普通高等教育"十一五"国家级规划教材

线性代数

Linear Algebra

李尚志

高等教育出版社
HIGHER EDUCATION PRESS

普通高等教育“十一五”国家级规划教材

线性代数

Xianxing Daishu

李尚志

高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书是在作者主持的国家精品课程“线性代数(非数学专业)”的建设过程中形成的教材，是作者主持的国家级教学成果奖二等奖项目“数学建模思想融入基础课教学”的重要成果之一。

本书不是“奉天承运皇帝诏曰”从天而降的抽象定义和推理，而是一部由创造发明的系列故事组成的连续剧。每个故事从颇具悬念的问题开始，在解决问题的过程中将所要学习的知识一步一步“发明”出来。随着剧情的发展，知识的引入如“随风潜入夜”，知识的应用如“润物细无声”，都成为自然而然的了。

本书适合作为大学本科非数学类专业线性代数、工科高等代数课程的教材，也可作为需要或关心线性代数和矩阵论知识的科技工作者、工程技术人员、大专院校师生及其他读者的参考书。

图书在版编目(CIP)数据

线性代数/李尚志编著. —北京：高等教育出版社，2011.6

ISBN 978-7-04-031795-4

Ⅰ.①线… Ⅱ.①李… Ⅲ.①线性代数-高等学校-教材

Ⅳ.①O151.23

中国版本图书馆CIP数据核字(2011)第056928号

策划编辑 兰莹莹　责任编辑 张彦云　封面设计 张 楠　版式设计 王 莹
责任绘图 尹文军　责任校对 殷 然　责任印制 朱学忠

出版发行 高等教育出版社
社　　址 北京市西城区德外大街4号
邮政编码 100120
印　　刷 山东省高唐印刷有限责任公司
开　　本 787×960 1/16
印　　张 19
字　　数 350 000
购书热线 010—58581118

咨询电话 400—810—0598
网　　址 http://www.hep.edu.cn
　　　　 http://www.hep.com.cn
网上订购 http://www.landraco.com
　　　　 http://www.landraco.com.cn
版　　次 2011年6月第1版
印　　次 2011年6月第1次印刷
定　　价 26.00元

物 料 号 31795—00

前 言

有人说，文学的永恒主题是爱与死。

数学的永恒主题是什么？问题的回答很可能众说纷纭，并且随着科学的发展而不断变化。但无论如何，可以说函数与方程是数学的重要主题，至少是中学数学与大学数学的重要主题。

最简单的函数是一次函数，最简单的方程是一次方程。

中学数学中已经学习过一元的一次函数与一次方程。

大学数学最重要的两门课程是微积分和线性代数。

微积分是把非一次的函数与方程化成一次函数与一次方程来研究。

线性代数干什么？“线性”就是“一次”，线性代数的主要内容就是研究多元的一次方程组与一次函数组。一次方程组也称为线性方程组。常数项为 0 的多元一次函数 $y = a_1x_1 + \cdots + a_nx_n$ 称为线性函数，n 个 n 元线性函数组成的函数组 $y_i = a_{i1}x_1 + \cdots + a_{in}x_n$ $(1 \leqslant i \leqslant n)$ 称为线性变换。大学非数学类专业线性代数课程的两大主题就是多元线性方程组与线性变换。线性方程组与线性变换可以由方程组或函数组中的系数排成矩阵来表示，通过矩阵运算来求解方程及解决相关的应用问题，其中最重要最基本的矩阵运算是初等变换和乘法这两个算法。用这两个算法为两大主题服务，是线性代数课程教学的基本任务。算法只有两个，要解决的问题却是千千万万，这就需要练就一双慧眼，从不同问题中看出共同点，将众多的问题转化为两个算法可以解决的形式。算法有一定之规，不难学会，难的是练就这双慧眼，提高实现这种转化的能力，这需要在长期实践中努力，不是仅仅通过几十学时的课堂教学所能做到的。但是，课堂教学至少应当让学生有一个良好的开端，朝正确的方向前进，而不要背道而驰，南辕北辙。

本书是为大学本科非数学类专业线性代数、工科高等代数课程编写的教材，基本目标不是将学生培养成数学理论的研究人员，而是让他们熟练掌握矩阵的初等变换与乘法这两种算法，逐步学会用这两种算法来解决学习和工作中与线性方程组和线性函数组有关的问题。我们不是以“奉天承运皇帝诏曰”的方式从天而降概念、算法和定理，也不是在学生不知有何用处的情况下先学好算法再拿去应用，而是从问题出发，在尝试解决问题的过程中将所需的算法“发明”出来。书中所有推理和证明的目的都是为了训练学生应用算法解决问

题的能力，使学生练就从纷纭复杂的问题中看出通向已有算法的出路的一双慧眼，而不是为了“数学的严密性”。数学的严密性，本身就是为了搞清楚相关结论和算法成立的理由和适用范围，保证算法在一定范围内的正确性和可靠性，既要防止超出适用范围发生错误，又要在适用范围内让结论和算法大显身手。考虑到很多非数学类专业的线性代数课程课时比较少，我们将推理和证明写得比较简略，尽量通过具体例子来体现普遍规律。有些结论和算法的证明和推理难度较大，我们就将它们写成附录，仅供教师或一部分感兴趣的学生参考，不作为课程学习内容，学生知道结论、会算会用就行了，暂时不必知其所以然。第 2 至 5 章有 6 节的标题用星号 * 标注出来，供工科高等代数课程选用，不作为线性代数课程的学习内容。线性代数课程只需在没有标 * 的章节中选择教学内容，并且还可以根据不同层次再略去一些内容。简而言之，本书为非数学类专业所有的教学和学生设置了共同的起点，共同的前进方向，但有几个不同的下车站。不同院校不同专业的教学可根据学生的不同层次选择不同的下车站，有的可以一直到终点站才下车，有的则可以选择适当的中途站，提前下车。

既然线性代数研究的是最简单的方程和函数，算法又比微积分少得多，按道理应当容易学。但实际情况是：很多学生学起来并不轻松。主要的困难是太抽象。比如，微积分中的导数可以理解为切线的斜率、运动的速度，定积分可以理解为求图形的面积、由速度求路程，这都比较具体。但线性代数从一开始就是一个接一个从天而降的抽象定义，使初学者难以理解。比如：行列式为什么要这样定义？矩阵为什么要这样相乘？向量到底是有方向和大小的量，还是数组，还是定义了加法和数乘的任意非空集合中的元素？线性相关、线性无关是什么意思，有什么用处？这些问题都让学生甚至很多讲授线性代数课程的教师迷惑不解。如果问学生为什么要学习行列式、矩阵和向量，为什么要学矩阵乘法和初等变换，很多学生的回答将是：教材上写了这些内容，考试要考这些内容，考试不及格就不能毕业，找不到工作，我敢不学吗？很多学生被线性代数折腾得很苦，死记硬背定义和算法，不知道这些定义和算法除了应付考试还有什么别的用处。即使是将定义和算法背熟了、考好了、毕业了的学生，在工程应用和理论研究中遇到问题需要解决的时候也不会应用线性代数的知识和算法。例如，需要解线性方程组时只会用中学的加减消去法而不会用矩阵的初等变换或求逆；需要根据顶点坐标计算平行四边形和三角形面积的时候不知道用二阶行列式；需要判断线性方程组是否有唯一解的时候不知道这就是判断系数矩阵各列是否线性无关；需要计算平面上和空间中的旋转时不知道用矩阵作乘法来实现；需要计算空间中的旋转轴时不知道这就是求特征值为 1 的特征向量，等等。

针对这种情况，本教材不从定义出发而从问题出发引入概念，引导学生在

尝试解决这些问题的过程中将所要讲授的知识重新“发明”出来。选择的问题，尽量是在现实生活和学生今后的工作中有用的问题，并且希望是学生感兴趣、容易懂的，因此也不选择需要较多专业知识或者综合性太强的实际问题。本书第一个例子是求前 n 个正整数的 k 次方和 $S_n = 1^k + 2^k + \cdots + n^k$ 的公式。这个问题看起来很困难，通过将 S_n 看成满足条件 $S_n - S_{n-1} = n^k$ 的多项式函数 $f(n) = a_1 n + a_2 n^2 + \cdots + a_{k+1} n^{k+1}$，就归结为求解待定系数 $a_1, a_2, \cdots, a_{k+1}$ 满足的三角形线性方程组，变得容易了。很自然提出问题：如果线性方程组不是三角形方程组，怎么求解？解决问题的思路也很自然：变成三角形方程组来求解。本书第 1 章第二个例子是求二次函数满足三组对应值，归结为解三元一次方程组。用中学的加减消去法求解这个三元一次方程组，引入了线性组合、方程组的同解变形、初等变换等概念和方法，再进一步将字母省去，将方程组用矩阵表示，方程组的初等变换用矩阵的初等行变换表示，得到了矩阵消去法。我们不是先讲矩阵的初等变换再用其来解方程组，而是为了解方程组“发明”出矩阵的初等行变换，这样引入初等变换就比较自然，学生容易理解和接受。但这还不够，矩阵的初等变换一旦发明出来，其用途就不限于解线性方程组，以后还要用来解决许许多多其他的问题，例如：判断线性相关和无关、求极大线性无关组和秩、计算行列式、求二次型的标准型、…… 人们买东西的时候都希望花一笔钱买来的一样东西有多样的用途，在科学研究中，用从许多事物中总结出来的结论和方法来解决千千万万个不同的问题，这就是抽象的威力，这样的抽象是好事。可是，如果在教学中不举任何一个实例，不讲任何一个应用，只让学生像念咒语一样死记硬背来应付考试，当然就会导致学生害怕抽象，这不是真正的抽象而是冒牌的抽象。因而不是抽象的错，也不是学生的错，而是教学方式的错。

本书第 2 章仍然围绕线性方程组这个主题，开始第一个例子就是讨论多项式函数 $f(x)$ 各项系数 $a_0, a_1, \cdots, a_m$ 满足的线性方程组什么时候有唯一解，在第 2.4 节中得到了求唯一解的公式，在第 3 章中又通过计算这个方程的系数行列式 (也就是范德蒙德 (Vandermonde) 行列式) 再次讨论了唯一解条件。为了讨论线性方程组的唯一解条件，我们先将二元和三元线性方程组写成向量形式，转化为平面和空间向量共线、共面的几何问题来讨论，再将共线、共面条件用向量的代数运算来描述，推广到 n 维空间，引出了线性相关、线性无关、维数、基、坐标等重要概念，并且在讨论方程组的解不唯一的情况时引出了秩、子空间等重要概念。很多学校因为课时不够而不讲线性空间。我们只花了课时讨论线性方程组的解的唯一性条件、解集合的大小和构造，与此同时也就讲完了线性空间的主要内容，不需要另外花课时，岂不两全其美？

第 2 章通过对平面向量共线与共面的几何问题的代数描述引出了线性相

关、线性无关等代数概念，解决了线性方程组解的唯一性及解集合的结构问题。第 3 章 (行列式) 仍然围绕线性方程组这个主题，仍然在几何和代数之间左右逢源。对 2 维和 3 维的情形，向量是否共线或共面可以通过平行四边形面积或平行六面体体积是否为 0 来描述。我们先由平行四边形面积和平行六面体体积引入二阶和三阶行列式，由它们的几何定义得到代数性质，由代数性质得出代数算法。将代数性质和代数算法推广，就得到 n 阶行列式。我们所用的三条代数性质是：第一，既然矩形的面积是长宽相乘，长方体或正方体的体积是长宽高相乘，将平行四边形面积和平行六面体体积看成两边或三条棱所代表的向量的某种乘积，按分配律和对于数乘的结合律展开。第二，更加理所当然的是，如果平行四边形或平行六面体有两条边或两条棱重合，面积或体积就等于 0。第三，单位正方形的面积和单位正方体的体积等于 1。我们正是从这三条性质推出了二阶和三阶行列式的计算公式，并且将算法和计算公式推广到了 n 阶行列式。不过，在这一章中我们并没有强调行列式的定义，行列式性质的证明也尽量淡化，将其中比较困难的证明放到附录中仅供参考，我们只突出了行列式性质的应用，尤其是初等变换对行列式的影响，特别是第三类初等变换不改变行列式的性质，通过初等变换计算行列式以及证明相关的定理 (主要是用行列式判断线性相关和无关的定理，及关于线性方程组唯一解的判定和唯一解公式的克拉默 (Cramer) 法则)。

矩阵的分块运算，是以华罗庚为代表的中国代数学家们进行科学研究的一个重要武器。然而，在很多线性代数教材和课堂中，分块运算却被认为没有什么用处，这是因为在这些教材和课堂中讲分块运算的时候没有给它用武之地，没有举出有用的例子。本教材中淡化分块运算的定义和证明，但将分块运算大用特用，让它大显神通。第 2 章中，在引入矩阵乘法之前，我们将向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合 $x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m$ 写成由向量组成的“行向量” $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 与系数组成的列向量 $\boldsymbol{X}$ 的乘积的形式 $\boldsymbol{AX}$，就是按分块运算的方式定义了矩阵 $\boldsymbol{A}$ 与列向量 $\boldsymbol{X}$ 的乘积。其后我们又将 S 的若干个线性组合组成的向量组 $(\boldsymbol{AX}_1,\cdots,\boldsymbol{AX}_p)$ 写成 $\boldsymbol{A}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_p)$，这就是用分块运算一般地定义了两个矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{X}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_p)$ 的乘积。第 4 章正式定义矩阵乘法之后，又用分块运算讨论矩阵乘法的各种重要性质，包括对角矩阵、纯量矩阵、单位矩阵的乘法性质，矩阵乘法的分配律和结合律。矩阵乘法的结合律，本来是被认为最繁琐最无法直观理解的，通过分块运算得到了自然的解释。在第 4 章的以后各节中，矩阵的分块运算还被用来计算逆矩阵、将初等变换转化为矩阵乘法、证明关于矩阵秩的不等式、计算行列式。

当我们通过字母运算得到 $(a+b)^2=a^2+2ab+b^2$ 时，不必考虑当 a,b 为分数时是否需要通分、当 a,b 符号相反时是否需要将绝对值相减、当 a,b 是无理

数时是否需要极限运算, 只要利用所有的复数共同满足的运算律: 乘法对于加法的分配律、加法结合律、乘法交换律, 就得到了所需的结论。这体现了运算律的威力, 也体现了抽象的威力。如果将 a, b 换成同阶方阵 $\boldsymbol{A}, \boldsymbol{B}$, 也不必考虑矩阵乘法与数的乘法有多大的区别, 只需考虑所用的三个运算律是否仍成立。矩阵乘法仍满足对加法的分配律, 矩阵加法仍满足结合律, 但矩阵乘法不满足交换律。不过, 某些特殊的方阵 $\boldsymbol{A}, \boldsymbol{B}$ 作乘法时可以交换: $\boldsymbol{AB} = \boldsymbol{BA}$, 对这样的 $\boldsymbol{A}, \boldsymbol{B}$, 不但完全平方公式 $(\boldsymbol{A}+\boldsymbol{B})^2 = \boldsymbol{A}^2 + 2\boldsymbol{AB} + \boldsymbol{B}^2$ 成立, 而且可以用牛顿二项式定理展开 $(\boldsymbol{A}+\boldsymbol{B})^n$。在第 4 章中, 以若尔当 (Jordan) 形矩阵

$$\boldsymbol{J} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \lambda \boldsymbol{I} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

为例, 用牛顿二项式定理计算 $\boldsymbol{J}^{10}$; 当 $\lambda = 1$ 时用泰勒 (Taylor) 展开式 $(1+x)^{-1} = 1 - x + x^2 + \cdots$ 求逆 $\boldsymbol{J}^{-1}$; 用 $(1+x)^{\frac{1}{n}} = 1 + \dfrac{1}{n}x + \dfrac{\frac{1}{n}\left(\frac{1}{n}-1\right)}{2}x^2 + \cdots$ 求 $\boldsymbol{J}$ 的一个“n 次方根” $\boldsymbol{X}$ 使其满足 $\boldsymbol{X}^n = \boldsymbol{J}$; 还可以用泰勒 (Taylor) 展开式 $\mathrm{e}^x = 1 + x + \dfrac{1}{2!}x^2 + \cdots$ 计算微分方程组 $\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X} = \boldsymbol{JX}$ 的解 $\boldsymbol{X} = \mathrm{e}^{\boldsymbol{J}t}C$。很自然引出一个问题: 怎样求任意方阵 $\boldsymbol{A}$ 的 n 次幂、方根和指数函数 $\mathrm{e}^{\boldsymbol{A}}$? 这就为将 $\boldsymbol{A}$ 相似于对角矩阵或若尔当 (Jordan) 标准形埋下了伏笔。

我参加面试 2010 年入学的研究生的时候, 问了一个问题: 求矩阵

$$\boldsymbol{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

的 2010 次幂。几乎所有的线性代数教材上都有对这个矩阵 $\boldsymbol{A}$ 求 $\boldsymbol{A}^n$ 的习题, 按道理学生应当会做。让我震惊的是, 将近十名在笔试中考高分的非数学类专业的学生没有一个能够给出答案。几乎每个考生都说, 先算 $\boldsymbol{A}$ 的平方, 再算 $\boldsymbol{A}$ 的 3 次方, 观察规律, 然后用数学归纳法证明。我问: 你们的教科书上都有这道题, 为什么以前没有算过 $\boldsymbol{A}$ 的平方与 3 次方, 没有观察过规律? 他们没有回答。我猜想, 大概他们的老师也认为这个题只能先算平方、3 次方观察规律, 再用数学归纳法证明, 认为这样的题目没有什么意思, 因此没有布置学生做, 学生也就没有做过。然而, 只要知道用矩阵 $\boldsymbol{A}$ 左乘列向量 $\boldsymbol{X}$ 的效果

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

是将点 $P(x,y)$ 旋转角 α 变到点 $P'(x',y')$, 就可立即知道: $\boldsymbol{A}^n$ 乘 $\boldsymbol{X}$ 的效果是将旋转 α 的动作重复进行 n 次, 总效果是旋转 $n\alpha$, 因而

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}^n = \begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}。$$

可是, 这些学生都不知道 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 是平面旋转, 因此都觉得要计算 $\boldsymbol{A}$ 的几千次幂是一件不可能完成的事情。为什么不知道, 是因为他们的老师没有告诉他们。为什么不告诉? 老师的解释可能是: 我们的课时不够, 不能讲线性变换。实际上, 不需要多少课时, 只要一句话就可以讲线性变换: 将列向量 $\boldsymbol{X}$ 乘方阵 $\boldsymbol{A}$ 得到列向量 $\boldsymbol{Y}$, 就是线性变换。将平面或空间每个点 P 通过旋转或轴对称变到点 P', 从 P 的坐标 $\boldsymbol{X}$ (写成列向量) 到 P' 的坐标 $\boldsymbol{Y}$ 的变换可以用某个方阵 $\boldsymbol{A}$ 左乘 $\boldsymbol{X}$ 来实现: $\boldsymbol{Y} = \boldsymbol{AX}$, 因此是线性变换。

本书没有花很多篇幅讨论线性变换, 只是在第 4.2 节中以平面上的旋转与轴对称两个例子来加以说明。线性变换 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 的各种性质 (保加法、保数乘、矩阵 $\boldsymbol{A}$ 的各列是自然基的像) 都由矩阵乘法的性质得出, 这一节对线性变换的讲解就变成了对矩阵乘法的训练, 训练学生理解矩阵乘法的代数性质的几何意义。第 4.3 节讲矩阵的逆, 第一个例子就是求平面旋转矩阵 $\boldsymbol{A}$ 的逆, 不是通过代数运算求逆, 而是通过几何观点求逆: 既然 $\boldsymbol{A}$ 表示旋转角 α, 旋转 $-\alpha$ 的矩阵就是 $\boldsymbol{A}^{-1}$。第 4.4 节的一个例子中, 将椭圆通过线性变换 $(x,y) \mapsto \left(x, \dfrac{a}{b}y\right)$ 拉伸成圆, 由圆的面积及内接四边形的最大面积来求椭圆面积及内接四边形的最大面积。通过这个简单易懂的例子解释了变换矩阵的行列式的几何意义, 引出了行列式的乘法性质。这个“拉伸”变换的矩阵是对角矩阵, 为第 5 章的特征向量和相似对角化埋下了伏笔。第 5 章第 5.5 节前两个例子与此一脉相承, 用非对角矩阵将斜着摆放的椭圆拉伸成圆。通过计算被拉伸的向量引出了特征向量, 并让学生一开始就知道了特征向量的几何意义。本书特征向量的例子, 一类是通过计算方阵 $\boldsymbol{A}$ 的特征向量研究对应的线性变换 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 的几何性质, 既有将向量拉长或缩短的变换的例子, 也有空间中旋转和对称变换的例子; 另一类是以计算 $\boldsymbol{A}$ 的函数如 $\boldsymbol{A}^n, \mathrm{e}^{\boldsymbol{A}t}$ 为目标的代数例子, 解释了将方阵 $\boldsymbol{A}$ 相似于对角矩阵或若尔当形矩阵的必要性和算法。

线性代数的算法只对行数和列数很少的矩阵才能用手算实现。而在实际工作中, 经常需要处理几十、几百甚至更多行和列的矩阵, 难以用手工实现算法, 必须求助于计算机及其软件。已经有很多软件能够满足这样的要求, 例如 MATLAB 与 Mathematica。西安电子科技大学陈怀琛教授主持的教育部教

改项目“用 MATLAB 和建模实践改造工科线性代数课程”在利用信息技术工具改革线性代数课程方面进行了有益的探索和实践。本书作者也组织了一些教师参加这个项目。本书以数学实验的方式附上了线性代数一些主要算法的 MATLAB 命令和 Mathematica 命令。其中 MATLAB 的例子是青年教师李娅为本书选择和整理的，谨在此感谢。读者可以“依样画葫芦”运行这些命令，再“照葫芦画瓢”替换其中的数据计算别的习题。关于 MATLAB 命令和 Mathematica 命令更详细的知识，请参考相关的书籍和资料，例如本书参考文献 [2],[3]。

李尚志

2010 年 12 月 31 日

目　　录

第 1 章　线性方程组的解法

§1.1　线性方程组的初等变换

1.1.1　多元线性方程组

中学数学学习了二元一次方程组的解法, 但在科学研究和实际应用中经常需要解更多未知数的一次方程组.

例 1　已知正整数 n, 求

(1) $1^2+2^2+\cdots+n^2$;

(2) $1^4+2^4+\cdots+n^4$.

分析　中学数学给出了前 n 个正整数的平方和 $1^2+2^2+\cdots+n^2$ 的公式, 并且用数学归纳法证明了公式的正确性, 但没有教学生怎么"发明"出这个公式, 也不能类似地得到更高次数的幂的和 $1^k+2^k+\cdots+n^k$ 的公式.

本例是已知数列通项公式 $u_n=n^k$ 求前 n 项和 S_n, 这很困难. 反过来, 已知前 n 项和 S_n 求 u_n 却很容易: $u_n=S_n-S_{n-1}$ (当 $n\geqslant 2$), $u_1=S_1=1$.

如果将 S_n 看成 n 的函数 $S_n=f(n)$, 则上述等式就是函数 $f(n)$ 应当满足的必要条件:

$$f(n)-f(n-1)=u_n=n^k\ (\forall\ n\geqslant 2),\qquad f(1)=u_1=1.$$

不难想到, 如果 $f(n)$ 是 n 的多项式函数, 则 $f(n)-f(n-1)$ 也是多项式, 并且次数比 $f(n)$ 低一次. 如果能用待定系数法求 $k+1$ 次多项式 $f(n)=a_0+a_1n+\cdots+a_{k+1}n^{k+1}$ 的系数, 使 $f(n)$ 满足恒等式

$$f(n)-f(n-1)=a_1[n-(n-1)]+\cdots+a_{k+1}[n^{k+1}-(n-1)^{k+1}]=n^k,$$

则

$$\begin{aligned}S_n&=1^k+2^k+\cdots+n^k\\&=(f(1)-f(0))+(f(2)-f(1))+\cdots+(f(n)-f(n-1))\\&=f(n)-f(0).\end{aligned}$$

只要取 $f(0)=a_0=0$, 即可使 $S_n=f(n)$ 对所有的正整数 n 成立.

解　(1) 设 $f(n)=a_1n+a_2n^2+a_3n^3$, 满足条件

$$f(n)-f(n-1)=a_1[n-(n-1)]+a_2[n^2-(n-1)^2]+a_3[n^3-(n-1)^3]$$

$$= a_1 + a_2(2n-1) + a_3(3n^2-3n+1)$$
$$= (a_1 - a_2 + a_3) + (2a_2 - 3a_3)n + 3a_3n^2 = n^2, \text{ 即}$$

$$\begin{cases} a_1 - a_2 + a_3 = 0, & ① \\ 2a_2 - 3a_3 = 0, & ② \\ 3a_3 = 1. & ③ \end{cases} \tag{1.1.1}$$

由方程 ③ 解出 $a_3 = \dfrac{1}{3}$, 代入 ② 解出 $a_2 = \dfrac{1}{2}$, 再代入 ① 解出 $a_1 = \dfrac{1}{6}$, 得到

$$S_n = \frac{1}{6}n + \frac{1}{2}n^2 + \frac{1}{3}n^3 = \frac{2n^3+3n^2+n}{6} = \frac{n(n+1)(2n+1)}{6}.$$

(2) 将 $S_n = f(n) = a_1n + a_2n^2 + a_3n^3 + a_4n^4 + a_5n^5$ 代入 $f(n) - f(n-1) = n^4$, 整理并比较对应项系数得

$$\begin{cases} a_1 - a_2 + a_3 - a_4 + a_5 = 0, & ① \\ 2a_2 - 3a_3 + 4a_4 - 5a_5 = 0, & ② \\ 3a_3 - 6a_4 + 10a_5 = 0, & ③ \\ 4a_4 - 10a_5 = 0, & ④ \\ 5a_5 = 1. & ⑤ \end{cases} \tag{1.1.2}$$

由下而上依次从各方程中解出 $a_5 = \dfrac{1}{5}, a_4 = \dfrac{1}{2}, a_3 = \dfrac{1}{3}, a_2 = 0, a_1 = -\dfrac{1}{30}$, 得到

$$S_n = \frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 - \frac{1}{30}n = \frac{n(6n^4+15n^3+10n^2-1)}{30}$$ □

例 1 中的 (1.1.1),(1.1.2) 分别是 3 元一次方程组与 5 元一次方程组.

一般地, n 个未知数 $x_1, \cdots, x_n$ 的如下形式的方程

$$a_1x_1 + \cdots + a_nx_n = b$$

称为 n 元一次方程, 也称为 n 元线性方程, 其中 $a_1, \cdots, a_n, b$ 是已知的常数, $a_1, \cdots, a_n$ 是一次项系数, b 是常数项. 具有同样 n 个未知数 $x_1, \cdots, x_n$ 的若干个一次方程组成的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.1.3}$$

称为 n 元一次方程组, 也称 n **元线性方程组** (linear equations in n variables).

注意, 在线性方程组中, 并不要求方程的个数 m 等于未知数个数 n, $m<n, m=n, m>n$ 等三种情况都允许. $m=1$ 的情形也允许. 也就是说: 一个线性方程也可以组成线性方程组.

如果将方程组 (1.1.3) 中的 n 个字母 $x_1,\cdots,x_n$ 分别替换成某 n 个已知数 $c_1,\cdots,c_n$, 得到的 m 个等式

$$a_{i1}c_1+\cdots+a_{in}c_n=b_i \quad (1\leqslant i\leqslant m)$$

全部成立, 就称 $c_1,\cdots,c_n$ 组成的有序数组 $(c_1,\cdots,c_n)$ 是方程组 (1.1.3) 的一个**解** (solution). 方程组的全体解组成的集合称为方程组的解集.

1.1.2 线性方程组的同解变形

例 1 的线性方程组 (1.1.1),(1.1.2) 具有特殊的形状: 从上到下每个方程比上一个方程少含一个未知数, 最后一个方程只含一个未知数. 将各方程的等号上下对齐, 同一未知数的项也上下对齐, 则等号左边左下角是空白 (其中各项系数都是 0), 其余各项组成一个摆放在右上角的三角形. 我们称这样的线性方程组为"上三角形". 例 1 的解法可以推广到一般的上三角形方程组: 从最后一个方程解出所含的未知数的值, 代入上一个方程再解出一个未知数的值, 由下而上依次将已求出的未知数的值代入上一个方程, 可以依次求出所有的未知数的值, 从而求出方程组的解.

任给的线性方程组 U 不一定是三角形, 但如果能够变成例 1 那样的三角形方程组 T, 就可以求出方程组 T 的解. 但是, 必须首先保证 T 与 U 同解, 由 T 得到的才是 U 的全部解.

例 2 二次函数 $y=f(x)$ 的图像经过三个已知点 $(1,1),(2,2),(3,0)$. 求 $f(4)$.

分析 设所求二次函数为 $f(x)=ax^2+bx+c$, 其中 a,b,c 是待定常数. 则

$$\begin{cases} a+\ \ b+c=1, & ① \\ 4a+2b+c=2, & ② \\ 9a+3b+c=0. & ③ \end{cases} \tag{1.1.4}$$

尝试用中学的加减消去法解这个三元一次方程组.

方程② − 方程①, 得 $\quad 3a+b=1.$ ④

方程③ − 方程②, 得 $\quad 5a+b=-2.$ ⑤

方程⑤ − 方程④, 得 $\quad 2a=-3.$ ⑥

由 ⑥ 可解出 a, 代入 ④ 可解出 b. 将 a,b 代入 ① 可解出 c. 这其实是从方程组

$$\begin{cases} a+\ \ b+c=1, & ① \\ 3a+b\qquad\ =1, & ④ \\ 2a\qquad\qquad=-3 & ⑥ \end{cases} \tag{1.1.5}$$

解出 a,b,c. 方程组 (1.1.5) 等号左边的右下角是空白, 左上角组成三角形, 可以由下而上依次求出各未知数的值. 但有一个问题: 由方程组 (1.1.5) 求出的是否是原方程组 (1.1.4) 的解集?

新方程 ④,⑤ 由原方程 ①,②,③ 相减得到, 新方程 ⑥ 又由 ④,⑤ 相减得到. 原方程组 (1.1.4) 的解使原方程 ①,②,③ 成为等式, 这些等式相减得到的 ④,⑤ 以及 ⑥ 仍然是等式. 这说明了原方程组 (1.1.4) 的解一定是新方程组 (1.1.5) 的解. 反过来, 如果 (1.1.5) 中的方程 ①,④,⑥ 乘常数再相加能够反过来得到原方程 ②,③, 则 (1.1.5) 的解也是原方程组 (1.1.4) 的解, 就可以保证原方程组 (1.1.4) 与新方程组 (1.1.5) 同解.

由原方程组 (1.1.4) 直接变成 (1.1.5), 变化太大, 不容易看出 (1.1.5) 怎样变回 (1.1.4). 我们将这个大变化分成若干个小步骤来实现, 使每一小步得到的方程组容易变回前一个方程组.

任意方程组 U 中各方程分别乘常数再相加得到的新方程称为方程组 U 的**线性组合** (linear combination), 由 U 的若干个线性组合组成的方程组 W 也称为 U 的线性组合. 容易验证, 方程组 U 的解一定是它的线性组合 W 的解. 反过来, 如果 W 可以通过线性组合变回 U, 则方程组 W 的解也是 U 的解, $U\to W$ 是同解变形.

定理 1.1.1 对线性方程组 U 进行以下三类变形, 得到的新方程组 W 与原方程组 U 互为线性组合, $U\longrightarrow W$ 是同解变形.

(1) 将第 i 个方程与第 j 个方程互相交换位置: $U \xrightarrow{(i,j)} W$.

(2) 将第 i 个方程两边同乘非零常数 λ: $U \xrightarrow{\lambda(i)} W$.

(3) 将第 i 个方程的 λ 倍加到第 j 个方程: $U \xrightarrow{\lambda(i)+(j)} W$.

证明 易见 W 的每个方程都是 U 的线性组合.

反过来, W 可以通过同样类型的变形变回 U :

(1) $U \xrightarrow{(i,j)} W \xrightarrow{(i,j)} U$.

(2) $U \xrightarrow{\lambda(i)} W \xrightarrow{\lambda^{-1}(i)} U$.

(3) $U \xrightarrow{\lambda(i)+(j)} W \xrightarrow{-\lambda(i)+(j)} U$.

可见 U 也是 W 的线性组合.

这就证明了方程组 U 与 W 互为线性组合. 由 U 到 W 的变形是同解变形. □

定理 1.1.1 中所说的三类变形称为方程组的**初等变换** (elementary transformations).

例 2 解法 1　所求二次函数 $f(x)=ax^2+bx+c$ 的系数满足方程组

$$\begin{cases} a+\ \ b+c=1, \\ 4a+2b+c=2, \\ 9a+3b+c=0. \end{cases} \qquad (U)$$

利用初等变换将原方程组 (U) 变成三角形方程组.

$$(U)\xrightarrow{-(2)+(3)}\begin{cases} a+\ \ b+c\ =1, \\ 4a+2b+c=2, \\ 5a+\ \ b\ \ \ \ \ \ =-2 \end{cases}\xrightarrow{-(1)+(2)}\begin{cases} a+b+c=1, \\ 3a+b\ \ \ \ \ =1, \\ 5a+b\ \ \ \ \ =-2 \end{cases}$$

$$\xrightarrow{-(2)+(3)}\begin{cases} a+b+c=1, \\ 3a+b\ \ \ \ \ =1, \\ 2a\ \ \ \ \ \ \ \ \ \ =-3. \end{cases} \qquad (T)$$

从方程组 (T) 由下而上解出 $a=-\dfrac{3}{2},b=\dfrac{11}{2},c=-3.$

所求函数 $f(x)=-\dfrac{3}{2}x^2+\dfrac{11}{2}x-3,\ f(4)=-5.$ □

例 2 的方程组化成三角形方程组 (T) 之后, 还可以用初等变换将 (T) 进一步化简, 使每个方程都只包含一个未知数, 这些未知数系数都是 1:

$$(T)\xrightarrow{-\frac{1}{2}(3)+(1),-\frac{3}{2}(3)+(2)}\begin{cases} \qquad b+c\ =\dfrac{5}{2}, \\ \qquad b\ \ \ \ \ \ =\dfrac{11}{2}, \\ 2a\ \ \ \ \ \ \ \ \ \ \ =-3 \end{cases}$$

$$\xrightarrow{-(2)+(1)}\begin{cases} \qquad\qquad c\ =-3, \\ \qquad b\ \ \ \ \ \ =\dfrac{11}{2}, \\ 2a\ \ \ \ \ \ \ \ \ \ \ =-3 \end{cases}\xrightarrow{\frac{1}{2}(3)}\begin{cases} \qquad\qquad c\ =-3, \\ \qquad b\ \ \ \ \ \ =\dfrac{11}{2}, \\ a\ \ \ \ \ \ \ \ \ \ \ \ =-\dfrac{3}{2}, \end{cases} \qquad (\Lambda)$$