alt}}^{(n)}$, $\alpha_{\lambda T} = \lambda^n \alpha_T = \lambda^n (\det T)\alpha$. 于是 $\det(\lambda T) = \lambda^n \det T$.
- 设 $T \in \mathcal{L}(V)$, $e_1, \ldots, e_n$ 是由 $T$ 的特征向量（分别对应于特征值 $\lambda_1, \ldots, \lambda_n$）构成的 $V$ 的基. 若 $\alpha \in V_{\text{alt}}^{(n)}$, 那么

$$\alpha_T(e_1, \ldots, e_n) = \alpha(\lambda_1 e_1, \ldots, \lambda_n e_n) = (\lambda_1 \cdots \lambda_n)\alpha(e_1, \ldots, e_n).$$

若 $\alpha \neq 0$, 那么由 9.39 得 $\alpha(e_1, \ldots, e_n) \neq 0$. 于是上式表明

$$\det T = \lambda_1 \cdots \lambda_n.$$

我们接下来的任务, 是定义方阵的行列式并给出其公式. 为此, 我们将每个方阵都与一算子关联起来, 然后定义方阵的行列式是其所关联的算子的行列式.

**9.43 定义：矩阵的行列式（determinant of a matrix）、$\det A$**

设 $n$ 是正整数，且 $A$ 是各元素均属于 $\mathbf{F}$ 的 $n \times n$ 方阵．令 $T \in \mathcal{L}(\mathbf{F}^n)$ 关于 $\mathbf{F}^n$ 的标准基的矩阵等于 $A$．$A$ 的**行列式**，记为 $\det A$，定义为 $\det A = \det T$．

**9.44 例：矩阵的行列式**

- 若 $I$ 是 $n \times n$ 恒等矩阵，那么它对应的 $\mathbf{F}^n$ 上的算子是 $\mathbf{F}^n$ 上的恒等算子 $I$．于是由 9.42 的第一例可得恒等矩阵的行列式是 1．
- 设 $A$ 是对角矩阵，对角线上元素为 $\lambda_1, \ldots, \lambda_n$．那么其对应的 $\mathbf{F}^n$ 上算子有一组特征向量是 $\mathbf{F}^n$ 的标准基，其中各向量分别对应于特征值 $\lambda_1, \ldots, \lambda_n$．于是由 9.42 的最后一例可得 $\det A = \lambda_1 \cdots \lambda_n$．

对于下面结果，将 $\mathbf{F}^n$ 中每个由 $n$ 个向量 $v_1, \ldots, v_n$ 构成的组看成由若干个 $n \times 1$ 列向量构成的组．这样，记号 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 就表示第 $k$ 列是 $v_k$（$k = 1, \ldots, n$）的 $n \times n$ 方阵．

**9.45 行列式是交错多重线性型**

设 $n$ 是正整数．将 $\mathbf{F}^n$ 中向量组 $v_1, \ldots, v_n$ 对应到 $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 的映射是 $\mathbf{F}^n$ 上的交错 $n$ 重线性型．

**证明** 令 $e_1, \ldots, e_n$ 是 $\mathbf{F}^n$ 的标准基，并假设 $v_1, \ldots, v_n$ 是 $\mathbf{F}^n$ 中的向量组．令 $T \in \mathcal{L}(\mathbf{F}^n)$ 是满足 $Te_k = v_k$（$k = 1, \ldots, n$）的算子．于是 $T$ 关于 $e_1, \ldots, e_n$ 的矩阵是 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$．于是由矩阵的行列式的定义，$\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \det T$．

令 $\alpha$ 是 $\mathbf{F}^n$ 上的交错 $n$ 重线性型且使得 $\alpha(e_1, \ldots, e_n) = 1$．那么

$$\begin{aligned}\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} &= \det T \\ &= (\det T)\alpha(e_1, \ldots, e_n) \\ &= \alpha(Te_1, \ldots, Te_n) \\ &= \alpha(v_1, \ldots, v_n),\end{aligned}$$

其中第三行来自算子的行列式的定义．上式表明，将 $\mathbf{F}^n$ 中向量组 $v_1, \ldots, v_n$ 对应到 $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 的映射是 $\mathbf{F}^n$ 上的交错 $n$ 重线性型 $\alpha$． ■

上述结论有几个重要推论．例如，由该结论立刻可知，有两列相同的矩阵的行列式等于 0．我们稍后会回过头来讨论其他推论，在此之前我们先给出方阵的行列式的公式．回忆一下，如果 $A$ 是一个矩阵，那么 $A_{j,k}$ 表示 $A$ 第 $j$ 行第 $k$ 列中的元素．

**9.46 矩阵的行列式的公式**

设 $n$ 是正整数且 $A$ 是 $n \times n$ 方阵．那么

$$\det A = \sum_{(j_1, \ldots, j_n) \in \operatorname{perm} n} (\operatorname{sign}(j_1, \ldots, j_n)) A_{j_1,1} \cdots A_{j_n,n}.$$

**证明** 运用 9.36，其中取 $V=\mathbf{F}^n$、$e_1,\ldots,e_n$ 为 $\mathbf{F}^n$ 的标准基、$\alpha$ 为 $\mathbf{F}^n$ 上将 $v_1,\ldots,v_n$ 对应到 $\det\begin{pmatrix} v_1 & \cdots & v_n\end{pmatrix}$ 的交错 $n$ 重线性型（见 9.45）. 若各 $v_k$ 等于 $A$ 的第 $k$ 列，则 9.36 中的各 $b_{j,k}$ 等于 $A_{j,k}$. 最后，

$$\alpha(e_1,\ldots,e_n)=\det\begin{pmatrix} e_1 & \cdots & e_n\end{pmatrix}=\det I=1.$$

于是 9.36 中的公式就成为了本结果所述的公式. ■

---

**9.47 例：行列式公式的具体示例**

- 若 $A$ 是 $2\times 2$ 矩阵，那么 9.46 中的公式变为

$$\det A=A_{1,1}A_{2,2}-A_{2,1}A_{1,2}.$$

- 若 $A$ 是 $3\times 3$ 矩阵，那么 9.46 中的公式变为

$$\begin{aligned}\det A=&A_{1,1}A_{2,2}A_{3,3}-A_{2,1}A_{1,2}A_{3,3}-A_{3,1}A_{2,2}A_{1,3}\\&-A_{1,1}A_{3,2}A_{2,3}+A_{3,1}A_{1,2}A_{2,3}+A_{2,1}A_{3,2}A_{1,3}.\end{aligned}$$

---

9.46 中的公式包含 $n!$ 个求和项. 因为随 $n$ 增长，$n!$ 会急剧增加，所以即便对于中等大小的 $n$ 阶行列式，用 9.46 中的公式来求其值也不可行. 例如，10! 超过三百万，100! 接近 $10^{158}$，即使用最快的计算机也无法计算如此多项之和. 很快我们将看到一些比直接运用 9.46 中的求和式更快的行列式求值方法.

**9.48 上三角矩阵的行列式**

设 $A$ 是上三角矩阵，其对角线上各元素是 $\lambda_1,\ldots,\lambda_n$. 那么 $\det A=\lambda_1\cdots\lambda_n$.

**证明** 若 $(j_1,\ldots,j_n)\in\operatorname{perm} n$ 满足 $(j_1,\ldots,j_n)\neq(1,\ldots,n)$，那么对某些 $k\in\{1,\ldots,n\}$ 有 $j_k>k$，这表明 $A_{j_k,k}=0$. 于是只有 $(1,\ldots,n)$ 这个排列为 9.46 中求和式贡献了非零项. 因为对各 $k=1,\ldots,n$ 有 $A_{k,k}=\lambda_k$，故由此可得 $\det A=\lambda_1\cdots\lambda_n$. ■

## 行列式的性质

行列式的定义可引出行列式可乘性的如下妙证.

**9.49 行列式是可乘的**

(a) 设 $S,T\in\mathcal{L}(V)$. 那么 $\det(ST)=(\det S)(\det T)$.

(b) 设 $A$ 和 $B$ 是大小相同的方阵. 那么

$$\det(AB)=(\det A)(\det B).$$

**证明**

(a) 令 $n=\dim V$. 设 $\alpha\in V_{\text{alt}}^{(n)}$ 且 $v_1,\ldots,v_n\in V$. 那么

$$\begin{aligned}\alpha_{ST}(v_1,\ldots,v_n)&=\alpha(STv_1,\ldots,STv_n)\\&=(\det S)\alpha(Tv_1,\ldots,Tv_n)\\&=(\det S)(\det T)\alpha(v_1,\ldots,v_n),\end{aligned}$$

其中第一个等式是由 $\alpha_{ST}$ 的定义而来，第二个等式是由 $\det S$ 的定义而来，第三个等式则源于 $\det T$ 的定义．上式就表明 $\det(ST) = (\det S)(\det T)$．

(b) 令 $S, T \in \mathcal{L}(\mathbf{F}^n)$ 使得 $\mathcal{M}(S) = A$ 及 $\mathcal{M}(T) = B$（本证明中所有算子的矩阵都是关于 $\mathbf{F}^n$ 的标准基）．那么 $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = AB$（见 3.43）．于是

$$\det(AB) = \det(ST) = (\det S)(\det T) = (\det A)(\det B).$$

其中第二个等号源于 (a) 中结论． ■

算子的行列式决定了该算子是否可逆．

**9.50 可逆 $\Longleftrightarrow$ 行列式非零**

算子 $T \in \mathcal{L}(V)$ 是可逆的，当且仅当 $\det T \neq 0$. 此外，若 $T$ 是可逆的，那么 $\det(T^{-1}) = \frac{1}{\det T}$.

**证明** 先设 $T$ 可逆．于是 $TT^{-1} = I$．现由 9.49 得

$$1 = \det I = \det(TT^{-1}) = (\det T)\left(\det(T^{-1})\right).$$

因此 $\det T \neq 0$ 且 $\det(T^{-1})$ 是 $\det T$ 的乘法逆元．

为了证明另一方向，设 $\det T \neq 0$．设 $v \in V$ 且 $v \neq 0$．令 $v, e_2, \ldots, e_n$ 是 $V$ 的一个基，并令 $\alpha \in V_{\mathrm{alt}}^{(n)}$ 满足 $\alpha \neq 0$．那么 $\alpha(v, e_2, \ldots, e_n) \neq 0$（由 9.39）．则

$$\alpha(Tv, Te_2, \ldots, Te_n) = (\det T)\alpha(v, e_2, \ldots, e_n) \neq 0,$$

从而 $Tv \neq 0$．因此 $T$ 可逆． ■

任何 $n \times n$ 矩阵 $A$ 可逆（可逆矩阵的定义见 3.80）当且仅当与 $A$（通过 $\mathbf{F}^n$ 的标准基）关联的 $\mathbf{F}^n$ 上算子是可逆的．于是上面结果就表明，方阵 $A$ 可逆当且仅当 $\det A \neq 0$．

**9.51 特征值和行列式**

设 $T \in \mathcal{L}(V)$ 且 $\lambda \in \mathbf{F}$．那么 $\lambda$ 是 $T$ 的特征值，当且仅当 $\det(\lambda I - T) = 0$．

**证明** 数 $\lambda$ 是 $T$ 的特征值当且仅当 $T - \lambda I$ 不可逆（见 5.7），这等价于 $\lambda I - T$ 不可逆，进而等价于 $\det(\lambda I - T) = 0$（由 9.50）． ■

设 $T \in \mathcal{L}(V)$ 且 $S : W \to V$ 是可逆线性映射．为证明 $\det(S^{-1}TS) = \det T$，我们可尝试利用 9.49 和 9.50 写出

$$\begin{aligned}\det(S^{-1}TS) &= \left(\det S^{-1}\right)(\det T)(\det S)\\ &= \det T.\end{aligned}$$

若 $W = V$，那么可用这个证法；但如果 $W \neq V$，那么此方法没有意义，因为行列式只对由一个向量空间到其自身的线性映射有定义，而 $S$ 将 $W$ 映射到 $V$，则 $\det S$ 无定义．下面的证明规避了这个问题，因而对 $W \neq V$ 情形也成立．

**9.52 行列式是相似不变量**

设 $T \in \mathcal{L}(V)$ 且 $S : W \to V$ 是可逆线性映射．那么

$$\det(S^{-1}TS) = \det T.$$

**证明** 令 $n = \dim W = \dim V$. 设 $\tau \in W_{\mathrm{alt}}^{(n)}$. 定义 $\alpha \in V_{\mathrm{alt}}^{(n)}$ 为：对 $v_1, \ldots, v_n \in V$,

$$\alpha(v_1, \ldots, v_n) = \tau(S^{-1}v_1, \ldots, S^{-1}v_n).$$

设 $w_1, \ldots, w_n \in W$. 那么

$$\begin{aligned}\tau_{S^{-1}TS}(w_1, \ldots, w_n) &= \tau(S^{-1}TSw_1, \ldots, S^{-1}TSw_n)\\ &= \alpha(TSw_1, \ldots, TSw_n)\\ &= \alpha_T(Sw_1, \ldots, Sw_n)\\ &= (\det T)\alpha(Sw_1, \ldots, Sw_n)\\ &= (\det T)\tau(w_1, \ldots, w_n).\end{aligned}$$

结合上式与算子 $S^{-1}TS$ 的行列式的定义可知 $\det(S^{-1}TS) = \det T$. ■

对于 $V = \mathbf{F}^n$ 且 $e_1, \ldots, e_n$ 是 $\mathbf{F}^n$ 的标准基这个特殊情形，由矩阵的行列式的定义知下面结论成立. 下面结论中等式左侧并不依赖于基的选取，这表明等式右侧也与基的选取无关.

**9.53 算子的行列式等于其矩阵的行列式**

设 $T \in \mathcal{L}(V)$ 且 $e_1, \ldots, e_n$ 是 $V$ 的基. 那么

$$\det T = \det \mathcal{M}\big(T, (e_1, \ldots, e_n)\big).$$

**证明** 令 $f_1, \ldots, f_n$ 是 $\mathbf{F}^n$ 的标准基. 令 $S : \mathbf{F}^n \to V$ 是满足 $Sf_k = e_k$（$k = 1, \ldots, n$）的线性映射. 于是 $\mathcal{M}\big(S, (f_1, \ldots, f_n), (e_1, \ldots, e_n)\big)$ 和 $\mathcal{M}\big(S^{-1}, (e_1, \ldots, e_n), (f_1, \ldots, f_n)\big)$ 都等于 $n \times n$ 恒等矩阵. 因此运用 3.43 两次得

$$\mathcal{M}\big(S^{-1}TS, (f_1, \ldots, f_n)\big) = \mathcal{M}\big(T, (e_1, \ldots, e_n)\big), \tag{9.54}$$

于是

$$\begin{aligned}\det T &= \det(S^{-1}TS)\\ &= \det \mathcal{M}\big(S^{-1}TS, (f_1, \ldots, f_n)\big)\\ &= \det \mathcal{M}\big(T, (e_1, \ldots, e_n)\big),\end{aligned}$$

其中第一行源于 9.52，第二行源于矩阵的行列式的定义，第三行则由式 (9.54) 得到. ■

下面结论给出了另一种看待行列式的方式，它比行列式定义或 9.46 中的公式都更直观. 我们可以将下面结论所给出的行列式的特性，作为有限维复向量空间上算子的行列式的定义，这样的话，现行定义就变成了这个新定义的一个结论.

**9.55 若 $\mathbf{F} = \mathbf{C}$，那么行列式等于特征值之积**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 那么 $\det T$ 等于 $T$ 的特征值之积（其中每个特征值出现次数等于其重数）.

**证明** 存在 $V$ 的一个基，使得 $T$ 关于该基有上三角矩阵，且该矩阵对角线上各元素是 $T$ 的特征值，各特征值出现次数等于其重数——见 8.37. 于是 9.53 和 9.48 表明，$\det T$ 等于 $T$ 的特征值之积（其中每个特征值出现次数等于其重数）. ■

接下来这条结果表明，对于方阵的转置、算子的对偶及内积空间上算子的伴随这些操作，行列式都具有很好的性质.

**9.56 转置、对偶或伴随的行列式**

(a) 设 $A$ 是方阵. 那么 $\det A^{\mathrm{t}} = \det A$.

(b) 设 $T \in \mathcal{L}(V)$. 那么 $\det T' = \det T$.

(c) 设 $V$ 是内积空间且 $T \in \mathcal{L}(V)$. 那么

$$\det(T^*) = \overline{\det T}.$$

**证明**

(a) 令 $n$ 为正整数. 定义 $\alpha : (\mathbf{F}^n)^n \to \mathbf{F}$ 为: 对于所有 $v_1, \ldots, v_n \in \mathbf{F}^n$,

$$\alpha(v_1, \ldots, v_n) = \det\Big(\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}\Big).$$

9.46 中的矩阵的行列式公式表明 $\alpha$ 是 $\mathbf{F}^n$ 上的 $n$ 重线性型.

设 $v_1, \ldots, v_n \in \mathbf{F}^n$ 且对某个 $j \neq k$ 有 $v_j = v_k$. 若 $B$ 是 $n \times n$ 矩阵, 那么 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}} B$ 不等于恒等矩阵, 因为 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}} B$ 的第 $j$ 行和第 $k$ 行相等. 于是 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}$ 不可逆, 这表明 $\alpha(v_1, \ldots, v_n) = 0$. 因此 $\alpha$ 是 $\mathbf{F}^n$ 上的交错 $n$ 重线性型.

注意到将 $\alpha$ 作用于 $\mathbf{F}^n$ 的标准基会得到 1. 因为 $\mathbf{F}^n$ 上交错 $n$ 重线性型构成的向量空间的维数是 1（由 9.37）, 所以这表明 $\alpha$ 就是行列式函数[3]. 于是 (a) 成立.

(b) 由 (a)、9.53 及 3.132 可得等式 $\det T' = \det T$.

(c) 选取 $V$ 的一个规范正交基. $T^*$ 关于该基的矩阵是 $T$ 关于该基的矩阵的共轭转置（由 7.9）. 于是由 9.53、9.46 及 (a) 可得 $\det(T^*) = \overline{\det T}$. ■

**9.57 有助于计算行列式的若干结论**

(a) 如果一方阵中任意两行或两列相等, 则该矩阵的行列式等于 0.

(b) 设 $A$ 是方阵, 且 $B$ 是通过交换 $A$ 中任意两行或两列所得的矩阵. 则 $\det A = -\det B$.

(c) 若将方阵的某行或某列乘以一个标量, 则其行列式的值也会被乘以同一个标量.

(d) 若将一方阵中某列的标量倍加到该矩阵中另一列, 则其行列式的值不变.

(e) 若将一方阵中某行的标量倍加到该矩阵中另一行, 则其行列式的值不变.

**证明** 上述结果中的所有论断成立, 都是由于 $v_1, \ldots, v_n \mapsto \det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 和 $v_1, \ldots, v_n \mapsto \det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}$ 这两个映射都是 $\mathbf{F}^n$ 上的交错 $n$ 重线性型【见 9.45 和 9.56 (a)】.

例如, 为证明 (d), 设 $v_1, \ldots, v_n \in \mathbf{F}^n$ 且 $c \in \mathbf{F}$. 那么

$$\begin{aligned}
&\det\begin{pmatrix} v_1 + cv_2 & v_2 & \cdots & v_n \end{pmatrix} \\
&\quad = \det\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} + c\det\begin{pmatrix} v_2 & v_2 & v_3 & \cdots & v_n \end{pmatrix} \\
&\quad = \det\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},
\end{aligned}$$

其中第一个等式源于行列式的多重线性性质, 第二个等式则源于行列式的交错性质. 上式表明, 将一矩阵第二列的标量倍加到该矩阵第一列不改变其行列式的值. 同样的结论对于任意的两列也是成立的. 于是 (d) 成立.

(e) 的证明可由 (d) 和 9.56(a) 得出. (a)、(b) 和 (c) 的证明所用手段类似, 留给读者完成. ■

[3] 意即 $\alpha\big(\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}\big)$ 就是 $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 的行列式.

对于各元素都为具体的数的矩阵，上述结果可引出一种比直接应用 9.46 中的公式更快的计算行列式的方法. 具体而言，应用高斯消元法，即交换两行【由 9.57 (b)，这将把行列式变为原来的 −1 倍】、将一行乘以非零常数【由 9.57 (c)，这将把行列式乘以相同常数】以及将一行的若干倍加到另一行【由 9.57 (e)，这不改变行列式的值】，可将矩阵化为上三角矩阵，而上三角矩阵的行列式就是对角线上各元素之积（由 9.48）. 如果你的软件能记录交换各行的次数，以及将一行乘以常数时所用的常数，那么就可以反算出原矩阵的行列式.

因为数 $\lambda \in \mathbf{F}$ 是算子 $T \in \mathcal{L}(V)$ 的特征值当且仅当 $\det(\lambda I - T) = 0$（由 9.51），你可能由此想到一种快速求解特征值的方法：取 $V$ 的一个基，令 $A = \mathcal{M}(T)$，求出 $\det(\lambda I - A)$，然后求解方程 $\det(\lambda I - A) = 0$ 来得出 $\lambda$. 然而，这个过程往往并不高效，除非 $\dim V = 2$（或 3 和 4，如果你愿意使用三次或四次求根公式的话）. 有个问题是，当矩阵中包含符号（例如 $\lambda I - A$ 中的 $\lambda$）时，上段所述的计算行列式的过程就不适用了. 出现此问题，是因为在高斯消元过程中需要确定某些量是否等于 0，而这在涉及符号 $\lambda$ 的表达式中比较复杂.

回忆一下，有限维内积空间上的算子是幺正的，当且仅当它是保范数的（见 7.51 及其后的一段）. 幺正算子的每个特征值的绝对值都是 1（由 7.54）. 于是幺正算子的特征值之积的绝对值是 1. 因此（至少在 $\mathbf{F} = \mathbf{C}$ 情况下）幺正算子的行列式的绝对值是 1（由 9.55）. 下面结论给出了不依赖于 $\mathbf{F} = \mathbf{C}$ 这个前提的证明.

**9.58 每个幺正算子的行列式绝对值都为 1**

设 $V$ 是内积空间，$S \in \mathcal{L}(V)$ 是幺正算子. 那么 $|\det S| = 1$.

**证明** 因为 $S$ 是幺正的，所以 $I = S^*S$（见 7.53）. 于是

$$1 = \det(S^*S) = (\det S^*)(\det S) = \overline{(\det S)}(\det S) = |\det S|^2,$$

其中第二个等号源于 9.49 (a)，第三个等号源于 9.56 (c). 上式表明 $|\det S| = 1$. ■

内积空间上正算子的行列式与将这类算子对应成非负实数的类比吻合得很好.

**9.59 每个正算子都具有非负行列式**

设 $V$ 是内积空间且 $T \in \mathcal{L}(V)$ 是正算子. 那么 $\det T \geq 0$.

**证明** 由谱定理（7.29 或 7.31），$V$ 有由 $T$ 的特征向量构成的规范正交基. 于是由 9.42 的最后一例，$\det T$ 等于 $T$ 的特征值之积（其中可能有重复的特征值）. $T$ 的每个特征值都是非负数（由 7.38）. 于是我们得出结论 $\det T \geq 0$. ■

设 $V$ 是内积空间且 $T \in \mathcal{L}(V)$. 回忆一下，$T^*T$ 的特征值的非负平方根构成的组（各数出现次数等于其对应特征值的重数）被称为 $T$ 的奇异值构成的组（见 7E 节）.

**9.60 $|\det T| = (T$ 的奇异值之积$)$**

设 $V$ 是内积空间且 $T \in \mathcal{L}(V)$. 那么

$$|\det T| = \sqrt{\det(T^*T)} = (T \text{ 的奇异值之积}).$$

**证明** 我们有

$$|\det T|^2 = \overline{(\det T)}(\det T) = \big(\det(T^*)\big)(\det T) = \det(T^*T),$$

其中第二个等号源于 9.56 (c)，最后一个等号源于 9.49 (a). 将上式两边同时开平方根就证明了 $|\det T| = \sqrt{\det(T^*T)}$.

令 $s_1, \dots, s_n$ 表示 $T$ 的奇异值构成的组. 于是 $s_1^2, \dots, s_n^2$ 是 $T^*T$ 的特征值构成的组（可能有重复），它们对应由 $T^*T$ 的特征向量构成的 $V$ 的规范正交基. 因此由 9.42 的最后一例得

$$\det(T^*T) = s_1^2 \cdots s_n^2.$$

于是 $|\det T| = s_1 \cdots s_n$，原命题得证. ■

实内积空间上的算子 $T$ 作用于体积，会将其乘以 $T$ 的奇异值之积（由 7.111）. 于是由 7.111 和 9.60 立即可得接下来的结果. 这个结果解释了为什么在多变量微积分的变量替换公式中会出现行列式的绝对值.

**9.61 $T$ 将体积变为其 $|\det T|$ 倍**

设 $T \in \mathcal{L}(\mathbf{R}^n)$ 且 $\Omega \subseteq \mathbf{R}^n$. 那么

$$\operatorname{volume} T(\Omega) = |\det T|\,(\operatorname{volume} \Omega).$$

♡

对于有限维复向量空间上的算子，我们现将其行列式与前面见过的一个多项式联系起来.

**9.62 若 $\mathbf{F} = \mathbf{C}$，那么 $T$ 的特征多项式等于 $\det(zI - T)$**

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 令 $\lambda_1, \dots, \lambda_m$ 表示 $T$ 的所有互异特征值，并令 $d_1, \dots, d_m$ 表示它们的重数. 那么

$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

♡

**证明** 存在 $V$ 的一个基，使得 $T$ 关于该基有上三角矩阵，且各 $\lambda_k$ 恰在其对角线出现 $d_k$ 次（由 8.37）. 关于该基，$zI - T$ 有上三角矩阵，且各 $z - \lambda_k$ 恰在其对角线出现 $d_k$ 次. 从而由 9.48 可得欲证等式. ■

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$. 在 8.26 中，$T$ 的特征多项式被定义作 9.62 中的等式右端. 我们之前并未对有限维实向量空间上的算子定义特征多项式，因为这类算子可能没有特征值，用 9.62 中的等式右端来定义就不合适了.

我们现呈上特征多项式的一个新定义，它是由 9.62 所启发的. 这个新定义对实向量空间和复向量空间都成立. 9.62 中的等式说明，$\mathbf{F} = \mathbf{C}$ 时，这个新定义等价于我们之前所作的定义（8.26）.

**9.63 定义：特征多项式（characteristic polynomial）**

设 $T \in \mathcal{L}(V)$. 定义为

$$z \mapsto \det(zI - T)$$

的多项式被称为 $T$ 的**特征多项式**.

♣

9.46 中的公式表明，算子 $T \in \mathcal{L}(V)$ 的特征多项式是一个次数为 $\dim V$ 的首一多项式. $T$ 的特征多项式在 $\mathbf{F}$ 中的零点恰为 $T$ 的特征值（由 9.51）.

之前我们证明了复数情形下的凯莱-哈密尔顿定理（8.29）. 现在我们可将这条结果拓展至实向量空间上的算子了.

**9.64 凯莱-哈密尔顿定理（Cayley-Hamilton theorem）**

设 $T \in \mathcal{L}(V)$ 且 $q$ 是 $T$ 的特征多项式. 那么 $q(T) = 0$.

**证明** 若 $\mathbf{F} = \mathbf{C}$，那么由 9.62 和 8.29 得等式 $q(T) = 0$.

现设 $\mathbf{F} = \mathbf{R}$. 取定 $V$ 的一个基，并令 $A$ 是 $T$ 关于该基的矩阵. 令 $S$ 是 $\mathbf{C}^{\dim V}$ 上的算子，满足 $S$（关于 $\mathbf{C}^{\dim V}$ 的标准基）的矩阵是 $A$. 对所有 $z \in \mathbf{R}$，我们有

$$q(z) = \det(zI - T) = \det(zI - A) = \det(zI - S).$$

于是 $q$ 是 $S$ 的特征多项式. 现由 $\mathbf{F} = \mathbf{C}$ 的情形（本证明第一句话）得 $0 = q(S) = q(A) = q(T)$. ■

凯莱-哈密尔顿定理（9.64）表明，算子 $T \in \mathcal{L}(V)$ 的特征多项式是最小多项式的多项式倍（由 5.29）. 于是，如果 $T$ 的最小多项式的次数等于 $\dim V$，那么 $T$ 的特征多项式等于 $T$ 的最小多项式. 绝大多数算子，比如超过 99.999% 的 $4 \times 4$ 矩阵对应的算子【矩阵各元素均为整数且处于 $[-100, 100]$ 区间，见式 (5.25) 下面一段】，都满足这种情况.

之前我们在复数情形下证明了下面结果中的最后一句话（见 8.54）. 现在我们可给出在实向量空间和复向量空间上都适用的证明.

**9.65 特征多项式、迹和行列式**

设 $T \in \mathcal{L}(V)$. 令 $n = \dim V$. 那么 $T$ 的特征多项式可写为

$$z^n - (\operatorname{tr} T)z^{n-1} + \cdots + (-1)^n(\det T).$$

**证明** 关于 $z$ 的多项式函数的常数项，就是该多项式在 $z = 0$ 时所取的值. 于是 $T$ 的特征多项式的常数项就等于 $\det(-T)$，也就等于 $(-1)^n \det T$（由 9.42 的第三例可得）.

取定 $V$ 的基，并令 $A$ 是 $T$ 关于该基的矩阵. $zI - T$ 关于该基的矩阵就是 $zI - A$. 9.46 给出的 $\det(zI - A)$ 的公式中，来自恒等排列 $\{1, \ldots, n\}$ 的项是

$$(z - A_{1,1}) \cdots (z - A_{n,n}).$$

上述表达式中 $z^{n-1}$ 的系数是 $-(A_{1,1} + \cdots + A_{n,n})$，就等于 $-\operatorname{tr} T$. $\det(zI - A)$ 的公式中来自 $\operatorname{perm} n$ 的其他各元素的项至多包含 $n - 2$ 个形如 $z - A_{k,k}$ 的因子，因此对于 $T$ 的特征多项式中 $z^{n-1}$ 的系数没有贡献. ■

在下面这条结果中，将 $n\times n$ 矩阵 $A$ 的各列看成 $\mathbf{F}^n$ 的元素．那么下面所示的范数就产生于 $\mathbf{F}^n$ 上的标准内积．回忆一下，下面证明中的记号 $R_{\cdot,k}$ 意为矩阵 $R$ 的第 $k$ 列（定义于 3.44）.

下面这条结果由雅克·阿达马（Jacques Hadamard，1865-1963）于 1893 年证明.

**9.66 阿达马不等式（Hadamard's inequality）**

设 $A$ 是 $n\times n$ 矩阵．令 $v_1,\dots,v_n$ 表示 $A$ 的各列．那么

$$|\det A|\leq\prod_{k=1}^{n}\|v_k\|.$$

**证明** 若 $A$ 不可逆，则 $\det A=0$，因此欲证不等式在此情形下成立.

于是假设 $A$ 可逆．QR 分解（7.58）告诉我们，存在一个幺正矩阵 $Q$ 和一个对角线上元素均为正数的上三角矩阵 $R$，使得 $A=QR$．我们有

$$\begin{aligned}|\det A|&=|\det Q|\,|\det R|\\&=|\det R|\\&=\prod_{k=1}^{n}R_{k,k}\\&\leq\prod_{k=1}^{n}\|R_{\cdot,k}\|\\&=\prod_{k=1}^{n}\|QR_{\cdot,k}\|\\&=\prod_{k=1}^{n}\|v_k\|,\end{aligned}$$

其中第一行源于 9.49 (b)，第二行源于 9.58，第三行源于 9.48，第五行成立则是由于 $Q$ 是等距映射. ■

下面给出阿达马不等式的几何解释．设 $\mathbf{F}=\mathbf{R}$．令 $T\in\mathcal{L}(\mathbf{R}^n)$ 是满足 $Te_k=v_k$（$k=1,\dots,n$，$e_1,\dots,e_n$ 是 $\mathbf{R}^n$ 的标准基）的算子．那么 $T$ 将长方体 $P(e_1,\dots,e_n)$ 映成平行体 $P(v_1,\dots,v_n)$【要回顾这两个记号和术语，请看 7.102 和 7.105】．因为长方体 $P(e_1,\dots,e_n)$ 的体积是 1，所以这表明（由 9.61）平行体 $P(v_1,\dots,v_n)$ 的体积是 $|\det T|$，也就等于 $|\det A|$．于是上述阿达马不等式就可解释成，在各边长为 $\|v_1\|,\dots,\|v_n\|$ 的所有平行体中，体积最大的是各边都正交的平行体（其体积为 $\prod_{k=1}^{n}\|v_k\|$）.

阿达马不等式取得等号的充要条件见于本节习题 18.

接下来这条结果中的矩阵被称为**范德蒙矩阵（Vandermonde matrix）**．范德蒙矩阵在多项式插值、离散傅里叶变换和数学中其他领域有重要应用．下面结果的证明很好地展现了在矩阵和线性映射之间来回切换的威力.

**9.67　范德蒙矩阵的行列式**

设 $n > 1$ 及 $\beta_1, \dots, \beta_n \in \mathbf{F}$. 那么

$$\det\begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{n-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & \beta_2^{n-1} \\ & & \ddots & & \\ 1 & \beta_n & \beta_n^2 & \cdots & \beta_n^{n-1} \end{pmatrix} = \prod_{1 \le j < k \le n} (\beta_k - \beta_j).$$

**证明**　令 $1, z, \dots, z^{n-1}$ 是 $\mathcal{P}_{n-1}(\mathbf{F})$ 的标准基，并令 $e_1, \dots, e_n$ 是 $\mathbf{F}^n$ 的标准基. 定义线性映射 $S: \mathcal{P}_{n-1}(\mathbf{F}) \to \mathbf{F}^n$ 为

$$Sp = \big(p(\beta_1), \dots, p(\beta_n)\big).$$

令 $A$ 表示上面结果框所示的范德蒙矩阵. 注意到

$$A = \mathcal{M}\big(S, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n)\big).$$

令 $T: \mathcal{P}_{n-1}(\mathbf{F}) \to \mathcal{P}_{n-1}(\mathbf{F})$ 是 $\mathcal{P}_{n-1}(\mathbf{F})$ 上的算子，它满足 $T1 = 1$ 以及对于 $k = 1, \dots, n-1$,

$$Tz^k = (z - \beta_1)(z - \beta_2) \cdots (z - \beta_k).$$

令 $B = \mathcal{M}\big(T, (1, z, \dots, z^{n-1}), (1, z, \dots, z^{n-1})\big)$. 那么 $B$ 是对角线上元素均为 1 的上三角矩阵. 从而 $\det B = 1$（由 9.48）.

令 $C = \mathcal{M}\big(ST, (1, z, \dots, z^{n-1}), (e_1, \dots, e_n)\big)$. 于是 $C = AB$（由 3.81），这表明

$$\det A = (\det A)(\det B) = \det C.$$

由 $C$、$S$ 和 $T$ 的定义可得，$C$ 等于

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \beta_2 - \beta_1 & 0 & \cdots & 0 \\ 1 & \beta_3 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & 0 \\ & & & \ddots & \\ 1 & \beta_n - \beta_1 & (\beta_n - \beta_1)(\beta_n - \beta_2) & \cdots & (\beta_n - \beta_1)(\beta_n - \beta_2)\cdots(\beta_n - \beta_{n-1}) \end{pmatrix}.$$

则 $\det A = \det C = \prod\limits_{1 \le j < k \le n} (\beta_k - \beta_j)$，其中我们用到了 9.56 (a) 和 9.48. ■

## 习题 9C

**1** 证明或给出一反例：$S, T \in \mathcal{L}(V) \implies \det(S + T) = \det S + \det T$.

**2** 设方阵 $A$ 的第一列除第一个元素 $A_{1,1}$ 可能不为零外全为零. 删去 $A$ 的第一行和第一列，得到矩阵 $B$. 证明：$\det A = A_{1,1} \det B$.

**3** 设 $T \in \mathcal{L}(V)$ 是幂零的. 证明：$\det(I + T) = 1$.

**4** 设 $S \in \mathcal{L}(V)$. 证明：$S$ 是幺正的，当且仅当 $|\det S| = \|S\| = 1$.

**5** 设 $A$ 是分块上三角矩阵

$$A=\begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix}.$$

其中，对角线上的每个 $A_k$ 都是方阵. 证明：

$$\det A=(\det A_1)\cdots(\det A_m).$$

**6** 设 $A=\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ 是 $n\times n$ 矩阵，其中 $v_k$ 表示 $A$ 的第 $k$ 列. 证明：如果 $(m_1,\dots,m_n)\in\operatorname{perm} n$，那么

$$\det\begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix}=\bigl(\operatorname{sign}(m_1,\dots,m_n)\bigr)\det A.$$

**7** 设 $T\in\mathcal{L}(V)$ 可逆. 令 $p$ 表示 $T$ 的特征多项式，令 $q$ 表示 $T^{-1}$ 的特征多项式. 证明：

$$q(z)=\frac{1}{p(0)}z^{\dim V}p\left(\frac{1}{z}\right)$$

对所有非零 $z\in\mathbf{F}$ 成立.

**8** 设 $T\in\mathcal{L}(V)$ 是没有特征值的算子（这暗示了 $\mathbf{F}=\mathbf{R}$）. 证明：$\det T>0$.

**9** 设 $V$ 是偶数维实向量空间，$T\in\mathcal{L}(V)$，且 $\det T<0$. 证明：$T$ 至少有两个相异特征值.

**10** 设 $V$ 是奇数维实向量空间，$T\in\mathcal{L}(V)$. 在不使用最小多项式的情况下，证明：$T$ 有特征值.

注　在不使用行列式和特征多项式的情况下，本题结论先前被证明过——见 5.34.

**11** 证明或给出一反例：如果 $\mathbf{F}=\mathbf{R}$，$T\in\mathcal{L}(V)$，$\det T>0$，那么 $T$ 有平方根.

注　如果 $\mathbf{F}=\mathbf{C}$，$T\in\mathcal{L}(V)$，$\det T\neq 0$，那么 $T$ 有平方根（见 8.41）.

**12** 设 $S,T\in\mathcal{L}(V)$ 且 $S$ 可逆. 定义 $p:\mathbf{F}\to\mathbf{F}$ 为

$$p(z)=\det(zS-T).$$

证明：$p$ 是次数为 $\dim V$ 的多项式，且该多项式中 $z^{\dim V}$ 的系数是 $\det S$.

**13** 设 $\mathbf{F}=\mathbf{C}$，$T\in\mathcal{L}(V)$，$n=\dim V>2$. 令 $\lambda_1,\dots,\lambda_n$ 表示 $T$ 的特征值，各特征值出现次数等于其重数.

(a) 求 $T$ 的特征多项式中 $z^{n-2}$ 的系数表达式（用 $\lambda_1,\dots,\lambda_n$ 表示）.

(b) 求 $T$ 的特征多项式中 $z$ 的系数表达式（用 $\lambda_1,\dots,\lambda_n$ 表示）.

**14** 设 $V$ 是内积空间，$T$ 是 $V$ 上的正算子. 证明：

$$\det\sqrt{T}=\sqrt{\det T}.$$

**15** 设 $V$ 是内积空间，$T\in\mathcal{L}(V)$. 利用极分解，证明：

$$|\det T|=\sqrt{\det(T^*T)}.$$

该证明应不同于早先给出的证明（见 9.60）.

**16** 设 $T\in\mathcal{L}(V)$. 定义 $g:\mathbf{F}\to\mathbf{F}$ 为 $g(x)=\det(I+xT)$. 证明：$g'(0)=\operatorname{tr}T$.

注　为本题找一个简洁的解法，而不用矩阵行列式的那个虽明确却复杂的公式.

**17** 设 $a,b,c$ 是正数. 求椭球

$$\left\{(x,y,z)\in\mathbf{R}^3:\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}<1\right\}$$

的体积. 按这一方式求解：找一个体积已知的集合 $\Omega \subseteq \mathbf{R}^3$，再求 $\mathbf{R}^3$ 上的一算子 $T$ 使得 $T(\Omega)$ 等于以上的椭球.

**18** 设 $A$ 是可逆方阵. 证明：阿达马不等式（9.66）取等，当且仅当 $A$ 的每一列都与其他列正交.

**19** 设 $V$ 是内积空间，$e_1,\ldots,e_n$ 是 $V$ 的规范正交基，$T \in \mathcal{L}(V)$ 是正算子.

(a) 证明：$\det T \le \prod_{k=1}^{n} \langle Te_k, e_k\rangle$.

(b) 证明：如果 $T$ 可逆；那么，(a) 中的不等式取等，当且仅当对于任一 $k = 1,\ldots,n$ 都有 $e_k$ 是 $T$ 的特征向量.

**20** 设 $A$ 是 $n\times n$ 矩阵，设 $c$ 使得对于所有 $j,k \in \{1,\ldots,n\}$ 都有 $\left|A_{j,k}\right| \le c$. 证明：

$$|\det A| \le c^n n^{n/2}.$$

注 矩阵行列式的公式（9.46）表明 $|\det A| \le c^n n!$. 然而，本题给出的估计比这好得多. 例如，如果 $c = 1$ 而 $n = 100$，那么 $c^n n! \approx 10^{158}$，但本题给出的估计值是 $10^{100}$，小得多. 如果 $n$ 是 2 的整数次幂，那么以上不等式是不能再改进的了.

**21** 设 $n$ 是正整数，$\delta : \mathbf{C}^{n,n} \to \mathbf{C}$ 是一函数，满足以下条件：

$$\delta(AB) = \delta(A)\cdot\delta(B)$$

对所有 $A, B \in \mathbf{C}^{n,n}$ 成立，且对于任一对角矩阵 $A \in \mathbf{C}^{n,n}$ 有 $\delta(A)$ 等于 $A$ 的对角线元素之积. 证明：

$$\delta(A) = \det A$$

对所有 $A \in \mathbf{C}^{n,n}$ 成立.

注 回顾一下，$\mathbf{C}^{n,n}$ 表示元素都在 $\mathbf{C}$ 中的 $n\times n$ 矩阵构成的集合. 本题表明，定义在方阵上的函数中，行列式是唯一一个可乘且在对角矩阵上具有我们所预期性质的函数. 本题和 8D 节习题 10 类似，那道题证明的是迹由它的代数性质唯一确定.

我发现，在上初等数学课时，出于教学上的考量，我越来越将行列式置于次要的地位. 我常有这样的感受：学生能娴熟地运用公式来化简冗长的表达式，却往往没能通晓这些公式的**意义**，他们只顾运算整理的各种技巧，从而没能深入钻研细节，进而就妨碍了自己达到完全的掌握.

——菲利克斯 · 克莱因（Felix Klein）《高观点下的初等数学：几何》

（*Elementary Mathematics from an Advanced Standpoint: Geometry*）

# 9D 张量积

## 两向量空间的张量积

接下来这个主题的动机，源于我们需要建立向量 $v \in V$ 和向量 $w \in W$ 的乘积. 这一乘积将表示为 $v \otimes w$，念作 "$v$ 圈乘 $w$"（"$v$ tensor $w$"）[4]，也是某个被称为 $V \otimes W$（同样念作 "$V$ 圈乘 $W$"）的新向量空间的元素.

我们已经有了向量空间 $V \times W$（见章节 3E），它被称为 $V$ 和 $W$ 的乘积. 然而，$V \times W$ 与我们此处的意图不符，因为它并没有提供一种自然的方式来让 $V$ 中的元素乘以 $W$ 中的元素. 我们希望的是，张量积满足乘法的一些常见性质. 比如，满足分配性质，意即，如果 $v_1, v_2, v \in V$ 且 $w_1, w_2, w \in W$，就有

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad 和 \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

在 TEX 中键入 \otimes 即可打出 $\otimes$.

我们还希望标量乘法和这种新的乘法"和谐共处"，意即

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$$

对所有 $\lambda \in \mathbf{F}$、$v \in V$ 和 $w \in W$ 成立.

另外，如果 $V$ 的任一基和 $W$ 的任一基相结合都可以得到 $V \otimes W$ 的基，那就好了. 具体而言，如果 $e_1, \ldots, e_m$ 是 $V$ 的基，$f_1, \ldots, f_n$ 是 $W$ 的基，那么我们希望，由 $e_j \otimes f_k$（以任意顺序）构成的组是 $V \otimes W$ 的基，其中 $j$ 取遍 1 到 $m$ 而 $k$ 取遍 1 到 $n$. 这意味着，$\dim(V \otimes W)$ 应该等于 $(\dim V)(\dim W)$. 回顾一下，$\dim(V \times W) = \dim V + \dim W$（见 3.92），这表明 $V \times W$ 这一乘积与我们此处的意图不符.

为了用一种自然的方式由 $V$ 和 $W$ 得到维数为 $(\dim V)(\dim W)$ 的向量空间，我们关注双线性泛函构成的向量空间，其定义如下.

**9.68 定义：$V \times W$ 上的双线性泛函（bilinear functional）、向量空间 $\mathcal{B}(V, W)$**

- $V \times W$ 上的**双线性泛函**是函数 $\beta : V \times W \to \mathbf{F}$，使得：对于任一 $w \in W$，$v \mapsto \beta(v, w)$ 都是 $V$ 上的线性泛函；对于任一 $v \in V$，$w \mapsto \beta(v, w)$ 都是 $W$ 上的线性泛函.
- $V \times W$ 上的双线性泛函构成的向量空间，记为 $\mathcal{B}(V, W)$.

♣

如果 $W = V$，那么 $V \times W$ 上的双线性泛函就是双线性型了（见 9.1）.

$\mathcal{B}(V, W)$ 上的加法和数乘运算，定义为函数通常的加法和标量乘法. 你可以验证，这些运算使 $\mathcal{B}(V, W)$ 成为向量空间，其加法恒等元是从 $V \times W$ 到 $\mathbf{F}$ 的零函数.

**9.69 例：双线性泛函**

- 设 $\varphi \in V'$ 且 $\tau \in W'$. 定义 $\beta : V \times W \to \mathbf{F}$ 为 $\beta(v, w) = \varphi(v)\tau(w)$. 那么 $\beta$ 是 $V \times W$ 上的双线性泛函.

转下页

[4]念作"圈乘"，是为了和"点乘"、"叉乘"等相一致. 也可以念作"$v$ 张量乘 $w$"，与 tensor 一词直接对应，略微拗口但更加精确.

- 设 $v \in V$ 且 $w \in W$. 定义 $\beta : V' \times W' \to \mathbf{F}$ 为 $\beta(\varphi, \tau) = \varphi(v)\tau(w)$. 那么 $\beta$ 是 $V' \times W'$ 上的双线性泛函.
- 定义 $\beta : V \times V' \to \mathbf{F}$ 为 $\beta(v, \varphi) = \varphi(v)$. 那么 $\beta$ 是 $V \times V'$ 上的双线性泛函.
- 设 $\varphi \in V'$. 定义 $\beta : V \times \mathcal{L}(V) \to \mathbf{F}$ 为 $\beta(v, T) = \varphi(Tv)$. 那么 $\beta$ 是 $V \times \mathcal{L}(V)$ 上的双线性泛函.
- 设 $m$ 和 $n$ 是正整数. 定义 $\beta : \mathbf{F}^{m,n} \times \mathbf{F}^{n,m} \to \mathbf{F}$ 为 $\beta(A, B) = \operatorname{tr}(AB)$. 那么 $\beta$ 是 $\mathbf{F}^{m,n} \times \mathbf{F}^{n,m}$ 上的双线性泛函.

**9.70 双线性泛函构成的向量空间的维数**

$\dim \mathcal{B}(V, W) = (\dim V)(\dim W)$.

**证明** 令 $e_1, \ldots, e_m$ 是 $V$ 的基, $f_1, \ldots, f_n$ 是 $W$ 的基. 对于双线性泛函 $\beta \in \mathcal{B}(V, W)$, 令 $\mathcal{M}(\beta)$ 是第 $j$ 行第 $k$ 列元素为 $\beta(e_j, f_k)$ 的 $m \times n$ 矩阵. $\beta \mapsto \mathcal{M}(\beta)$ 这一映射是从 $\mathcal{B}(V, W)$ 到 $\mathbf{F}^{m,n}$ 的线性映射.

对于矩阵 $C \in \mathbf{F}^{m,n}$, 在 $V \times W$ 上定义双线性泛函 $\beta_C$ 为

$$\beta_C(a_1 e_1 + \cdots + a_m e_m, b_1 f_1 + \cdots + b_n f_n) = \sum_{k=1}^{n} \sum_{j=1}^{m} C_{j,k} a_j b_k,$$

其中 $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbf{F}$.

从 $\mathcal{B}(V, W)$ 到 $\mathbf{F}^{m,n}$ 的线性映射 $\beta \mapsto \mathcal{M}(\beta)$, 与从 $\mathbf{F}^{m,n}$ 到 $\mathcal{B}(V, W)$ 的线性映射 $C \mapsto \beta_C$ 互为彼此的逆. 这是因为, 对所有 $\beta \in \mathcal{B}(V, W)$ 有 $\beta_{\mathcal{M}(\beta)} = \beta$, 而对所有 $C \in \mathbf{F}^{m,n}$ 有 $\mathcal{M}(\beta_C) = C$ (你应该自行验证).

因此, 这两个映射都是同构, 而且它们所关联的这两个空间维数相同. 于是, $\dim \mathcal{B}(V, W) = \dim \mathbf{F}^{m,n} = mn = (\dim V)(\dim W)$. ■

在数学文献中, 也会出现 $V \otimes W$ 的一些其他定义. 至少在有限维的语境下, 这些定义是彼此等价的, 因为维数相同的任意两个向量空间都是同构的.

以上结果指出, $\mathcal{B}(V, W)$ 的维数正是我们所寻求的, $\mathcal{L}(V, W)$ 和 $\mathbf{F}^{\dim V, \dim W}$ 的也是. 这就可能诱使我们把 $V \otimes W$ 定义为 $\mathcal{B}(V, W)$、$\mathcal{L}(V, W)$ 或者 $\mathbf{F}^{\dim V, \dim W}$. 然而, 这些定义都不能引出 $v \otimes w$ ($v \in V, w \in W$) 的一个与基无关的定义.

接下来这条定义, 尽管乍一看有点奇怪而抽象, 但是有个巨大的优势, 就是以一种与基无关的方式定义了 $v \otimes w$. 我们将 $V \otimes W$ 定义为在 $V' \times W'$ 上的双线性泛函构成的向量空间, 而不是在 $V \times W$ 上的 (尽管你可能忍不住想这样定义).

**9.71 定义: 张量积 (tensor product)、$V \otimes W$、$v \otimes w$**

- **张量积** $V \otimes W$ 定义为 $\mathcal{B}(V', W')$.
- 对于 $v \in V$ 和 $w \in W$, **张量积** $v \otimes w$ 是 $V \otimes W$ 的元素[5], 定义为

$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w)$$

对所有 $(\varphi, \tau) \in V' \times W'$ 成立.

[5]注意, 这里并没有说 $V \otimes W$ 的元素都可以写成 $v \otimes w$ 的形式.

我们可以很快地证明 $V \otimes W$ 的这一定义能让它有我们想要的维数.

**9.72 两向量空间的张量积的维数**

$\dim(V \otimes W) = (\dim V)(\dim W)$.

**证明** 因为向量空间和它的对偶维数相同（根据 3.111），所以我们有 $\dim V' = \dim V$ 和 $\dim W' = \dim W$. 因此，9.70 告诉我们 $\mathcal{B}(V', W')$ 的维数等于 $(\dim V)(\dim W)$. ■

为了理解两向量 $v \in V$ 和 $w \in W$ 的张量积 $v \otimes w$ 的定义，就要关注 $v \otimes w$ 究竟是何种类型的对象. $V \otimes W$ 的元素是 $V' \times W'$ 上的双线性泛函，特别地，它是从 $V' \times W'$ 到 $\mathbf{F}$ 的函数. 因此，输入 $V' \times W'$ 中的每个元素，它都输出 $\mathbf{F}$ 中的一个元素. 以上定义就具有这样的性质，因为将 $v \otimes w$ 作用于 $V' \times W'$ 的一个元素 $(\varphi, \tau)$，就能得到 $\varphi(v)\tau(w)$ 这个数.

$v \otimes w$ 这有点抽象的本质，应该无关紧要. 重点在于这些对象的性质. 接下来这条结果表明，向量的张量积具有我们预期的双线性性质.

**9.73 张量积的双线性**

设 $v, v_1, v_2 \in V$，$w, w_1, w_2 \in W$，$\lambda \in \mathbf{F}$. 那么就有

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad \text{和} \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$$

以及

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w).$$

**证明** 设 $(\varphi, \tau) \in V' \times W'$. 那么

$$\begin{aligned}
\big((v_1 + v_2) \otimes w\big)(\varphi, \tau) &= \varphi(v_1 + v_2)\tau(w) \\
&= \varphi(v_1)\tau(w) + \varphi(v_2)\tau(w) \\
&= (v_1 \otimes w)(\varphi, \tau) + (v_2 \otimes w)(\varphi, \tau) \\
&= (v_1 \otimes w + v_2 \otimes w)(\varphi, \tau).
\end{aligned}$$

因此，$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$.

另外两个等式可类似这样证明. ■

根据定义，“组”是有序的. 顺序有时很重要，比如组成一算子关于一个基的矩阵时. 而对于本节里这种有两个下标的组，比如接下来这条结果中的 $\{e_j \otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$，顺序就无关紧要了，我们也不做具体说明——只是怎么方便怎么排.

下面这条结果的 (a) 中，$V \otimes W$ 中元素的线性无关性，揭示了这一观点：$V \otimes W$ 中的向量之间的关系，要么是来自张量积的双线性（见 9.73），要么可能是由于 $V$ 或 $W$ 中的向量组的线性相关性，此外就再没有别的了.

**9.74 $V \otimes W$ 的基**

设 $e_1, \ldots, e_m$ 是 $V$ 中的一组向量，$f_1, \ldots, f_n$ 是 $W$ 中的一组向量.

(a) 如果 $e_1, \ldots, e_m$ 和 $f_1, \ldots, f_n$ 都是线性无关组，那么

$$\{e_j \otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$$

是 $V \otimes W$ 中的线性无关组.

转下页

(b) 如果 $e_1,\ldots,e_m$ 是 $V$ 的基，$f_1,\ldots,f_n$ 是 $W$ 的基，那么组 $\{e_j\otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$ 是 $V\otimes W$ 的基.

**证明** 为证明 (a)，设 $e_1,\ldots,e_m$ 和 $f_1,\ldots,f_n$ 都是线性无关组. 由这一线性无关性以及线性映射引理（3.4）可得，存在 $\varphi_1,\ldots,\varphi_m\in V'$ 和 $\tau_1,\ldots,\tau_n\in W'$ 使得

$$\varphi_j(e_k)=\begin{cases}1 & 若\ j=k,\\ 0 & 若\ j\neq k\end{cases}\qquad 和 \qquad \tau_j(f_k)=\begin{cases}1 & 若\ j=k,\\ 0 & 若\ j\neq k,\end{cases}$$

其中第一个式子中 $j,k\in\{1,\ldots,m\}$，第二个式子中 $j,k\in\{1,\ldots,n\}$.

设 $\{a_{j,k}\}_{j=1,\ldots,m;k=1,\ldots,n}$ 是一组标量，使得

$$\sum_{k=1}^{n}\sum_{j=1}^{m}a_{j,k}(e_j\otimes f_k)=0. \tag{9.75}$$

注意到，$(e_j\otimes f_k)(\varphi_M,\tau_N)$ 在 $j=M$ 且 $k=N$ 情况下等于 1，其余情况下等于 0. 那么，将 9.75 两侧都作用于 $(\varphi_M,\tau_N)$，就得到 $a_{M,N}=0$，证明了 $\{e_j\otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$ 线性无关.

这样一来，根据 (a) 和 $\dim(V\otimes W)=(\dim V)(\dim W)$【见 9.72】，以及“长度恰好的线性无关组是一个基”这一结果（见 2.38），可证得 (b). ■

根据以上结果中的 (b)，$V\otimes W$ 的每个元素都是形如 $v\otimes w$ 的元素的有限和，其中 $v\in V, w\in W$. 然而，如果 $\dim V>1$ 且 $\dim W>1$，那么由习题 4 可知

$$\{v\otimes w:(v,w)\in V\times W\}\neq V\otimes W.$$

**9.76 例：$\mathbf{F}^m$ 的元素和 $\mathbf{F}^n$ 的元素所成的张量积**

设 $m$ 和 $n$ 是正整数. 令 $e_1,\ldots,e_m$ 表示 $\mathbf{F}^m$ 的标准基，令 $f_1,\ldots,f_n$ 表示 $\mathbf{F}^n$ 的标准基. 设

$$v=(v_1,\ldots,v_m)\in\mathbf{F}^m \quad 和 \quad w=(w_1,\ldots,w_n)\in\mathbf{F}^n.$$

那么

$$\begin{aligned}v\otimes w&=\left(\sum_{j=1}^{m}v_je_j\right)\otimes\left(\sum_{k=1}^{n}w_kf_k\right)\\&=\sum_{k=1}^{n}\sum_{j=1}^{m}(v_jw_k)(e_j\otimes f_k).\end{aligned}$$

因此，关于 $\mathbf{F}^m\otimes\mathbf{F}^n$ 的【根据 9.74 (b) 得到的】基 $\{e_j\otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$，$v\otimes w$ 的系数是 $\{v_jw_k\}_{j=1,\ldots,m;k=1,\ldots,n}$. 如果我们不把这些数写成组，而是写成一个 $m\times n$ 矩阵，其中第 $j$ 行第 $k$ 列元素为 $v_jw_k$，那么我们就可以把 $v\otimes w$ 等同为这个 $m\times n$ 矩阵：

$$\begin{pmatrix}v_1w_1 & \cdots & v_1w_n\\ & \ddots & \\ v_mw_1 & \cdots & v_mw_n\end{pmatrix}.$$

以上例子指出了一种等同关系，你可通过本节习题 5、6 来练习如何运用这种等同关系.

我们现在定义双线性映射. 不同于双线性泛函，双线性映射的目标空间可以是任一向量空间，而不只是标量域.

**9.77 定义：双线性映射（bilinear map）**

从 $V\times W$ 到向量空间 $U$ 的**双线性映射**是这样一个函数 $\Gamma: V\times W \to U$，其使得 $v \mapsto \Gamma(v,w)$ 对任一 $w \in W$ 都是从 $V$ 到 $U$ 的线性映射，$w \mapsto \Gamma(v,w)$ 对任一 $v \in V$ 都是从 $W$ 到 $U$ 的线性映射.

**9.78 例：双线性映射**

- $V \times W$ 上的每个双线性泛函都是从 $V \times W$ 到 $\mathbf{F}$ 的双线性映射.
- 定义为 $\Gamma(v,w) = v \otimes w$ 的函数 $\Gamma: V \times W \to V \otimes W$ 是从 $V \times W$ 到 $V \otimes W$ 的双线性映射（根据 9.73）.
- 定义为 $\Gamma(S,T) = ST$ 的函数 $\Gamma: \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V)$ 是从 $\mathcal{L}(V) \times \mathcal{L}(V)$ 到 $\mathcal{L}(V)$ 的双线性映射.
- 定义为 $\Gamma(v,T) = Tv$ 的函数 $\Gamma: V \times \mathcal{L}(V,W) \to W$ 是从 $V \times \mathcal{L}(V,W)$ 到 $W$ 的双线性映射.

接下来这条结果表明，张量积使我们能够将 $V \times W$ 上的双线性映射转换为 $V \otimes W$ 上的线性映射（反之亦然）. 在数学文献中，下面结果的 (a) 称为张量积的“泛性质”（universal property）.

**9.79 化双线性映射为线性映射**

设 $U$ 是向量空间.

(a) 设 $\Gamma: V \times W \to U$ 是双线性映射，那么存在唯一的线性映射 $\hat{\Gamma}: V \otimes W \to U$ 使得

$$\hat{\Gamma}(v \otimes w) = \Gamma(v,w)$$

对所有 $(v,w) \in V \times W$ 成立.

(b) 反之，设 $T: V \otimes W \to U$ 是线性映射，那么存在唯一的双线性映射 $T^{\#}: V \times W \to U$ 使得

$$T^{\#}(v,w) = T(v \otimes w)$$

对所有 $(v,w) \in V \times W$ 成立.

**证明** 令 $e_1,\ldots,e_m$ 是 $V$ 的基，令 $f_1,\ldots,f_n$ 是 $W$ 的基. 根据线性映射引理（3.4）以及 9.74 (b)，存在唯一的线性映射 $\hat{\Gamma}: V \otimes W \to U$ 使得

$$\hat{\Gamma}(e_j \otimes f_k) = \Gamma(e_j, f_k)$$

对所有 $j \in \{1,\ldots,m\}$ 和 $k \in \{1,\ldots,n\}$ 成立.

现在设 $(v,w) \in V \times W$. 存在 $a_1,\ldots,a_m,b_1,\ldots,b_n \in \mathbf{F}$ 使得 $v = a_1e_1 + \cdots + a_me_m$ 而 $w = b_1f_1 + \cdots + b_nf_n$. 因此

$$\begin{aligned}\hat{\Gamma}(v\otimes w) &= \hat{\Gamma}\left(\sum_{k=1}^{n}\sum_{j=1}^{m}(a_j b_k)(e_j\otimes f_k)\right)\\ &= \sum_{k=1}^{n}\sum_{j=1}^{m}a_j b_k\hat{\Gamma}(e_j\otimes f_k)\\ &= \sum_{k=1}^{n}\sum_{j=1}^{m}a_j b_k\Gamma(e_j, f_k)\\ &= \Gamma(v,w),\end{aligned}$$

即待证等式成立，其中第二行成立是因为 $\hat{\Gamma}$ 是线性的，第三行成立是因为 $\hat{\Gamma}$ 的定义，第四行成立是因为 $\Gamma$ 是双线性的.

根据 9.74 (b)，可得满足 $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ 的线性映射 $\hat{\Gamma}$ 具有唯一性，也就完成了 (a) 的证明.

为证明 (b)，定义函数 $T^{\#}: V\times W\to U$ 为 $T^{\#}(v,w)=T(v\otimes w)$ 对所有 $(v,w)\in V\times W$ 成立. 由张量积的双线性（见 9.73）和 $T$ 的线性，可得 $T^{\#}$ 是双线性的.

显然，满足条件的 $T^{\#}$ 是唯一的. ■

要证明 9.79 (a)，我们不能只简单地定义 $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ 对所有 $v\in V$ 和 $w\in W$ 成立（然后将 $\hat{\Gamma}$ “线性地”推广至整个 $V\otimes W$ 上），因为 $V\otimes W$ 的元素并不能唯一地表示成形如 $v\otimes w$ 的元素的有限和. 我们的证明用了 $V$ 的基和 $W$ 的基以避免这个问题.

虽然我们在 9.79 (a) 的证明中对 $\hat{\Gamma}$ 的构造依赖于 $V$ 和 $W$ 的基，但是 $\hat{\Gamma}(v\otimes w)=\Gamma(v,w)$ 对所有 $v\in V$ 和 $w\in W$ 成立，这又表明 $\hat{\Gamma}$ 不依赖于 $V$ 和 $W$ 的基的选取.

## 内积空间的张量积

下面的结果包含三种内积——$V\otimes W$ 上的、$V$ 上的和 $W$ 上的. 但是我们都用相同的符号 $\langle\cdot,\cdot\rangle$ 来表示.

**9.80 两内积空间所成张量积上的内积**

设 $V$ 和 $W$ 是内积空间. 那么，$V\otimes W$ 上存在唯一一种内积使得

$$\langle v\otimes w, u\otimes x\rangle=\langle v,u\rangle\langle w,x\rangle$$

对所有 $v,u\in V$ 和 $w,x\in W$ 成立.

**证明** 设 $e_1,\ldots,e_m$ 是 $V$ 的规范正交基，$f_1,\ldots,f_n$ 是 $W$ 的规范正交基. 在 $V\otimes W$ 上定义一内积为

$$\left\langle\sum_{k=1}^{n}\sum_{j=1}^{m}b_{j,k}e_j\otimes f_k, \sum_{k=1}^{n}\sum_{j=1}^{m}c_{j,k}e_j\otimes f_k\right\rangle=\sum_{k=1}^{n}\sum_{j=1}^{m}b_{j,k}\overline{c_{j,k}}. \tag{9.81}$$

验证式 (9.81) 定义了 $V\otimes W$ 上的一内积是很简单的【利用 9.74 (b)】，这就留给读者了.

设 $v,u\in V$，$w,x\in W$. 令 $v_1,\ldots,v_m\in\mathbf{F}$ 使得 $v=v_1e_1+\cdots+v_me_m$，而 $u$、$w$ 和 $x$ 的表达式以此类推. 那么

$$
\begin{aligned}
\langle v\otimes w, u\otimes x\rangle &= \left\langle \sum_{j=1}^{m} v_j e_j \otimes \sum_{k=1}^{n} w_k f_k, \sum_{j=1}^{m} u_j e_j \otimes \sum_{k=1}^{n} x_k f_k \right\rangle \\
&= \left\langle \sum_{k=1}^{n}\sum_{j=1}^{m} v_j w_k e_j \otimes f_k, \sum_{k=1}^{n}\sum_{j=1}^{m} u_j x_k e_j \otimes f_k \right\rangle \\
&= \sum_{k=1}^{n}\sum_{j=1}^{m} v_j \overline{u_j} w_k \overline{x_k} \\
&= \left(\sum_{j=1}^{m} v_j \overline{u_j}\right)\left(\sum_{k=1}^{n} w_k \overline{x_k}\right) \\
&= \langle v, u\rangle \langle w, x\rangle.
\end{aligned}
$$

因为 $V\otimes W$ 的每个元素都可以写成形如 $v\otimes w$ 的元素的线性组合【根据 9.74 (b)】，所以 $V\otimes W$ 上只有一种内积使得 $\langle v\otimes w, u\otimes x\rangle = \langle v,u\rangle\langle w,x\rangle$ 对所有 $v,u\in V$ 和 $w,x\in W$ 成立. ■

下面定义 $V\otimes W$ 上一自然的内积，而 9.80 已经证明了这一定义的合理性. 如果不加证明的话，我们就不能简单地定义 $\langle v\otimes w, u\otimes x\rangle$ 为 $\langle v,u\rangle\langle w,x\rangle$（然后分别用各位置的可加性将定义推广至 $V\otimes W$ 上），因为 $V\otimes W$ 的元素并不能唯一地表示为形如 $v\otimes w$ 的元素的有限和.

**9.82 定义：两内积空间所成张量积上的内积（inner product on tensor product of two inner product spaces）**

设 $V$ 和 $W$ 是内积空间. $V\otimes W$ 上的内积是唯一使得

$$\langle v\otimes w, u\otimes x\rangle = \langle v,u\rangle\langle w,x\rangle$$

对所有 $v,u\in V$ 和 $w,x\in W$ 成立的，从 $(V\otimes W)\times(V\otimes W)$ 到 $\mathbf{F}$ 的函数 $\langle\cdot,\cdot\rangle$. ♣

在上式中取 $u=v$ 和 $x=w$，再开平方根，可得

$$\|v\otimes w\| = \|v\|\,\|w\|$$

对所有 $v\in V$ 和所有 $w\in W$ 成立.

在 9.80 的证明中，内积的构造依赖于 $V$ 的规范正交基 $e_1,\dots,e_m$ 和 $W$ 的规范正交基 $f_1,\dots,f_n$. 式 (9.81) 意味着，$\{e_j\otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$ 是 $V\otimes W$ 的双下标规范正交组，从而是 $V\otimes W$ 的规范正交基【根据 9.74 (b)】. 接下来这条结果之所以重要，是因为其中所用的规范正交基，可以不同于 9.80 中定义内积所用的规范正交基. 虽然在 9.80 的证明和下面这一结果中这些基的记号是一样的，但是要将它们视为两组不同的规范正交基. [6]

**9.83 $V\otimes W$ 的规范正交基**

设 $V$ 和 $W$ 是内积空间，且 $e_1,\dots,e_m$ 是 $V$ 的一个规范正交基，$f_1,\dots,f_m$ 是 $W$ 的一个规范正交基. 那么

$$\{e_j\otimes f_k\}_{j=1,\dots,m;k=1,\dots,n}$$

是 $V\otimes W$ 的一个规范正交基. ♡

[6]换句话说，不要误以为 9.83 只适用于 9.80 中定义内积时所用的那一对规范正交基.

**证明** 我们已经知道，$\{e_j \otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$ 是 $V \otimes W$ 的基【根据 9.74 (b)】. 所以我们只需要验证规范正交性. 为此，设 $j, M \in \{1,\ldots,m\}$ 和 $k, N \in \{1,\ldots,n\}$. 那么

$$\langle e_j \otimes f_k, e_M \otimes f_N \rangle = \langle e_j, e_M \rangle \langle f_k, f_N \rangle = \begin{cases} 1 & \text{若 } j = M \text{ 且 } k = N, \\ 0 & \text{其他.} \end{cases}$$

因此，双下标组 $\{e_j \otimes f_k\}_{j=1,\ldots,m;k=1,\ldots,n}$ 确实是 $V \otimes W$ 的一个规范正交基. ■

关于 $V \otimes W$ 上的内积结构如何与 $V$ 和 $W$ 上的算子相互作用的一例，见习题 11.

## 多个向量空间的张量积

两个有限维向量空间所成张量积的性质，我们已经讨论过了. 现在我们将注意力转向，多个有限维向量空间所成的张量积. 这一推广不需要什么新观点，只需要一些稍微更复杂的记号. 读者如果能很好地理解两个向量空间的张量积，就应该能进一步理解多于两个向量空间的张量积.

因此，这一小节不给出证明. 所提供的定义和结果表述，应该足以让读者用已学过的关于两向量空间所成张量积的知识填充其中的细节.

首先是以下的记号假设.

**9.84 记号：$V_1, \ldots, V_m$**

本小节剩余部分中，$m$ 表示大于 1 的整数，而 $V_1, \ldots, V_m$ 都表示有限维向量空间. ♣

我们将要定义 $m$ 重线性泛函，这一概念推广了双线性泛函（见 9.68）. 回顾一下，“泛函”一词指的是映射到标量域 $\mathbf{F}$，以及，“$m$ 重线性型”这一术语用于特殊情形 $V_1 = \cdots = V_m$（见 9.25）. $\mathcal{B}(V_1, \ldots, V_m)$ 这一记号推广了我们原先的 $\mathcal{B}(V, W)$.

**9.85 定义：$m$ 重线性泛函（$m$-linear functional）、向量空间 $\mathcal{B}(V_1, \ldots, V_m)$**

- $V_1 \times \cdots \times V_m$ 上的 $m$ 重线性泛函是满足如下性质的函数 $\beta: V_1 \times \cdots \times V_m \to \mathbf{F}$：取 $\beta(\cdot, \ldots, \cdot)$ 中任一位置并将该位置之外的向量固定，则 $\beta$ 都将成为该位置上的线性泛函.
- $V_1 \times \cdots \times V_m$ 上的 $m$ 重线性泛函构成的向量空间，记为 $\mathcal{B}(V_1, \ldots, V_m)$. ♣

---

**9.86 例：$m$ 重线性泛函**

对于每个 $k \in \{1, \ldots, m\}$，设 $\varphi_k \in V_k{}'$. 定义 $\beta: V_1 \times \cdots \times V_m \to \mathbf{F}$ 为

$$\beta(v_1, \ldots, v_m) = \varphi_1(v_1) \times \cdots \times \varphi_m(v_m).$$

那么 $\beta$ 是 $V_1 \times \cdots \times V_m$ 上的 $m$ 重线性泛函.

---

接下来这条结果，可以仿照 9.70 的证明方法来证.

**9.87 $m$ 重线性泛函所成向量空间的维数**

$\dim \mathcal{B}(V_1, \ldots, V_m) = (\dim V_1) \times \cdots \times (\dim V_m).$ ♡

现在我们可以定义，多个向量空间的张量积，以及这些向量空间的元素的张量积. 以下定义和我们先前在 $m=2$ 情形下所做的定义（9.71）完全类似.

**9.88 定义：张量积（tensor product）、$V_1\otimes\cdots\otimes V_m$、$v_1\otimes\cdots\otimes v_m$**

- **张量积** $V_1\otimes\cdots\otimes V_m$ 定义为 $\mathcal{B}(V_1',\ldots,V_m')$.
- 对于 $v_1\in V_1,\ldots,v_m\in V_m$，**张量积** $v_1\otimes\cdots\otimes v_m$ 是 $V_1\otimes\cdots\otimes V_m$ 的元素，定义为

$$(v_1\otimes\cdots\otimes v_m)(\varphi_1,\ldots,\varphi_m)=\varphi_1(v_1)\cdots\varphi_m(v_m)$$

  对所有 $(\varphi_1,\ldots,\varphi_m)\in V_1'\times\cdots\times V_m'$ 成立.

♣

接下来这条结果的证明，可以参照 $m=2$ 情形下的类似结果的证明方式（见 9.72）.

**9.89 张量积的维数**

$\dim(V_1\otimes\cdots\otimes V_m)=(\dim V_1)\cdots(\dim V_m)$.

♡

接下来这条结果推广了 9.74.

**9.90 $V_1\otimes\cdots\otimes V_m$ 的基**

设 $\dim V_k=n_k$ 且 $e_1^k,\ldots,e_{n_k}^k$ 是 $V_k$ 的基（$k=1,\ldots,m$）. 那么

$$\{e_{j_1}^1\otimes\cdots\otimes e_{j_m}^m\}_{j_1=1,\ldots,n_1;\cdots;j_m=1,\ldots,n_m}$$

是 $V_1\otimes\cdots\otimes V_m$ 的基.

♡

设 $m=2$，$e_1^1,\ldots,e_{n_1}^1$ 是 $V_1$ 的基，$e_1^2,\ldots,e_{n_2}^2$ 是 $V_2$ 的基. 那么，$V_1\otimes V_2$ 的元素，关于以上结果中的基 $\{e_{j_1}^1\otimes e_{j_2}^2\}_{j_1=1,\ldots,n_1;j_2=1,\ldots,n_2}$ 的系数，可以用 $n_1\times n_2$ 矩阵表示，其中第 $j_1$ 行第 $j_2$ 列的元素是 $e_{j_1}^1\otimes e_{j_2}^2$ 的系数. 因此，为了表示 $V_1\otimes V_2$ 的元素，我们需要一个矩阵——由两个下标规定的数组.

如果 $m>2$，那么以上结果就表明我们需要一个由 $m$ 个下标规定的数组，来表示 $V_1\otimes\cdots\otimes V_m$ 的一个任意元素. 所以，当我们处理的对象由包含多个下标的数组指定时，可能就会出现张量积.

接下来这条定义推广了双线性映射的概念（见 9.77）. 和双线性映射一样，其目标空间也是任意向量空间.

**9.91 定义：$m$ 重线性映射（$m$-linear map）**

从 $V_1\times\cdots\times V_m$ 到向量空间 $U$ 的$m$ **重线性映射**是满足如下性质的函数 $\Gamma: V_1\times\cdots\times V_m\to U$：固定 $\Gamma(\cdot,\ldots,\cdot)$ 中任一位置以外的向量，都能得到该位置上的线性映射.

♣

接下来这条结果，可以参照 9.79 的证明方式来证.

**9.92 化 $m$ 重线性映射为线性映射**

设 $U$ 是向量空间.

(a) 设 $\Gamma: V_1\times\cdots\times V_m \to U$ 是 $m$ 重线性映射,那么存在唯一的线性映射 $\hat{\Gamma}: V_1\otimes\cdots\otimes V_m \to U$ 使得

$$\hat{\Gamma}(v_1\otimes\cdots\otimes v_m)=\Gamma(v_1,\ldots,v_m)$$

对所有 $(v_1,\ldots,v_m)\in V_1\times\cdots\times V_m$ 成立.

(b) 反之,设 $T: V_1\otimes\cdots\otimes V_m \to U$ 是线性映射,那么存在唯一的 $m$ 重线性映射 $T^{\#}$ : $V_1\times\cdots\times V_m \to U$ 使得

$$T^{\#}(v_1,\ldots,v_m)=T(v_1\otimes\cdots\otimes v_m)$$

对所有 $(v_1,\ldots,v_m)\in V_1\times\cdots\times V_m$ 成立.

多个内积空间的张量积,见本节习题 12、13.

## 习题 9D

**1** 设 $v\in V$ 和 $w\in W$. 证明:$v\otimes w=0$,当且仅当 $v=0$ 或 $w=0$.

**2** 给出一例:$\mathbf{R}^3$ 中六个不同的向量 $v_1,v_2,v_3,w_1,w_2,w_3$,使得

$$v_1\otimes w_1+v_2\otimes w_2+v_3\otimes w_3=0,$$

而 $v_1\otimes w_1, v_2\otimes w_2, v_3\otimes w_3$ 当中,任一元素都不是该组中另一元素的标量倍.

**3** 设 $v_1,\ldots,v_m$ 是 $V$ 中的线性无关组. 又设 $w_1,\ldots,w_m$ 是 $W$ 中的组,使得

$$v_1\otimes w_1+\cdots+v_m\otimes w_m=0.$$

证明:$w_1=\cdots=w_m=0$.

**4** 设 $\dim V>1$ 且 $\dim W>1$. 证明:

$$\{v\otimes w:(v,w)\in V\times W\}$$

不是 $V\otimes W$ 的子空间.

注 本题意味着,如果 $\dim V>1$ 且 $\dim W>1$,那么

$$\{v\otimes w:(v,w)\in V\times W\}\neq V\otimes W.$$

**5** 设 $m$ 和 $n$ 是正整数. 对于 $v\in\mathbf{F}^m$ 和 $w\in\mathbf{F}^n$,按照例 9.76,将 $v\otimes w$ 等同于一个 $m\times n$ 矩阵. 在这一等同关系下,证明:

$$\left\{v\otimes w: v\in\mathbf{F}^m \text{ 且 } w\in\mathbf{F}^n\right\}$$

是秩至多为一的 $m\times n$ 矩阵(其元素在 $\mathbf{F}$ 中)构成的集合.

**6** 设 $m$ 和 $n$ 是正整数. 类比第 5 题,描述秩至多为二的 $m\times n$ 矩阵(其元素在 $\mathbf{F}$ 中)构成的集合.

**7** 设 $\dim V>2$ 且 $\dim W>2$. 证明:

$$\left\{v_1\otimes w_1+v_2\otimes w_2: v_1,v_2\in V \text{ 且 } w_1,w_2\in W\right\}\neq V\otimes W.$$

**8** 设 $v_1,\ldots,v_m\in V$ 且 $w_1,\ldots,w_m\in W$，使得

$$v_1\otimes w_1+\cdots+v_m\otimes w_m=0.$$

设 $U$ 是向量空间，$\Gamma:V\times W\to U$ 是双线性映射．证明：

$$\Gamma(v_1,w_1)+\cdots+\Gamma(v_m,w_m)=0.$$

**9** 设 $S\in\mathcal{L}(V)$ 且 $T\in\mathcal{L}(W)$．证明：$V\otimes W$ 上存在唯一一个这样的算子：对于所有 $v\in V$ 和 $w\in W$，将 $v\otimes w$ 映成 $Sv\otimes Tw$.

注　本题给出的 $V\otimes W$ 上这一算子也经常被称为 $S\otimes T$（这里滥用了记号[7]）.

**10** 设 $S\in\mathcal{L}(V)$ 且 $T\in\mathcal{L}(W)$．证明：$S\otimes T$ 是 $V\otimes W$ 上的可逆算子，当且仅当 $S$ 和 $T$ 都是可逆算子．并证明：如果 $S$ 和 $T$ 都是可逆算子，那么 $(S\otimes T)^{-1}=S^{-1}\otimes T^{-1}$（这里我们用到了第 9 题小注部分提到的记号）.

**11** 设 $V$ 和 $W$ 是内积空间．证明：如果 $S\in\mathcal{L}(V)$ 且 $T\in\mathcal{L}(W)$，那么 $(S\otimes T)^*=S^*\otimes T^*$（这里我们用到了第 9 题小注部分提到的记号）.

**12** 设 $V_1,\ldots,V_m$ 是有限维内积空间．证明：$V_1\otimes\cdots\otimes V_m$ 上存在唯一一种内积，使得

$$\langle v_1\otimes\cdots\otimes v_m,u_1\otimes\cdots\otimes u_m\rangle=\langle v_1,u_1\rangle\cdots\langle v_m,u_m\rangle$$

对 $V_1\times\cdots\times V_m$ 中所有 $(v_1,\ldots,v_m)$ 和 $(u_1,\ldots,u_m)$ 成立.

注　注意，上式意味着

$$\|v_1\otimes\cdots\otimes v_m\|=\|v_1\|\times\cdots\times\|v_m\|$$

对所有 $(v_1,\ldots,v_m)\in V_1\times\cdots\times V_m$ 成立.

**13** 设 $V_1,\ldots,V_m$ 是有限维内积空间，并用第 12 题中定义的内积，使 $V_1\otimes\cdots\otimes V_m$ 成为内积空间．设对任一 $k=1,\ldots,m$，都有 $e_1^k,\ldots,e_{n_k}^k$ 是 $V_k$ 的规范正交基．证明：向量组

$$\left\{e_{j_1}^1\otimes\cdots\otimes e_{j_m}^m\right\}_{j_1=1,\ldots,n_1;\cdots;j_m=1,\ldots,n_m}$$

是 $V_1\otimes\cdots\otimes V_m$ 的规范正交基.

[7]为什么说“这里滥用了记号”呢？因为按照正文中给出的定义，“某 $\otimes$ 某”（这里的“某”指的是某向量，而线性映射本身当然也是向量）得到的是双线性泛函，但这里却用 $S\otimes T$ 表示算子．记号滥用其实有损于严谨性和准确性，但有时也便于表达和理解，比如接下来几题就用 $S\otimes T$ 等避免了冗长的描述，还不用另外定义记号.

# 图片来源

- 第 ii 页：照片来自 Carrie Heeter 和 Bishnu Sarangi. 公有领域图片.
- 第 1 页：Pierre Louis Dumesnil 的原创画作，于 1884 年由 Nils Forsberg 临摹. 公有领域图片，于 2022 年 3 月 29 日下载自
  https://commons.wikimedia.org/wiki/File:René_Descartes_i_samtal_med_Sveriges_drottning,_Kristina.jpg.
- 第 22 页：公有领域图片，于 2022 年 2 月 4 日下载自
  https://commons.wikimedia.org/wiki/File:IAS_Princeton.jpg.
- 第 43 页：照片来自 Stefan Schäfer，采用知识共享署名-相同方式共享许可协议. 于 2022 年 3 月 28 日下载自
  https://commons.wikimedia.org/wiki/File:BurgDankwarderode2016.jpg.
- 第 100 页：照片来自 Alireza Javaheri，采用知识共享署名许可协议. 于 2023 年 3 月 12 日下载自
  https://commons.wikimedia.org/wiki/File:Hakim_Omar_Khayam_-_panoramio.jpg.
- 第 111 页：Alireza Javaheri 于 1863 年完成的雕像. 照片来自 Hans-Peter Postel，采用知识共享署名许可协议. 于 2022 年 3 月 14 日下载自
  https://commons.wikimedia.org/wiki/File:Leonardo_da_Pisa.jpg.
- 第 151 页：照片来自 Matthew Petroff，采用知识共享署名-相同方式共享许可协议. 于 2022 年 3 月 31 日下载自
  https://commons.wikimedia.org/wiki/File:George-peabody-library.jpg.
- 第 190 页：照片来自 Petar Milošević，采用知识共享署名-相同方式共享许可协议. 于 2022 年 3 月 30 日下载自
  https://en.wikipedia.org/wiki/Lviv.
- 第 249 页：照片来自 David Iliff，采用知识共享署名-相同方式共享许可协议. 于 2022 年 3 月 30 日下载自
  https://en.wikipedia.org/wiki/File:Long_Room_Interior,_Trinity_College_Dublin,_Ireland_-_Diliff.jpg.
- 第 279 页：照片来自 Daniel Schwen，采用知识共享署名-相同方式共享许可协议. 于 2019 年 7 月 9 日下载自
  https://commons.wikimedia.org/wiki/File:Mathematik_Göttingen.jpg.

# 符号索引

# 索引

各条目按照汉语拼音排序．部分条目下有若干子条目，子条目中用波浪线“~”指代其所属的条目．除部分不必或不便翻译的条目可能采用“原文【解释】”形式外，所有条目（含子条目）后附注英文原文．例如，“加法（addition）”条目下的“矩阵的 ~（of matrices）”，即表示“矩阵的加法”（addition of matrices）．

## Q



## R



## S



## T



## W



## X