线性代数

李尚志　编著

内 容 简 介

本书是国家十五规划教材，是作者主持的国家级精品课程《线性代数》的教材．适合作为大学本科数学专业《线性代数》（或称《高等代数》）课程的教材，也可作为各类大专院校师生的参考书，以及关心线性代数和矩阵论的知识的科技工作者或其他读者的自学读物或参考书．

本书具有如下特点：

1．不是从定义出发，而是从问题出发来展开课程内容，引导学生在分析和解决这些问题的过程中将线性代数的知识重新“发明”一遍，貌似抽象难懂的概念和定理也就成为显而易见．

2．“空间为体，矩阵为用”，自始至终强调几何与代数的相互渗透．

3．不板着面孔讲数学，努力采用生动活泼、学生喜闻乐见的语言．

前　言

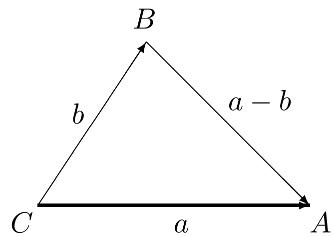
《线性代数》是大学最重要的数学基础课之一。其基本内容是线性空间和矩阵的理论。线性代数的知识，是学习数学和其它学科的重要基础，并且在科学研究各个领域和各行各业中有广泛的应用。该课程对于培养学生的逻辑推理和抽象思维能力、空间直观和想象能力具有重要的作用。

本书是国家十五规划教材，是作者主持的国家级精品课程《线性代数》的教材，适用于大学本科数学类专业的《线性代数》课 (或称为《高等代数》课) 作为教材.

《线性代数》的内容比《数学分析》或《微积分》少，但不少学生感到学起来并不更容易。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，这都比较自然，容易理解。而线性代数从一开始就是一个接一个从天而降的抽象定义，使初学者感到不好理解。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？让许多初学者迷惑不解。

抽象确实是学习线性代数的一个拦路虎。一提起抽象，给人的印象就是莫名其妙、晦涩难懂、脱离实际、没有用处，总之是一个令人害怕的贬义词。然而，抽象并不是线性代数特有的，也不是从大学开始的。比如，幼儿园的小孩就要学 $3+2=5$, 这是抽象还是具体？怎样教小孩 $3+2=5$？是先教加法的定义，然后再按照定义来做 $3+2=5$ 吗？加法的定义，幼儿园没教过，小学和中学也没教过. 然而小孩们却学会了加法，不是靠定义学会加法，而是通过例子学会了加法。比如，可以教小孩数自己的手指来学 $3+2=5$, 3 只手指加 2 只手指就是 5 只手指。也可以数铅笔， 3 支铅笔加 2 支铅笔就是 5 支铅笔。假如已经数过了手指，又数过了铅笔，一个细心而胆大的小孩发现手指是肉做的，铅笔是木头做的，举手问老师：“ 5 是肉做的还是木头做的？”老师怎样回答？假如又数了 5 个乒乓球，发现手指和铅笔是长的，乒乓球是圆的，再问：“ 5 是长的还是圆的？”老师又怎样回答？也许老师会斥责这个不听话的调皮小孩：“ 好好听课，不要胡说八道！”然而，这样的小孩才是聪明的小孩，会思考的小孩. 他注意到了 5 只手指、 5 只铅笔、 5 个乒乓球的差别，这确实是聪明的表现。但只是注意到差别还不够，还要让他学会忽略这种差别，将肉做的 5 只手指、木头做的 5 只铅笔“混为一谈”, 将 5 个长的物体 (手指和铅笔) 和 5 个圆的物体 (乒乓球)“混为一谈”, 忽略它们的差别而只关心它们的共同点：数量的多少，这才学会了 $3+2=5$, 这才能够将 $3+2=5$ 用到千千万万的其它例子，如 3 本书加 2 本书， 3 张桌子加 2 张桌子等。要让小孩学会忽略这些差别，不是一件容易的事情. 这正如郑板桥说的：“聪明难，糊涂亦难，由聪明而糊涂尤难。”这种忽略差别的过程，就是“由聪明而糊涂”的过程，也就是数学的抽象的过程。抽象不是从天而降，而是来自于实际，来自于具体的例子。然而，抽象又没有停留于实际，而是“脱离”了实际：它脱离了具体的例子，舍弃了不同例子的不同点而提取了它们的共同点，这样才能应用到更多更广泛的实际例子中。有一个电视节目的时事评论员常说：“许多看似不相干的事情，其实都是相互关联的。”我们可以说：“许多看似不相同的事情，其实都有共同点。”从不同的事情中发现共同点，研究共同点，得到放之四海而皆准的真理，用到更多的不

同事物中去，这就是抽象。这样的抽象不是没有用处，反而是神通更广大。数学由低级到高级的过程，就是抽象的程度由低到高的过程，也是应用的范围由狭窄到广泛的过程。幼儿园的 $3+2=5$ 忽略掉了大小、长短、原料的差别，只关心数量的多少。初中的 $(a-b)^2=a^2-2ab+b^2$ 将字母 a,b 所代表的数的多少也忽略掉了，只关心它们的共同的运算规律。更进一步的“糊涂”是：公式 $(a-b)^2=a^2-2ab+b^2$ 中的字母 a,b 可以不代表数而代表几何向量，将其中的乘法理解为向量的内积，公式照样成立．画出有向线段来表示公式中向量，如下图： $\overrightarrow{CA}=a,\overrightarrow{CB}=b$，则 $\overrightarrow{BA}=a-b$.



公式 $(a-b)^2=a^2+b^2-2a\cdot b$ 的几何意义就是

$$|BA|^2=|CA|^2+|CB|^2-2|CA||CB|\cos C.$$

这就是余弦定理！ 当 C 是直角时就是勾股定理！ 只不过一念之差，在乘法公式 $(a-b)^2=a^2+b^2-2ab$ 中“难得糊涂”，将数与向量“混为一谈”, 就立即得到了余弦定理和勾股定理，数学的抽象的威力由此可见一斑！

我在念研究生的时候，导师曾肯成教授经常给我们指定一些经典著作让我们读，并且轮流到讲台上去讲，他在下面听，向讲的人提出各种问题。通常他并不问某某定义怎样叙述、某某定理怎样证明，而是问：“书上为什么要写这个内容？不写可不可以？”这样的问题很难回答，但我们不得不努力去思考怎样回答。假如我们在讲线性代数时问：“书上为什么要写矩阵和向量空间？为什么要写线性相关和线性无关？可不可以删去？”应当怎样回答？前人在发明这些内容的时候都是为了解决一定的问题．写进教科书中的内容，更是经过历史的检验被证明是最重要最有用的东西．课本上的定义和定理的叙述，每一句话甚至每个字都是经过多少年多少人的反复推敲得到的，字字值千金．我们今天没有必要完全重复当年发明这些知识的过程，更没有必要去重复他们走过的弯路．但是，如果我们的学生只是去死记硬背这些定义或定理的条文，而对这些条文的来龙和去脉毫不了解，不了解它们产生的背景，不知道它们有什么用处，不了解这些抽象的概念和结论所能代表的一些具体例子，就不能体会这些定义和定理的深刻含义和强大威力，反而觉得它们莫名其妙，枯燥乏味，学起来困难，学了也不知道有什么用处.

有鉴于此，我们形成了如下的教学模式，也就是编写本教材的指导思想：不是从定义出发，而是从问题出发展开课程内容．我们围绕线性代数的主要内容，精选了一些有重大意义而又浅显易懂的问题作为组织课程内容的主要线索．引导学生一起来分析这些问题，尝试建立一定的数学工具来解决这些问题．这实际上也就是数学建模的过程 —— 将所要解决的问题 (实际问题或理论

问题) 用适当的数学语言加以描述, 转换为数学问题 (即数学模型), 用一定的数学工具 (已有的工具或发明新的工具) 来加以解决, 再将所得到的数学解翻译成为原来问题的解. 在解决原有问题的过程中又产生新的问题，需要建立新的概念、方法和技巧。这个过程本来是这些知识当初建立的过程，也是学生今后搞科学研究和应用要经历的过程. 我们不是将前人得到的知识灌输给学生，而是引导学生重新经历一次发明这些知识来解决问题的过程. 这样做的好处是：不只是背颂叙述知识的条文，而是体会到这些条文的来龙去脉，体会到这些知识的原创性的想法和实质，提前接受了从事科研工作的训练，培养了创新意识和素质。

现在我们要建立创新型国家, 培养具有创新精神的人才是一项重要任务. 对于培养创新精神，有一种看法认为大学低年级学生基础知识不够，还谈不上创新，只有到研究生阶段才能培养创新精神和素质. 其实，即使到了硕士研究生阶段，也很少有人能做出真正具有创新性的研究成果，主要还是在为以后做出创新性成果打好基础. 而培养创新精神和素质，也完全可以从本科生低年级就开始. 低年级大学生学的基础课内容，对人类来说是已有的知识，可能还是几百年前发明的知识，但对这些学生来说却是新的知识. 让他们将这些知识重新发明一遍，虽然发明出来的东西对人类不是新的成果，但对他们自己却是发明创造的一种模拟和演习，是一次创新的实践，对于培养创新精神和素质很有好处.

我们希望通过从问题出发的教学模式引起学生探索问题的兴趣，培养创新精神. 但是我们也明白，采用这种模式也可能有它的缺点和风险. 如果把握不好，有可能导致叙述冗长，主线不突出，干扰学生对主干内容的理解. 为了避免这一缺点，我们一方面努力做到选择问题恰当，由问题引入概念时适度，强调思路而不在细节上纠缠. 另一方面，为了减少风险，我们在教材编写体例上采取了一个特殊的处理方式：在每一章的第 $1, 2, \ldots$ 节前面设置第 0 节，其内容是提出问题，对问题做一定分析，引出本章的主要概念. 而从第 1 节仍是按传统的方式从定义开始叙述. 使用本教材的教师可以对第 0 节灵活处理. 比如可以直接从第 1 节开始讲授, 将第 0 节留给感兴趣的学生自己去阅读. 也可以将第 0 节的内容作提纲挈领的简单介绍，主要介绍解决问题的思路. 除了每一章的第 0 节以解决问题为线索引入教材内容之外，我们还从第二章开始将每章的最后一节设为“更多的例子”, 其中一部分是一些综合性、技巧性较高的问题的解答，而另一部分就是基本知识的一些有启发性的应用例以及扩展性的知识，希望通过这样的例子展现抽象的基础知识在解决各种问题中的强大威力，并且为如何利用这些知识解决各种问题提供范例. 教师在使用本教材时可以根据各自的具体情况灵活选用这些例子，不一定将它们全部讲完.

具体地说，在第一章我们首先选取了怎样解多元线性方程组来作为要解决的第一个问题，这是线性代数的重大问题之一，而且与中学数学中解二元一次方程组的加减消元法自然衔接，学生容易接受. 我们将中学数学对方程的变形提高到方程组的同解变形，将加减消去法提高到方程的线性组合、线性方程组的初等变换. 由于解线性方程组只用到加减乘除四则运算，在讨论方程组的系数与解之间的关系时很自然引入了数域的概念. 在将字母略去不写之后很自然地引入了数组向量来表示线性方程，用矩阵来表示线性方程组，用矩阵的初等行变换来解线性方程组. 矩阵的初等变换无疑是整个线性代数最重要的计算手段，借助于解线性方程组这一重大问题让学生得到

了充分的训练.

第二章引入了线性空间，这可能是本教材与以往教材差别最大的一部分. 在这一章中，围绕线性方程组中方程的个数与解集的大小之间的关系的讨论，以自然的方式引出了线性相关、线性无关、秩、基、维数等一批最重要而又最抽象难懂的概念，以方程作为向量的重要例子展开了对线性空间的主要内容的全面讨论. 我们不是先研究了抽象的线性空间的性质再应用到各种具体空间中去，而是反过来，先对数组空间得出了这些性质，然后指出对这些性质的证明其实不依赖于数组空间的特殊性质，而是只依赖于加法和数乘的运算律 (8 条公理), 因此可以适用于定义了加法和数乘的任何其它对象 — 抽象的线性空间. 这不但使抽象的线性空间的定义的引入比较自然，而且对于什么是数学的抽象、怎样进行数学的抽象、怎样由直观而不严格的想法 (方程的个数、解集的大小) 建立严格的数学概念 (线性相关、线性无关、秩、维数) 提供了一个重要的范例，让学生在以后的学习和研究中可以模仿. 在这一章的最后一节"更多的例子"中，利用子空间的思想求斐波那契数列的通项公式、设计幻方，利用同构的思想得出拉格朗日插值公式、中国剩余定理，都是应用抽象的代数概念来解决问题的很好的例子. 这些例子都有另外的专门的方法和技巧来解决，而在我们这里却不需要学习任何专门的方法和技巧，只需要将线性代数中的最简单的基本思想适当地应用，问题就迎刃而解.

由于第二章已经建立了线性空间的概念，第三章就可以从几何的背景引入行列式. 以往的行列式教学中有一件怪事：学生从空间解析几何中知道了三阶行列式可以代表平行六面体的体积，却不知道二阶行列式可以代表平行四边形的面积. 我们将 n 元线性方程组解释为一个几何问题：由 n 已知向量线性组合出另一个已知向量: $x_1\alpha_1+\cdots+x_n\alpha_n=\beta$, 求组合系数 $x_1,\dots,x_n$. 通过对 2 元一次线性方程组所对应的几何问题的分析和解答得到了 2 阶行列式以及二元一次方程组的 Crammer 法则，再将它推广到了 n 阶行列式及 n 元线性方程组的 Crammer 法则，将 n 阶行列式看成 n 维"体积", 将行列式中某一列的代数余子式组成的向量看成其余各列的"外积". 从几何的观点看来，行列式的定义和各种性质都显得理所当然，容易接受.

代数和几何相互渗透不可分割，这是线性代数的一个基本特点，也是本课程和教材的一个重要特点. 线性代数名曰代数，其实也是几何. 在某种意义上可以说：空间解析几何是 3 维空间的线性代数，而线性代数是 n 维空间的解析几何. 线性代数的主要内容，可以用"空间为体，矩阵为用"来概括. 它研究的对象是由向量组成的线性空间，这是几何对象. 研究的工具则是矩阵，这是代数工具. 学生如果能够将关于向量空间的几何问题转化为矩阵的问题，用矩阵运算加以解决，再"翻译"回几何的语言得到答案，就算是对线性代数的基本理论和方法有了一个较好的掌握. 几何与代数紧密不可分割，是同一个事物的两个方面. 然而这两个方面也各有所长，各有所短. 几何的优点是形象直观便于理解，缺点是不便于计算；矩阵的优点是便于计算，缺点是不便于理解. 因此，在处理问题时要发挥它们各自的长处，适当回避它们的缺点：几何观点主要用来建模，将几何语言转换为矩阵语言之后再用矩阵计算来解决问题，然后再用几何观点加以理解和解释. 当然，也有很多问题可以不用矩阵计算来解决，而用几何推理来解决. 对这方面的例子，我们同时给出了矩阵计算和几何推理两种解决方法. 在多数情况下，用矩阵来计算有比较"死板"的现成方法，容

易掌握，类似于解析几何方法；用几何推理比较灵活多变，比较有趣，然而掌握起来更困难一些，类似于综合几何的证明方法．对初学者，我们还是提倡他们先掌握将几何问题转化成矩阵、通过矩阵运算计算来处理的方法，在此基础上再去自由发挥，追求更为灵活多变的几何方法．我们这一指导思想贯穿于全书的所有各章．第一章解方程组是代数，但在第二章一开始就对线性方程组给了一个几何解释．整个第二章中，以方程组为主要的模型讨论了线性空间这一几何对象．以后两章的行列式和矩阵运算虽然都是代数，但都从几何模型引入，再讲代数算法．而更后面的线性变换、内积本来就是几何对象，重点讲怎样将它们归结为矩阵运算 (相似和相合) 来处理，而在讲矩阵运算时又时时指出这些运算的几何背景，始终在几何观点的指引下进行运算．为了实现几何与矩阵的左右逢源，我们突出了向量空间的同构的应用：在选取一定的基之后将每个向量对应于它的坐标，从而将有限维线性空间同构到数组空间来处理，在讲坐标变换、线性变换、内积时都坚持这一处理模式，将线性空间的各种问题都归结为数组空间的问题来处理，通过矩阵运算来解决．

矩阵的相似标准形的算法和证明，是线性代数中最困难的问题．本书第 7 章中对这一问题分别给出了两种不同的解决方式．第一种方式是通过解线性方程组尝试和探索求若当标准形的矩阵算法，从中提炼出若当标准形理论的几何证明，并在这个几何理论指导下给出了通过解线性方程组求若当标准形和过渡矩阵的一般算法．这样一种处理方式从理论上和算法上都可以成为一个独立完整的体系，不需要再用 λ 矩阵来补充．而在此之后再利用 λ 矩阵的相抵 (也称为“等价”) 来研究相似标准形则是在更高观点下的另外一种独立的体系．这个更高观点就是模的观点，将线性变换 $\mathbf{A}$ 作用的空间 V 看成数域上一元多项式环 $F[\lambda]$ 上的模，通过研究这个模的的分解来研究 $\mathbf{A}$ 的标准形矩阵，其中的计算归结为多项式环 $F[\lambda]$ 上的矩阵 (即 λ 矩阵) 的相抵．但是，在线性代数课程中一般不讲模的概念，因此就在讲了 λ 矩阵的相抵变换之后采用了一个较为“初等”的方式来将 λ 矩阵相抵的结论转化为数域上的矩阵的相似标准形的结论．这样的处理方式，回避了较高级的名词“模”而只用到较低级的知识 — 多项式矩阵的运算．这在逻辑上当然是对的．然而，由于它在回避“模”这个名称时将它所蕴涵的几何想法也完全抛弃，整个处理过程变得没有想法而只有运算．为什么在研究矩阵的相似的时候需要研究 λ 矩阵？ λ 矩阵的相抵怎样转化为矩阵的相似？显得莫名其妙，好象纯属偶然．我们的观点是，“模”这个高级的名称可以回避，但它的思想却并非难懂反而非常精彩，可以借用．实际上，高级的东西未必比低级的东西难懂，甚至往往还更容易懂．例如，用方程来解应用题就往往比算术方法容易，用微积分基本定理求图形的面积也比用初等几何的方法直接了当．类似地，用“模”的高级观点来处理线性变换也可以比仅用矩阵运算这样的较低级的观点更容易理解．我们采取的处理方式是：只回避“模”这个“吓人的”名称，而将它实质上的几何思想换一种平易近人的方式来叙述：将线性变换 $\mathbf{A}$ 在每个向量 α 上的作用看作 $\mathbf{A}$ 与 α 的乘法，将 $\mathbf{A}$ 看成向量的“系数”，允许 $\mathbf{A}$ 及其多项式成为矩阵的元素，这样得到的矩阵就是 λ 矩阵，其中的字母 λ 代表的就是线性变换 $\mathbf{A}$．所有的矩阵运算全部在这样的几何观点的指引下进行，每一步运算都有明确的目的，直到最后达到目标，整个过程显得自然而合理，水到渠成．

有一个成语“挂羊头卖狗肉”，说的是将低价产品 (狗肉) 挂一个高价招牌 (羊头)，假冒高价

产品 (羊肉) 来出售，目的是骗取更多的钱．我们正好相反，是“挂狗头卖羊肉”：卖的是高价的“肉”— 按模的思想来处理问题，但不挂“模”这样的高价招牌，免得它将学生吓着了，而换了一个较为平易近人的招牌 —“将线性变换 $\mathbf{A}$ 看成向量的系数”．这样做的目的是为了让学生免受高级名词的恐吓而又能享受到高级名词背后的实惠．这样的“挂狗头卖羊肉”的事情在第三章引入行列式的时候也做过：将行列式看作它的各个列向量的“某种乘积”, 可以按乘法的运算律展开，这实际上是说行列式是各个列向量的多重线性函数．然而，“多重线性函数”这样的高级名词听起来太吓人，我们用“按乘法的运算律展开”这样的中学数学中的初级术语来代替，实质内容相同，但更容易被学生接受．此外，第二章中用“有多余的方程”来讲线性相关，用“货真价实”来讲线性无关，用“方程的真正个数”来讲向量组的秩，也都是“挂狗头卖羊肉”的例子.

一般的教材都将多项式安排在全书的最前面，以便于应用．但是，由于多项式这部分内容的思想方法与其它内容很不相同，这样一种安排好象就是由多项式与线性代数两门不同风格的课程硬凑在一起．在中国科技大学数学系，经过反复考虑和多年的实践，将多项式这部分内容从线性代数中拿出去，与原有的初等数论一起组成一门课程《整数与多项式》，让学生在大学第一学期 (在《线性代数》之前) 学习．由于整数与多项式都有带余除法，很多性质非常类似，放到一起非常合理而融洽．但为了不开设《整数与多项式》课程的学校使用本书，书中还是写了多项式这一章，但没有放在全书最前面，而是放在第 5 章，在讲完线性空间和矩阵的代数运算之后，线性变换之前．之所以不放在第一章是为了一开始就开门见山展开线性代数的主要思想，将线性代数最基本的概念和基本方法 (线性空间及矩阵运算) 一气呵成．这些内容组成线性代数的入门阶段．而从线性变换开始进入一个更高级的新阶段，而多项式对于建立线性变换的理论起着重要作用，因此将多项式放在这两个阶段之间，线性变换之前．当然，这也有一个问题：在线性变换之前的线性空间和行列式计算也用到多项式，而那时还没有讲多项式．但这个问题不大，主要是因为在中学数学中已经学过多项式的一些初步知识，稍加补充就足够在这些章节中使用了．还要指出的是：在中学数学中将多项式的字母理解为未知数，而在大学数学中将多项式的字母看作“未定元”, 只作为一个符号而不代表具体的东西，这让学生不容易理解．如果将多项式放在第一章，实在是举不出例子说明多项式的字母除了代表未知数之外还能代表什么别的东西．然而，在矩阵之后讲多项式，就可以说多项式的字母不但可以代表未知数，还可以代表方阵，以后还可以代表别的对象 (如线性变换). 说它是符号或者未定元，不限定它代表什么对象，是为了允许它代表更多更广泛的对象．在多项式这一章的第 0 节 § 5.0 中，所举的例子就是将多项式的字母换成矩阵产生的令人惊喜的结果，从一个方面来说明为什么要将“未知数”上升为“未定元”．当然，将多项式放在线性代数的中间，确实有将线性代数拦腰斩断的感觉，这是一个缺点．为了降低这种安排的负面影响，我们在多项式这一章中也举了一些利用多项式来解决矩阵和行列式的例子．如果有的教师不喜欢这种安排，可以将多项式这一章提到最前面去，只要将涉及到矩阵或行列式的例子去掉或移到别的适当位置就行了.

本书的以上指导思想和做法是作者在二十多年的教学经验中逐步形成的．特别是已经按照以上所说的指导思想在对连续 8 届学生的教学实践中进行了探索，在探索过程中逐步将这些指导思

想转化为切实可行的实施办法，并且不断根据教学效果进行调整和完善，才形成了现在这样的教学模式，以及反映这种教学模式的教材．在教学过程中，参考了中国科学院数学与系统科学研究院许以超教授编写的《代数学引论》(上海科学技术出版社，1966), 使用了中国科技大学李炯生教授和同济大学查建国教授编写的《线性代数》(中国科学技术大学出版社，1989) 作为教材，在本书编写过程中从这些教材以及北京大学王萼芳教授和首都师范大学石生民教授修订、北京大学数学系几何与代数教研室前代数小组编写的《高等代数》(高等教育出版社，2003 年第 3 版) 中选用了一些习题，特在此对以上教授表示感谢．此外，上海大学王卿文教授为本书的编写提供了一些习题，也在此表示感谢．

目　录

第 1 章　线性方程组的解法

线性代数就是“一次”代数，一个重要论题就是解多元一次方程组．一次方程组也称为线性方程组．本章介绍了解一般的 n 元线性方程组的高斯消去法.

§ 1.0 解多元一次方程组的尝试

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

例 1　在平面直角坐标系中, 作一条抛物线 $y = ax^2 + bx + c$ 经过三个已知点 $(-3, 20), (1, 0), (2, 10)$, 求抛物线方程.

解. 求抛物线 $y = ax^2 + bx + c$ 的方程也就是求 a, b, c 的值．将三个已知点的坐标代入抛物线方程得到

$$\begin{cases} 9a - 3b + c = 20 & (1) \\ a + b + c = 0 & (2) \\ 4a + 2b + c = 10 & (3) \end{cases}$$

用加减消去法消去 c ：

$$(1)\text{式} - (2)\text{式：} \quad 8a - 4b = 20 \qquad (4)$$
$$(3)\text{式} - (2)\text{式：} \quad 3a + b = 10 \qquad (5)$$

再由 (4),(5) 两式消去 b:

$$(5)\text{式} + \frac{1}{4} \times (4)\text{式：} \quad 5a = 15 \qquad (6)$$

由 (6) 解出　$a = 3$.

代入 (5) 解出　$b = 1$.

再将 $a = 3, b = 1$ 代入 (2) 解出　$c = -4$.

经检验可知

$$\begin{cases} a = 3 \\ b = 1 \\ c = -4 \end{cases}$$

是原方程组的解．所求的抛物线方程为 $y = 3x^2 + x - 4$.　□

以上采用的仍然是中学数学解二元一次方程的加减消元法：由原方程分别乘以适当的常数再相加进行消元, 得到未知数更少的新方程. 新方程还可以分别乘以适当的常数再相加进一步消元, 得到含未知数更少的新方程．最后得到只含一个未知数的方程，就能求出解来.

具体说来，在例 1 中由原方程 (1),(2),(3) 得到了新方程 (4),(5),(6). 由 (6),(5),(2) 解出了未知数的值. 因此，可以说是将由 (1),(2),(3) 组成的原方程组 (I) 经过变形成为由 (2),(5),(6) 组成的新方程组：

$$\begin{cases}(1)\\(2)\\(3)\end{cases}\quad\Rightarrow\quad\begin{cases}(2)\\(5)\\(6)\end{cases}$$

由于新方程都是由原方程 (1),(2),(3) 得出来的，原方程组的解一定都是新方程组的解. 但是，反过来要问：由 (2),(5),(6) 得出的解一定是原方程组的解吗？

在中学数学中解分式方程或无理方程的时候，将原方程变形为新方程时，可以保证原方程的解都是新方程的解，但不能保证新方程的解一定是原方程的解，而可能出现增根. 而现在，同样地由原方程得出了新方程，由新方程 (5),(6) 与一个原方程 (2) 一起得出的解是否一定是另外两个原方程 (1),(3) 的解呢？是否可能是“增根”呢？这个问题的答案并不显然. 因此，在例 1 中将所得到的解代入原方程组进行了检验.

但是，如果由新方程组中的方程 (2),(5),(6) 也可以反过来得出另外两个原方程 (1),(3), 那就可以断定 (2),(5),(6) 的公共解一定是原方程组的解而不可能是“增根”, 不须代回原方程组检验.

事实上，由 $(3)-(2)=(5)$ 得 $(2)+(5)=(3)$, 可见由方程 (2),(5) 可以得到 (3). 另一方面，由 $(5)+\frac{1}{4}(4)=(6)$ 得 $4\times(6)-4\times(5)=(4)$, 再由 $(1)-(2)=(4)$ 知 $(2)+(4)=(1)$, 从而 $(2)+4\times(6)-4\times(5)=(1)$, 可见由 (2),(5),(6) 可以得到 (1).

这说明了由 (2),(5),(6) 组成的新方程组确实可以反过来得出原方程 (1),(2) 从而得出原方程组. 新方程组的解也是原方程组的解，因此原方程组与新方程组同解，不会发生“增根”.

但如果解每一个方程组都要这样来论证它不会发生增根，岂不太费事！我们希望能预先设计出方程组的一些简单的基本变形，经过这样的变形得到的新方程组可以反过来得到变形前的方程组，这就可以保证每次变形前后的方程组同解.

§ 1.1 线性方程组的同解变形

1. 线性方程组的定义

定义 1.1.1 n 个未知数 $x_1,x_2,\dots,x_n$ 的如下形式的方程

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b \tag{1.1.1}$$

称为 n 元一次方程，也称 n **元线性方程** (linear equation in n variables), 其中一次项系数 $a_1,\dots,a_n$ 和常数项 b 都是已知数.

如果 $c_1,c_2,\dots,c_n$ 是 n 个数，且将 $x_1=c_1,x_2=c_2,\dots,x_n=c_n$ 代入方程 (1.1.1) 能使方程变为等式，即 $a_1c_1+a_2c_2+\cdots+a_nc_n=b$ 成立，则这一组数 $(c_1,c_2,\dots,c_n)$ 称为方程 (1.1.1) 的一个 **解** (solution). 数组中的第 i 个数 c_i (即 x_i 的取值) 称为解的第 i 分量.

具有同样 n 个未知数 $x_1, x_2, \dots, x_n$ 的若干个线性方程组成的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1.1.2}$$

称为 n **元线性方程组** (linear equations in n variables). 如果一组数 $(c_1, c_2, \dots, c_n)$ 是方程组 (1.1.2) 中所有方程的公共解，也就是说：将 $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ 代入方程组的每一个方程，能使所有这些方程都变为等式，就称这组数 $(c_1, c_2, \dots, c_n)$ 为这个方程组的解. □

注意：在方程组中，一般并不要求方程的个数 m 与未知数个数 n 相等， $m < n, m = n, m > n$ 等三种情况都允许.

2．方程的线性组合

方程的加法: 将两个线性方程

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \qquad (1)$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \qquad (2)$$

左、右两边分别相加得到一个新的方程

$$(a_{11} + a_{21})x_1 + (a_{12} + a_{22})x_2 + \cdots + (a_{1n} + a_{2n})x_n = b_1 + b_2, \qquad (3)$$

称为原来两个方程 (1) 与 (2) 的和．同样可定义若干个方程的和.

方程乘常数: 将方程 $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ 乘以已知常数 λ, 也就是将它的每一项都乘以 λ, 得到一个新方程 $(\lambda a_1)x_1 + (\lambda a_2)x_2 + \cdots + (\lambda a_n)x_n = \lambda b$, 称为原方程的 λ 倍．方程与常数相乘，也称方程的数乘.

注意：方程所乘的常数也可以是 0. 此时得到的新方程 $0 = 0$ 是恒等式，任何一组数都是它的解.

方程的线性组合: 将 m 个方程

$$\begin{array}{cr} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 & (1) \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 & (2) \\ \cdots\cdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m & (m) \end{array}$$

分别乘以 m 个已知常数 $\lambda_1, \lambda_2, \dots, \lambda_m$, 再将所得的 m 个方程相加，得到的新方程

$$\alpha_1x_1 + \alpha_2x_2 + \cdots + \alpha_nx_n = \beta$$

称为原来的 n 个方程 $(1), (2), \dots, (m)$ 的一个 **线性组合** (linear combination), 其中 x_j 的系数 $\alpha_j = \lambda_1a_{1j} + \lambda_2a_{2j} + \cdots + \lambda_ma_{mj}$ $(1 \le j \le n)$ 由各方程中 x_j 的系数分别乘以 $\lambda_1, \lambda_2, \dots, \lambda_m$

再相加得到，常数项 $\beta=\lambda_1b_1+\lambda_2b_2+\cdots+\lambda_mb_m$ 由各方程的常数项分别乘以 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 再相加得到. 各方程所乘的的这些常数 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 称为这个线性组合的系数.

容易验证：如果一组数 $(c_1,c_2,\ldots,c_n)$ 是方程 $(1),(2),\ldots,(m)$ 的公共解，那么它也是这 m 个方程的任一个线性组合的解.

注意：线性组合的系数 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 中可以有些是 0, 甚至可以全部是 0. 如果某些系数是 0, 所得到的线性组合实际上也就是系数不为 0 的那些方程的线性组合. 反过来，从某一组方程 $u_1,u_2,\ldots,u_m$ 中取出一部分方程，比如说取 $u_1,\ldots,u_k$, $k<m$, 作线性组合 $u=\lambda_1u_1+\cdots+\lambda_ku_k$, 则 u 也是原来的所有 m 个方程的线性组合 $u=\lambda_1u_1+\cdots+\lambda_ku_k+0u_{k+1}+\cdots+0u_m$. 特别，如果取某个 $\lambda_i=1$, 其余 $\lambda_j=0$ $(i\geq i)$, 得到的线性组合

$$\lambda_1u_1+\lambda_2u_2+\cdots+\lambda_mu_m=\cdots+0u_{i-1}+u_i+0u_{i+1}+\cdots=u_i.$$

这也就是说，一组方程 $u_1,\ldots,u_m$ 中的每一个方程都是所有这些方程的线性组合。

如果方程组 (II) 中每个方程都是方程组 (I) 中的方程的线性组合，就称方程组 (II) 是方程组 (I) 的线性组合. 此时方程组 (I) 的每一组解也都是方程组 (II) 的解.

如果方程组 (I) 与方程组 (II) 互为线性组合，(即：方程组 (II) 是方程组 (I) 的线性组合，方程组 (I) 也是方程组 (II) 的线性组合), 就称这两个方程组 **等价** (equivalent). 此时两个方程组的解集合相同，或者说：这两个方程组同解. 并且说：将方程组 (I) 变成方程组 (II) 的过程是同解变形.

解方程组的基本方法，就是将方程组进行适当的同解变形，直到最后得到的方程组的解可以写出来为止.

3．基本的同解变形

定理 1.1.1 方程组的以下三种变形是同解变形：

1. 交换其中任意两个方程的位置，其余方程不变.

2. 将任一个方程乘以一个非零的常数 λ, 其余方程不变.

3. 将任一方程的常数倍加到另一方程上，其余方程不变.

证明 为叙述方便，将原方程组 (I) 的 m 个方程依次记为 $u_1,u_2,\ldots,u_m$, 经过变形得到的新方程组 (II) 的 m 个方程依次记为 $v_1,v_2,\ldots,v_m$.

第一步：证明经过这三种变形之后得到的新方程组是原方程组的线性组合.

变形 1: 设在原方程组中将第 i 个方程与第 j 个方程互换位置得到新方程组，其中 $i\neq j$. 则 $v_i=u_j$, $v_j=u_i$, 而 $v_k=u_k$ $(\forall\ k\neq i,j)$, 它们都是原方程 $u_1,\ldots,u_m$ 的线性组合.

变形 2: 设将原方程组的第 i 个方程乘以 λ 得到新方程组. 即 $v_i=\lambda u_i$, $v_k=u_k$ $(\forall\ k\neq i)$, 而 λu_i 和 u_k $(k\neq i)$ 都是原方程组的线性组合.

变形 3: 设将原方程组的第 i 个方程乘以 λ 加到第 j 个方程上得到新方程组，其中 $i\neq j$. 则 $v_j=u_j+\lambda u_i$, $v_k=u_k$ $(\forall\ k\neq j)$, 都是 $u_1,\ldots,u_m$ 的线性组合.

第二步：证明原方程组 (I) 也可以由新方程组 (II) 经过所说的三种类型的变形得到.

变形 1: 设在 (I) 中将第 i 个方程与第 j 个方程互换位置得到 (II), 则在 (II) 中将第 i 个方程与第 j 个方程互换位置得到 (I).

变形 2: 设在 (I) 中将第 i 个方程乘以非零常数 λ 得到 (II), 则在 (II) 中将第 i 个方程乘以非零常数 λ^{-1} 得到 (I).

变形 3: 设在 (I) 中将第 i 个方程乘以常数 λ 加到第 j 个方程上得到 (II), 则在 (II) 中将第 i 个方程乘以常数 $-\lambda$ 加到第 j 个方程上得到 (I).

第三步：第一步已证明 (II) 是 (I) 的线性组合. 第二步证明 (I) 可以由定义所说的三类变形得到，再用第一步的结论知 (I) 也是 (II) 的线性组合. 这就证明了 (I) 与 (II) 等价，因而 (I) 与 (II) 同解. □

定理 1.1.1 所说的线性方程组的三类同解变形，称为线性方程组的 **初等变换** (elementary transformation).

4. 用消去法解方程组

反复利用定理 1.1.1 中所说的三种初等变换，可以将线性方程组消元，求出解来.

例 1 解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 \\ x_1+2x_2 \qquad -5x_4=1 \\ 3x_1-x_2-x_3 \qquad =1 \\ x_1 \qquad +x_3+2x_4=-1 \end{cases}$$

解 将第一个方程的 -1 倍加到第二个方程上，可以将第二个方程的 x_1 的系数变成 0, 从而在第二个方程中消去未知数 x_1. 同样，将第一个方程的 -3 倍、 -1 倍分别加到第三、第四个方程上，可以在第三、四两个方程中消去未知数 x_1. 得到的新方程组为：

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 & (1) \\ \qquad -3x_3-9x_4=4 & (2) \\ -7x_2-10x_3-12x_4=10 & (3) \\ -2x_2-2x_3-2x_4=2 & (4) \end{cases} \tag{I}$$

方程组 (I) 中只有第一个方程含 x_1, 后面三个方程都不含 x_1. 保持第一个方程不变，用与刚才同样的方法对后三个方程组成的方程组进行同解变形，使其中只有最前面一个方程含未知数 x_2, 后面两个方程都不含 x_2. 由于 (I) 中方程 (2) 不含 x_2, 我们将它与含有 x_2 的方程 (4) 互换位置. 得

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 & (1) \\ -2x_2-2x_3-2x_4=2 & (2) \\ -7x_2-10x_3-12x_4=10 & (3) \\ \qquad -3x_3-9x_4=4 & (4) \end{cases} \tag{II}$$

在方程组 (II) 中，方程 (4) 本来就不含 x_2, 剩下还需用方程 (2) 来消去方程 (3) 的 x_2, 为此，只要将方程 (2) 的 $-\frac{7}{2}$ 倍加到方程 (3) 上即可. 但也可先将方程 (2) 乘 $-\frac{1}{2}$ 将方程 (2) 的

x_2 的系数化为 1, 然后再将新的方程 (2) 的 7 倍加到方程 (3) 上以消去 x_2:

$$\left\{\begin{array}{rl} x_1+2x_2+3x_3+4x_4= -3 & (1)\\ x_2\ +x_3\ +x_4= -1 & (2)\\ -3x_3-5x_4= 3 & (3)\\ -3x_3-9x_4= 4 & (4)\end{array}\right. \tag{III}$$

再将方程组 (III) 中的方程 (3) 的 -1 倍加到方程 (4) 上消去 x_3:

$$\left\{\begin{array}{rl} x_1+2x_2+3x_3+4x_4= -3 & (1)\\ x_2\ +x_3\ +x_4= -1 & (2)\\ -3x_3-5x_4= 3 & (3)\\ -4x_4= 1 & (4)\end{array}\right. \tag{IV}$$

最后得到的方程组 (IV) 中，方程 (4) 只含一个未知数 x_4, 可解出

$$x_4 = -\frac{1}{4}.$$

方程 (3) 只含 x_3, x_4, 而 x_4 已解出，于是由 (3) 解出

$$x_3 = \frac{3+5x_4}{-3} = -1 + (-\frac{5}{3})(-\frac{1}{4}) = -\frac{7}{12}.$$

再由 (2) 解出

$$x_2 = -1 - x_3 - x_4 = -1 - (-\frac{7}{12}) - (-\frac{1}{4}) = -\frac{1}{6}.$$

由 (1) 解出

$$x_1 = -3 - 2x_2 - 3x_3 - 4x_4 = -3 - 2(-\frac{1}{6}) - 3(-\frac{7}{12}) - 4(-\frac{1}{4}) = \frac{1}{12}.$$

这就得出原方程组的解为 $(\frac{1}{12}, -\frac{1}{6}, -\frac{7}{12}, -\frac{1}{4})$. $\square$

例 1 主要是利用初等变换由上而下进行消元，将方程组的左边化为“上三角形”; 再由下而上逐次求出所有未知数的值来. 实际上，在消元中起主要作用的是第 3 类初等变换，在第 3 类初等变换不能进行时使用第 1 类初等变换来创造消元的条件. 而第 2 类初等变换主要用来将未知数系数化为 1 以求出解来.

也可继续对已经化成的上三角形方程组 (IV) 作同解变形，由下而上进一步消元，使所有的方程都只含一个未知数，直接得出方程的解来. 具体作法如下：

将前述方程组 (IV) 中的方程 (4) 的 $-\frac{5}{4}$ 倍、$\frac{1}{4}$ 倍、1 倍分别加到方程 (3),(2),(1) 上，消去这三个方程中的未知数 x_4, 得

$$\left\{\begin{array}{rl} x_1+2x_2+3x_3\qquad = -2 & (1)\\ x_2\ +x_3\qquad = -\frac{3}{4} & (2)\\ -3x_3\qquad = \frac{7}{4} & (3)\\ -4x_4= 1 & (4)\end{array}\right. \tag{V}$$

将方程组 (V) 的方程 (3) 的 $\frac{1}{3}$ 倍、 1 倍分别加到方程 (2), (1) 上，消去这两个方程中的 x_3, 得

$$\begin{cases} x_1+2x_2 & = -\frac{1}{4} \quad (1) \\ \quad\ x_2 & = -\frac{1}{6} \quad (2) \\ \qquad -3x_3 & = \frac{7}{4} \quad (3) \\ \qquad\qquad -4x_4 & = 1 \quad (4) \end{cases} \tag{VI}$$

再将方程组 (VI) 的方程 (2) 的 -2 倍加到方程 (1) 上，消去方程 (1) 中的 x_2, 得

$$\begin{cases} x_1 & = -\frac{1}{12} \quad (1) \\ \quad x_2 & = -\frac{1}{6} \quad (2) \\ \qquad -3x_3 & = \frac{7}{4} \quad (3) \\ \qquad\qquad -4x_4 & = 1 \quad (4) \end{cases} \tag{VII}$$

方程组 (VI) 的左边成 “对角形”, 每个方程只含一个未知数. 将各方程分别除以所含未知数的系数，也就是将方程 (1) 至 (4) 分别乘以 $1, 1, -\frac{1}{3}, -\frac{1}{4}$ (这是第二类基本变形), 可以将各未知数的系数化为 1, 得到方程组

$$\begin{cases} x_1 & = -\frac{1}{12} \quad (1) \\ \quad x_2 & = -\frac{1}{6} \quad (2) \\ \qquad x_3 & = -\frac{7}{12} \quad (3) \\ \qquad\quad x_4 & = -\frac{1}{4} \quad (4) \end{cases} \tag{VIII}$$

它的解可以立即写出，为 $(\frac{1}{12}, -\frac{1}{6}, -\frac{7}{12}, -\frac{1}{4})$.

以上是方程组有唯一解的例子. 解的每个分量都是由方程组的系数经过加、减、乘、除四则运算得到. 如果原方程组的系数都是实数，由于实数集合对加、减、乘、除四则运算封闭 (当然做除法时除数不允许为 0), 方程组的唯一解的所有分量就都是实数. 同样，有理数集合对加、减、乘、除运算也封闭, 因此有理系数线性方程组的唯一解的分量也都是有理数. 还可以考虑其它的系数范围，只要它们对加、减、乘、除四则运算封闭.

定义 1.1.2 设 F 是复数集合的子集，包含 0 和 1, 并且在加、减、乘、除运算下封闭 (做除法时除数不为 0), 就称 F 是 **数域** (number field). □

例如，复数集合 C 、实数集合 R 、有理数集合 Q 都是数域.

按照这个术语，我们有：如果线性方程组的系数都在某个数域 F 的范围内, 并且这个方程组有唯一解，则解的分量也都在 F 的范围内.

习题 1.1

1. 用消元法解线性方程组

(1) $\begin{cases} x_1+2x_2+3x_3=1 \\ 2x_1+2x_2+5x_3=2 \\ 3x_1+5x_2+x_3=3 \end{cases}$; (2) $\begin{cases} \quad\;\; x_2+x_3+x_4=1 \\ x_1 \quad\; +x_3+x_4=2 \\ x_1+x_2 \quad\; +x_4=3 \\ x_1+x_2+x_3 \quad\; =4 \end{cases}$.

2. (1) 求证：如果复数集合的子集 P 包含至少一个非零数，并且对加、减、乘、除 (除数不为 0) 封闭，则 P 包含 $0, 1$, 从而是数域.

   (2) 求证：所有的数域都包含有理数域.

   (3) 求证：集合 $F=\{a+b\sqrt{2} \mid a, b \in Q\}$ 是数域. （其中 Q 是有理数域.）

   (4) 试求包含 $\sqrt[3]{2}$ 的最小的数域.

3. 证明：(1) 线性组合的传递性：如果方程组 II 是方程组 I 的线性组合，方程组 III 是方程组 II 的线性组合，则方程组 III 是方程组 I 的线性组合.

   (2) 等价的传递性：如果方程组 I 与方程组 II 等价，方程组 II 与方程组 III 等价，则方程组 I 与方程组 III 等价.

§ 1.2 矩阵消元法

以后，凡是谈到线性方程组，总假定它的系数全都在某个数域 F 中，称它为 F 上的线性方程组. 解这个线性方程组的过程就只涉及到 F 中的数之间的加、减、乘、除四则运算.

考察 §1.1 例 1 中解线性方程组的过程可以发现，在解方程组的过程中，实际上只对各方程中各项的系数进行了运算 (加、减、乘、除运算), 每次将代表未知数的字母抄写一遍实际上是一种累赘. 为了书写的简便，更为了突出解方程组中本质的东西 — 系数的运算，我们采用分离系数法，将线性方程组中代表未知数的字母略去，将等号也略去，只写出各方程的各系数. 将每个方程的各项系数从左到右依次写成一行，将各方程中同一个未知数的系数上下对齐，常数项也上下对齐，这样得到一个矩形数表，来表示这个方程组. 一般地，线性方程组

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

用矩形数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

来表示.（将数表用括号括起来是表示所有的数组成一个整体.）

定义 1.2.1 对任意自然数 m, n, 由数域 F 中 $m \times n$ 个数排成 m 行、n 列所得到的数表，称为 F 上的 $m \times n$ **矩阵** (matrix). 数表中的每个数称为矩阵的一个 **元素** (element), 也称为矩阵的一个 **分量** (entry), 其中排在第 i 行第 j 列的数称为矩阵的第 (i,j) 元或第 (i,j) 分量. F 上全体 $m \times n$ 矩阵的集合记作 $F^{m\times n}$. $\square$

按照这个定义，由 m 个 n 元线性方程组成的方程组用 $F^{m\times(n+1)}$ 中一个矩阵表示. 它的 m 行分别代表 m 个方程，前 n 列分别是 n 个未知数的系数，最后一列是常数项.

矩阵的第 i 行 $(a_{i1}, a_{i2}, \dots, a_{in}, b_i)$ 表示方程组中第 i 个方程 $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$. 这一行的 $n+1$ 个数中每个数到底是哪个未知数的系数或是常数项，完全由它在数组中的位置来决定. 因此，各数的排列顺序不能搞乱. 因此，我们说，这个线性方程是由它的系数组成的有序数组表示的.

定义 1.2.2 由数域 F 中 n 个数 a_i $(1 \le i \le n)$ 排成的有序数组 $(a_1, a_2, \dots, a_n)$ 称为 F 上的 n **维向量** (n-dimensional vector), 也称 n 维数组向量，a_i 称为它的第 i 分量. 所有分量都为 0 的向量 $(0, \dots, 0)$ 称为 **零向量** (zero vector), 记作 $\mathbf{0}$. (通常也将零向量记作 0, 从上下文可以知道它是表示零向量还是表示数 0, 不会混淆.) F 上全体 n 维向量组成的集合称为 F 上的 n **维向量空间** (n-dimensional vector space), 记作 F^n. $\square$

按照这个定义，F 上一个 n 元线性方程用 F^{n+1} 中的一个向量表示. 而 n 元线性方程组的每一组解是 F^n 中的一个向量. $F^{m\times n}$ 中每个矩阵的每一行是一个 n 维向量，每一列是一个 m 维向量.

在书写向量的时候，有时将 n 维向量的 n 个分量从左到右横写成一行，如 $(a_1, \dots, a_n)$, 此时称它为 **行向量** (row vector). n 维行向量也就是 1 行 n 列的矩阵，此时可将 F^n 写为 $F^{1\times n}$. 另一方面，也可将 n 维向量 $(a_1, \dots, a_n)$ 的 n 个分量从上到下竖写为一列

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

此时称它为 **列向量** (column vector), 它也就是 n 行 1 列的矩阵. 此时 F^n 可写为 $F^{n\times 1}$.

特别，每个线性方程用行向量表示. 方程组的解通常也可以用行向量表示，以节省书写篇幅. 但我们将看到，作理论分析时，用列向量来表示方程组的解有它的优越性.

将线性方程用向量表示、线性方程组用矩阵表示之后，线性方程的加法、数乘、线性组合等运算，以及线性方程组的初等变换，就对应于向量的如下运算和矩阵的如下基本变形.

定义 1.2.3 (1) (向量的加法) 设 $\alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n)$ 是同一个数域 F 上两个 n 维向量，将它们按分量相加得到的向量 $(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \in F^n$ 称为这两个向量的和，记作 $\alpha + \beta$. 同样可以定义多个（有限个）向量 $\alpha_1, \alpha_2, \dots, \alpha_m \in F^n$ 的和

$\alpha_1+\alpha_2+\cdots+\alpha_m\in F^n$, 它的第 j 分量 $(1\le j\le n)$ 等于各 α_i $(1\le i\le m)$ 的第 j 分量之和.

(2) (向量与数的乘法) 将任一 $\lambda\in F$ 遍乘任一 $\alpha=(a_1,a_2,\ldots,a_n)\in F^n$ 的各分量，所得到的向量 $(\lambda a_1,\lambda a_2,\ldots,\lambda a_n)\in F^n$ 称为 α 的 λ 倍，记作 $\lambda\alpha$.

(3) (向量的线性组合) 将 F^n 中的一组向量 $\alpha_1,\ldots,\alpha_m$ 分别乘以 F 中的一组数 $\lambda_1,\ldots,\lambda_m$ 再相加，得到的向量 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_m\in F^n$ 称为向量 $\alpha_1,\alpha_2,\ldots,\alpha_m$ 的线性组合. □

定义 1.2.4 设 A,B 是 $F^{m\times n}$ 中的两个矩阵. 如果 B 的每一行都是 A 的行的线性组合，A 的每一行也是 B 的行的线性组合，就称两个矩阵 **行等价** (row equivalent). □

定理 1.2.1 设 F 上的矩阵 A 经过以下变形之一变成矩阵 B, 则 A 与 B 行等价：

1. 将某两行互换位置;
2. 用 F 中某个非零的数乘以某行;
3. 将某行的常数倍加到另一行上. □

定义 1.2.5 定理 1.3.1 中所说的三类变形称为矩阵的 **初等行变换** (elementary transformation of rows).

矩阵的三类初等行变换对应于线性方程组的三类基本同解变形. 用基本同解变形对线性方程组消元的过程，也就是用初等行变换将尽可能多的矩阵元素化为零的过程.

为了叙述方便，我们用矩阵 A 到 B 的箭头来表示 A 经过初等行变换变为 B, 箭头上方注明所用的是哪一个变换：

$$(1)\ A\xrightarrow{(i,j)}B\quad(2)\ A\xrightarrow{\lambda(i)}B\quad(3)\ A\xrightarrow{\lambda(i)+(j)}B$$

箭头上的 (i,j) 表示将第 i 行与第 j 行互换， $\lambda(i)$ 表示用非零数 λ 乘以第 i 行， $\lambda(i)+(j)$ 表示将第 i 行的 λ 倍加到第 j 行上.

下面用矩阵的初等行变换重新来解 § 1.1 例 1 的方程组.

例 1 解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3\\ x_1+2x_2\qquad\ -5x_4=1\\ 3x_1-x_2-x_3\qquad\ =1\\ x_1\qquad\ +x_3+2x_4=-1\end{cases}.$$

解. 将方程组用矩阵表示为

$$\begin{pmatrix}1&2&3&4&-3\\1&2&0&-5&1\\3&-1&-1&0&1\\1&0&1&2&-1\end{pmatrix}.$$

然后用矩阵的初等行变换进行消元 (用箭头表示):

$$\xrightarrow{-(1)+(2),-3(1)+(3),-(1)+(4)}\begin{pmatrix}1&2&3&4&-3\\0&0&-3&-9&4\\0&-7&-10&-12&10\\0&-2&-2&-2&2\end{pmatrix}.$$

$$\xrightarrow{(2,4),-\frac{1}{2}(2),7(2)+(3),-2(2)+(1)}\begin{pmatrix}1&0&1&2&-1\\0&1&1&1&-1\\0&0&-3&-5&3\\0&0&-3&-9&4\end{pmatrix}.$$

注意，此处的消元顺序与 § 1.1 中略有不同，不但用第 2 行 (代表方程 (2)) 的倍数向下消去第 3 行的第 2 元素，而且还向上消去了第 1 行的第 2 元素，这样就使第 2 列除第 2 行的元素外都是 0, 也就是说：未知数 x_2 只在第 2 个方程中出现.

$$\xrightarrow{-(3)+(4)}\begin{pmatrix}1&0&1&2&-1\\0&1&1&1&-1\\0&0&-3&-5&3\\0&0&0&-4&1\end{pmatrix}$$

$$\xrightarrow{-\frac{1}{4}(4),-2(4)+(1),-(4)+(2),5(4)+(3)}\begin{pmatrix}1&0&1&0&-\frac{1}{2}\\0&1&1&0&-\frac{3}{4}\\0&0&-3&0&\frac{7}{4}\\0&0&0&1&-\frac{1}{4}\end{pmatrix}$$

$$\xrightarrow{-\frac{1}{3}(3),-(3)+(1),-(3)+(2)}\begin{pmatrix}1&0&0&0&\frac{1}{12}\\0&1&0&0&-\frac{1}{6}\\0&0&1&0&-\frac{7}{12}\\0&0&0&1&-\frac{1}{4}\end{pmatrix}$$

写成方程的形式，即

$$\begin{cases}x_1 \qquad\qquad = \frac{1}{12}\\ \quad x_2 \qquad\quad = -\frac{1}{6}\\ \qquad\quad x_3 \quad = -\frac{7}{12}\\ \qquad\qquad x_4 = -\frac{1}{4}\end{cases}.$$

它的解显然是 $(\frac{1}{12},-\frac{1}{6},-\frac{7}{12},-\frac{1}{4})$. □

并非所有的方程组都像例 1 那样有唯一解.

例 2 解线性方程组

$$\begin{cases}x_1+2x_2+3x_3+4x_4=-3\\x_1+2x_2-5x_4=1\\2x_1+4x_2-3x_3-19x_4=6\\3x_1+6x_2-3x_3-24x_4=7\end{cases}.$$

解

$$\begin{pmatrix}1&2&3&4&-3\\1&2&0&-5&1\\2&4&-3&-19&6\\3&6&-3&-24&7\end{pmatrix}\xrightarrow{-(1)+(2),-2(1)+(3),-3(1)+(4)}\begin{pmatrix}1&2&3&4&-3\\0&0&-3&-9&4\\0&0&-9&-27&12\\0&0&-12&-36&16\end{pmatrix}$$

$$\xrightarrow{-3(2)+(3),-4(2)+(4),(2)+(1),-\frac{1}{3}(2)}\begin{pmatrix}1&2&0&-5&1\\0&0&1&3&-\frac{4}{3}\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

重新写成方程的形式. 注意第 3, 4 两个方程中所有的未知数及常数项都被消掉了, 成为恒等式 $0=0$, 可以从方程组中删去而不会改变方程组的解. 于是方程组成为:

$$\begin{cases}x_1+2x_2\quad-5x_4=1\\ \qquad\qquad x_3+3x_4=-\frac{4}{3}\end{cases}$$

未知数 x_1 只含于第 1 个方程，未知数 x_3 只含于第 2 个方程. 将两个方程中含其余两个未知数 x_2, x_4 的项移到等号右边，方程组化为

$$\begin{cases}x_1=1-2x_2+5x_4 & (1)\\ x_3=-\frac{4}{3}-3x_4 & (2)\end{cases}$$

等号右边的未知数 x_2, x_4 可以任意取值. 让 x_2 取定任意值 t_1, x_4 取定任意值 t_2, 则由 (1),(2) 两式分别可算出 x_1, x_3 的值，得到一组解. 当 x_2, x_4 取遍所有的可能的值时就得到方程组的所有的解. 因此，方程组的解集合为

$$\{(1-2t_1+5t_2, t_1, -\frac{4}{3}-3t_2, t_2)\mid t_1, t_2\text{在允许范围内任意取值}\}.\qquad\square$$

例 2 的方程组的解的一般形式 $(1-2t_1+5t_2, t_1, -\frac{4}{3}-3t_2, t_2)$ 称为方程组的 **通解** (general solutions)，当其中的独立参数 t_1, t_2 取遍允许范围内所有可能的值时，就得到方程组的所有的解. 当 t_1, t_2 取定一组具体的值时，就得到方程组的一个解，称为方程组的一个 **特解** (special solution). 如果不加限制， t_1, t_2 可以独立取遍任意复数值. 但有时由于方程组所反映的问题本身的限制，只允许 t_1, t_2 在实数范围或有理数范围内取值. 注意通解的每个分量是独立参数 t_1, t_2 的一次多项式，其系数由原方程组的系数经加、减、乘、除得出. 如果原方程组的系数都在某个数域 F, 则这些一次多项式的系数也都在 F 内，只要独立参数也同样在 F 的范围内取值，则解的所有分量也都在 F 内. 因此，系数在数域 F 内的线性方程组的求解可以限制在 F 的范围内进行，它的通解中的独立参数允许取遍 F 内的所有的值.

例 3　解线性方程组

$$\begin{cases}x_1+2x_2+3x_3+4x_4=0\\ x_1+2x_2\qquad\ -5x_4=0\\ 2x_1+4x_2-3x_3-19x_4=0\\ 3x_1+6x_2-3x_3-24x_4=0\end{cases}.$$

解 注意这个方程组的左边与例 2 的方程组完全一样，只是右边的常数项全部换成了 0. 经过与例 2 完全相同的变换得到通解为 $(-2t_1+5t_2, t_1, -3t_2, t_2)$. □

由于例 3 的方程组中每个方程的常数项都是 0, 表示这个方程组的矩阵

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 0 \\ 1 & 2 & 0 & -5 & 0 \\ 2 & 4 & -3 & -19 & 0 \\ 3 & 6 & -3 & -24 & 0 \end{pmatrix}$$

的最后一列全部是 0, 在进行初等行变换的过程中这一列也始终是 0, 对于计算过程不产生影响. 因此可以将表示常数项的这一列略去不写，直接用各个方程中未知数的系数组成的矩阵

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 0 & -5 \\ 2 & 4 & -3 & -19 \\ 3 & 6 & -3 & -24 \end{pmatrix}$$

来表示方程组. 一般地说，任何一个线性方程组的各方程中未知数系数组成的矩阵称为这个方程组的 **系数矩阵** (coefficient matrix). 当常数项全部为 0 时，用系数矩阵可以完全表示这个线性方程组. 对系数矩阵进行初等行变换就可以求出方程组的解.

例 3 的方程组的常数项全部为 0, 显然 $(0,0,0,0)$ 是它的解，它也就是在通解中将独立参数 t_1, t_2 都取零值得到的解.

将例 3 的通解 $(-2t_1+5t_2, t_1, -3t_2, t_2)$ 写成列向量的形式，容易看出它可以写成两个向量的线性组合：

$$\begin{pmatrix} -2t_1 & +5t_2 \\ t_1 & \\ & -3t_2 \\ & t_2 \end{pmatrix} = t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 5 \\ 0 \\ -3 \\ 1 \end{pmatrix},$$

其中的组合系数 t_1, t_2 可以独立地取遍 F 中所有的值. 因此，例 3 方程组的解集由向量

$$\alpha_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad 与 \quad \alpha_2 = \begin{pmatrix} 5 \\ 0 \\ -3 \\ 1 \end{pmatrix}$$

在 F 上所有的线性组合组成. 向量 α_1, α_2 本身也都是解：在通解中取 $t_1=1, t_2=0$ 就得到解 α_1, 取 $t_1=0, t_2=1$ 就得到解 α_2.

一般地，如果线性方程组中所有的方程的常数项都是 0, 就称这样的线性方程组为 **齐次线性方程组** (homogeneous liear equations). (“齐次” 是所有的方程中的非零项都是一次项，次数

“整齐”– 都等于 1.）显然，n 维零向量 $(0,\dots,0)$ 一定是齐次线性方程组的解，称为 **零解** (zero solution), 也称为 **平凡解** (trivial solution). (说它是“平凡”解，意思是不需要解方程组就可以立即写出来的解，“得来全不费功夫”.) 不难验证，齐次线性方程组的解集由其中某几个特解的全体线性组合组成.

如果线性方程组中有至少一个方程的常数项不为 0, 就称为 **非齐次线性方程组** (inhomogeneous linear equations). 例如，例 1, 例 2 都是非齐次线性方程组. 显然，零向量不是非齐次线性方程组的解.

前面的例子的线性方程组都有解. 但线性方程组也可能无解.

例 4　解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 \\ x_1+2x_2-5x_4=1 \\ 2x_1+4x_2-3x_3-19x_4=6 \\ 3x_1+6x_2-3x_3-24x_4=6 \end{cases}.$$

解　经过与例 2 同样的方法消元之后化为

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3 & (1) \\ -3x_3-9x_4=4 & (2) \\ 0=0 & (3) \\ 0=-1 & (4) \end{cases}.$$

其中方程 (4) $0=-1$ 的等号不可能成立，方程组无解.　□

习题 1.2

1. 用矩阵消元法解线性方程组，并将其中 (1),(2) 的过程及结果与习题 1.1 第 1 题比较.

(1) $\begin{cases} x_1+2x_2+3x_3=1 \\ 2x_1+2x_2+5x_3=2 \\ 3x_1+5x_2+x_3=3 \end{cases}$;　(2) $\begin{cases} x_2+x_3+x_4=1 \\ x_1+x_3+x_4=2 \\ x_1+x_2+x_4=3 \\ x_1+x_2+x_3=4 \end{cases}$;

(3) $\begin{cases} 2x_1+x_2-5x_3+x_4=8 \\ x_1-3x_2-6x_4=9 \\ 2x_2-x_3+2x_4=-5 \\ x_1+4x_2-7x_3+6x_4=0 \end{cases}$;　(4) $\begin{cases} x_1+3x_2-5x_3-5x_4=2 \\ x_1+2x_2+2x_3-2x_4+x_5=-2 \\ 2x_1+x_2+3x_3-3x_4=2 \\ x_1-4x_2+x_3+x_4-x_5=3 \\ x_1+3x_3-x_4+x_5=1 \end{cases}$.

2. 在空间直角坐标系中, 求三个平面 $9x-3y+z=20$, $x+y+z=0$ 和 $-x+2y+z=-10$ 的公共点集合.

3. 已知两个变量 x, y 之间有某种函数关系 $y = f(x)$, 并且有如下对应值

x	1	2	3	4
y	2	7	16	29

问 y 是否可能是 x 的二次函数？如果可能，试求出满足要求的二次函数。

4. 在实数范围内解线性方程组

$$\begin{cases} x+3y+2z=4 \\ 2x+5y-3z=-1 \\ 4x+11y+z=7 \end{cases}.$$

这个方程组的解集在 3 维空间中的图象 Π 是什么？

将这个方程组的常数项全部变成 0, 得到的方程组的解集在 3 维空间中的图象 Π_0 是什么？Π_0 与 Π 有什么关系？

§ 1.3 一般线性方程组的消元解法

现在来叙述用消元法解一般的线性方程组的操作步骤，并讨论解的所有可能的情况.

一般的线性方程组具有形式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.3.1}$$

先将它写成矩阵形式

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{1.3.2}$$

然后按下面的步骤用初等行变换对矩阵消元.

1. 矩阵的第一列元素不能全为 0, 否则方程组中所有的方程都不含未知数 x_1. 如果 $a_{11} = 0$, 第一列必有某个 $a_{i1} \neq 0$, $i \geq 2$. 将矩阵的第 1 行与第 i 行互换，化为 $a_{11} \neq 0$ 的情形. 再将第 1 行乘以非零常数 a_{11}^{-1} 可化为 $a_{11} = 1$ 的情形. 对每个 $i \geq 2$, 将第 1 行的 $-a_{i1}$ 倍加到第 i 行上，可以将第 i 行的第一个元素化为 0. 这就将第 1 列中除了第 1 行为 1 外其余各行的元素全部化成了 0. 矩阵化为

$$\begin{pmatrix} 1 & a_{12}^{(1)} & \cdots & b_1^{(1)} \\ 0 & a_{22}^{(1)} & \cdots & b_2^{(1)} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{m2}^{(1)} & \cdots & b_m^{(1)} \end{pmatrix} \tag{1.3.3}$$

的形状. (这意味着除了第一个方程含有未知数 x_1 、其系数是 1 外，其余所有的方程都不含未知数 x_1.)

2. 如果矩阵 (1.3.3) 只有一行，或者第一行以下的各行 (即第 2 至 m 行）的前 n 列元素全都等于 0 (不管最后一列是否为 0), 则消元过程结束. (这意味着只有第一个方程含有未知数；其余各方程都不含未知数，左边为 0.)

若不然，在第 2 行到第 m 行的前 n 列元素组成的矩阵

$$\begin{pmatrix} 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{m2}^{(1)} & \cdots & a_{mn}^{(1)} \end{pmatrix}$$

中必有某个元素不为 0. 从中找出最左边的不为零的列，设为第 j_2 列. 当然 $j_2 > 1$. 则矩阵 (1.3.3) 具有形状

$$\begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1,j_2-1}^{(1)} & a_{1j_2}^{(1)} & \cdots & b_1^{(1)} \\ 0 & \cdots & 0 & a_{2j_2}^{(1)} & \cdots & b_2^{(1)} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & a_{m,j_2}^{(1)} & \cdots & b_m^{(1)} \end{pmatrix}$$

其中某个 $a_{i_2j_2}^{(1)} \neq 0$, $i_2 \geq 2$. 如果 $a_{2j_2}^{(1)} = 0$, 则 $i_2 > 2$, 将矩阵 (1.3.3) 的第 2 行与第 i_2 行互换可化为 $a_{2,j_2}^{(1)} \neq 0$ 的情形。再将矩阵 (1.3.3) 的第 2 行乘以 $(a_{2j_2}^{(1)})^{-1}$ 可化 $a_{2j_2}^{(1)}$ 为 1. 故总可设 $a_{2j_2}^{(1)} = 1$. 现在可以进行下一步的消元:

对每个 $i \neq 2$, (即 $i = 1$ 或 $3 \leq i \leq m$), 将第 2 行的 $-a_{ij_2}^{(1)}$ 倍加到第 i 行上可将 $a_{ij_2}^{(1)}$ 化为 0. 这样就将第 j_2 列中除第 2 行以外的元素全都化成了 0. 矩阵化为如下的形状

$$\begin{pmatrix} a_{11}^{(2)} & \cdots & a_{1,j_2-1}^{(2)} & 0 & a_{1,j_2+1}^{(2)} & \cdots & a_{1n}^{(2)} & b_1^{(2)} \\ 0 & \cdots & 0 & a_{2j_2}^{(2)} & a_{2,j_2+1}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & \cdots & 0 & 0 & a_{3,j_2+1}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & a_{m,j_2+1}^{(2)} & \cdots & a_{mn}^{(2)} & b_m^{(2)} \end{pmatrix} \tag{1.3.4}$$

其中 $a_{11}^{(2)} = a_{2j_2}^{(2)} = 1$.

3. 如果矩阵 (1.3.4) 只有 2 行，或者除了前 2 行以外的以下各行的前 n 列元素 $a_{ij}^{(2)}$ $(i \geq 3, 1 \leq j \leq n)$ 全都等于 0, 则消元过程结束. 若不然，在第 3 行到第 m 行的前 n 列元素组成的矩阵

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & a_{3,j_2+1}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & a_{m,j_2+1}^{(2)} & \cdots & a_{mn}^{(2)} \end{pmatrix}$$

中必有某个元素不为 0, 从中可以找到最左边的不为零的列，设为第 j_3 列. 当然 $j_3 > j_2$. 则矩阵

(1.3.4) 具有形状

$$\begin{pmatrix} a_{11}^{(2)} & \cdots & a_{1,j_2-1}^{(2)} & 0 & a_{1,j_2+1}^{(2)} & \cdots & a_{1j_3}^{(2)} & \cdots & b_1^{(2)} \\ 0 & \cdots & 0 & a_{2j_2}^{(2)} & a_{2,j_2+1}^{(2)} & \cdots & a_{2j_3}^{(2)} & \cdots & b_2^{(2)} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & a_{3,j_3}^{(2)} & \cdots & b_3^{(2)} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & a_{m,j_3}^{(2)} & \cdots & b_m^{(2)} \end{pmatrix}$$

其中某个 $a_{i_3j_3}^{(2)} \neq 0$, $i_3 \geq 3$. 如果 $a_{3j_3}^{(2)} = 0$, 则 $i_3 > 3$, 将第 3 行与第 i_3 行互换可化为 $a_{3j_3}^{(2)} \neq 0$ 的情形, 再将第 3 行乘以 $(a_{3j_3}^{(2)})^{-1}$ 可化为 $a_{3j_3}^{(2)} = 1$ 的情形. 故总可设 $a_{3j_3} = 1$. 现在可以进行下一步的消元:

对每个 $i \neq 3$, (即 $i \leq 2$ 或 $4 \leq i \leq m$), 将第 3 行的 $-a_{ij_3}^{(2)}$ 倍加到第 i 行上可将 a_{ij_3} 化为 0. 这样就将第 j_3 列中除第 3 行以外的元素全都化成了 0. 矩阵化为如下形状

$$\begin{pmatrix} a_{11}^{(3)} & \cdots & a_{1,j_2-1}^{(3)} & 0 & a_{1,j_2+1}^{(3)} & \cdots & a_{1,j_3-1}^{(3)} & 0 & a_{1,j_3+1}^{(3)} & \cdots & a_{1n}^{(3)} & b_1^{(3)} \\ 0 & \cdots & 0 & a_{2j_2}^{(3)} & a_{2,j_2+1}^{(3)} & \cdots & a_{2,j_3-1}^{(3)} & 0 & a_{2,j_3+1}^{(3)} & \cdots & a_{2n}^{(3)} & b_2^{(3)} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & a_{3j_3}^{(3)} & a_{3,j_3+1}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & a_{4,j_3+1}^{(3)} & \cdots & a_{4n}^{(3)} & b_4^{(3)} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & a_{m,j_3+1}^{(3)} & \cdots & a_{mn}^{(3)} & b_m^{(3)} \end{pmatrix}$$

其中 $a_{11}^{(3)} = a_{2j_2}^{(3)} = a_{3j_3}^{(3)} = 1$.

将上述过程重复 k 次之后, 矩阵被化为如下的阶梯形 (矩阵中的空格表示该位置的元素是 0):

$$\begin{pmatrix} a_{11}^{(k)} & \cdots & 0 & a_{1,j_2+1}^{(k)} & \cdots & 0 & a_{1,j_k+1}^{(k)} & \cdots \\ & & a_{2j_2}^{(k)} & \cdots & \cdots & 0 & a_{2,j_k+1}^{(k)} & \cdots \\ & & & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & & a_{kj_k}^{(k)} & a_{k,j_k+1}^{(k)} & \cdots \\ & & & & & & \vdots & \cdots \\ & & & & & & a_{m,j_k+1}^{(k)} & \cdots \end{pmatrix} \tag{1.3.5}$$

其中前 k 行每行的第一个非零元素 $a_{11}^{(k)} = a_{2j_2}^{(k)} = \cdots = a_{kj_k}^{(k)} = 1$, $1 < j_2 < \cdots < j_k \leq n$. $a_{11}^{(k)}, a_{2j_2}^{(k)}, \ldots, a_{kj_k}^{(k)}$ 这 k 个非零元素的左方、上方和下方全是 0, 而第 k 行以下各行 (即第 $k+1$ 至 m 行) 的前 j_k 列元素全是 0.

注意：我们约定，在书写矩阵时，空白位置都表示这个位置的矩阵元素是 0. 以上的矩阵中，左下角的空白位置的元素都是 0.

如果此时第 $k+1$ 至 m 行的前 n 列全是 0, 则消元过程结束。否则，重复前面的步骤，在第 $k+1$ 至 m 行组成的矩阵中找出最左边的非零列，设为第 j_{k+1} 列，$(j_k < j_{k+1} \leq n)$, 其中含有某个元素 $a_{i,j_{k+1}}^{(k)} \neq 0$ $(i \geq k+1)$. 当 $a_{k+1,j_{k+1}}^{(k)} = 0$ 时还可将矩阵的第 $k+1$ 行与第 i 行互换

使 $a^{(k)}_{k+1,j_{k+1}}\neq 0$, 并进而将第 $k+1$ 行乘以 $(a^{(k)}_{k+1,j_{k+1}})^{-1}$ 化为 $a^{(k)}_{k+1,j_{k+1}}=1$ 的情形。在此基础上进行下一步的消元：将第 $k+1$ 行的 $-a^{(k)}_{i,j_{k+1}}$ 倍加到每个第 i 行 $(i\neq k+1)$, 将 $a^{(k)}_{k+1,j_{k+1}}$ 的上方和下方的元素全部化为 0.

以上过程重复下去，直到将矩阵化为下面的形状为止 (矩阵中的空格表示该位置的元素是 0):

$$\begin{pmatrix}
\alpha_{11} & \cdots & 0 & \alpha_{1,j_2+1} & \cdots & 0 & \alpha_{1,j_r+1} & \cdots & \alpha_{1n} & \beta_1\\
 & & \alpha_{2j_2} & \cdots & \cdots & 0 & \alpha_{2,j_r+1} & \cdots & \alpha_{2n} & \beta_2\\
 & & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
 & & & & & \alpha_{rj_r} & \alpha_{r,j_r+1} & \cdots & \alpha_{rn} & \beta_r\\
 & & & & & & & & & \vdots\\
 & & & & & & & & & \beta_m
\end{pmatrix}\tag{1.3.6}$$

其中前 r 行每行的第一个非零元素 $\alpha_{11}=\alpha_{2j_2}=\cdots=\alpha_{kj_k}=1$, $1<j_2<\cdots<j_r\leq n$. $\alpha_{11},\alpha_{2j_2},\ldots,\alpha_{rj_r}$ 这 r 个非零元素的左方、上方和下方全是 0, 而第 r 行以下各行 (即第 $r+1$ 至 m 行）的前 n 列元素全是 0.

矩阵消元至此结束. 写出最后这个矩阵 (1.3.6) 所代表的方程组

$$\begin{cases}
x_1+\alpha_{12}x_2+\cdots & +\alpha_{1,j_2+1}x_{j_2+1}+\cdots & +\alpha_{1,j_r+1}x_{j_r+1}+\cdots & =\beta_1\\
\qquad x_{j_2} & +\alpha_{2,j_2+1}x_{j_2+1}+\cdots & +\alpha_{2,j_r+1}x_{j_r+1}+\cdots & =\beta_2\\
 & \cdots\cdots & \cdots\cdots & \\
 & & x_{j_r}+\alpha_{r,j_r+1}x_{j_r+1}+\cdots & =\beta_r\\
 & & 0 & =\beta_{r+1}\\
 & & & \vdots\\
 & & 0 & =\beta_m
\end{cases}\tag{1.3.7}$$

我们称上述形式的方程组具有最简形式. 矩阵消元法的过程就是将原方程组化为最简形式的过程. 方程组化为最简形式之后，可以立即判断它是否有解，当它有解时可以立即得出它的解来.

情况 1 矩阵 (1.3.6) 的最后 $m-r$ 行不全为零， $\beta_i\neq 0$ 对某个 $r+1\leq i\leq m$ 成立。此时第 i 个方程 $0=\beta_i$ 的等号不可能成立，方程组无解。

情况 2 矩阵 (1.3.6) 的最后 $m-r$ 行全为零，这些行所代表的 $m-r$ 个方程 $0=0$ 全都是恒等式，可以从最后的方程组中删去而不影响方程组的解. 最后的方程组只剩下 r 个方程. r 个未知数 $x_1,x_{j_2},\ldots,x_{j_r}$ 分别只在 r 个方程中的一个中出现而且系数为 1, 在其余方程中都不出现. 我们称这 r 个未知数为非独立未知数，其余 $n-r$ 个未知数都称为独立未知数. 当然，$r\leq n$. 而且当 $n=r$ 时不存在独立未知数，所有的未知数都是非独立的. 以下再分 $r=n$ 和 $r<n$ 两种不同情况讨论.

情况 2.1 $r=n$. 此时 n 个未知数全都是非独立未知数， $j_k=k$ 对 $1\leq k\leq n$ 成立，矩

阵 (1.3.6) 具有形式

$$\begin{pmatrix}1 & 0 & 0 & \cdots & 0 & \beta_1\\ 0 & 1 & 0 & \cdots & 0 & \beta_2\\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots\\ 0 & \cdots & \cdots & 0 & 1 & \beta_n\end{pmatrix}.$$

对应的方程组为

$$\begin{cases} x_1 = \beta_1 \\ x_2 = \beta_2 \\ \quad\vdots \\ x_n = \beta_n \end{cases},$$

显然具有唯一解 $(\beta_1, \beta_2, \ldots, \beta_n)$.

情况 2.2 $r < n$, 除 $x_1, x_{j_2}, \ldots, x_{j_r}$ 等 r 个非独立未知数外，剩下还有 $n-r$ 个独立未知数，设它们为 $x_{j_{r+1}}, \ldots, x_{j_n}$, 其中 $j_{r+1}, \ldots, j_n$ 是从前 n 个正整数 $1, 2, \ldots, n$ 中去掉 $1, j_2, \ldots, j_r$ 之后剩下的 $n-r$ 个整数. 对每个方程移项，只保留它所独有的那个非独立未知数在左边，而将其余各项 (独立未知数所在的项) 全部移到右边，则方程组化为如下形状

$$\begin{aligned} x_1 &= \beta_1 - \alpha_{1j_{r+1}} x_{j_{r+1}} - \cdots - \alpha_{1j_n} x_{j_n} \\ x_{j_2} &= \beta_2 - \alpha_{2j_{r+1}} x_{j_{r+1}} - \cdots - \alpha_{2j_n} x_{j_n} \\ &\cdots\cdots \quad \cdots\cdots \\ x_{j_r} &= \beta_r - \alpha_{rj_{r+1}} x_{j_{r+1}} - \cdots - \alpha_{rj_n} x_{j_n}. \end{aligned}$$

这 r 个方程将非独立未知数表示为独立未知数的一次函数. 将各个独立未知数 $x_{j_{r+1}}, \ldots, x_{j_n}$ 在允许值范围 — 数域 F 内任意取值 $t_1, \ldots, t_{n-r}$, 由上述 r 个表达式就可以算出非独立未知数的值，得到一个解，解的每个分量都是 $n-r$ 个独立参数 $t_1, \ldots, t_{n-r}$ 的一次多项式. 这就是方程组的通解. 当所有的参数 $t_1, \ldots, t_{n-r}$ 分别独立取遍 F 时，就得到所有的解.

由上面的讨论可以知道：

1. 矩阵消元法可以求出任何一个线性方程组的解。

2. 将原方程组化为最简形式之后，可以立即判断它是否有解、有解时是有唯一解还是有无穷多组解，在有解时可以立即写出它的通解。

方程组化成最简形式之后，等号左边的非独立未知数的个数 r 对于判断方程组是否有解以及解集的大小有重要的作用.

将最简形式的方程组中的恒等式 $0=0$ 删去，不影响方程组的解. 设剩下的方程个数为 $\tilde{r}$.

1. 如果 $r < \tilde{r}$, 则方程组无解.

2. 如果 $r = \tilde{r}$, 则方程组有解. 此时 r 就是具有最简形式的方程组中除去恒等式 $0=0$ 之后的方程个数.

当 $r=n$ 时方程组有唯一解.

当 $r<n$ 时方程组有无穷多解，并且通解中有 $n-r$ 个独立取值的自由参数.

容易想到，当方程组有无穷多组解时，自由参数 $n-r$ 的个数反映了解集的大小. 比如，当 $n=3$ 并且方程组的系数都是实数时，在空间中建立直角坐标系，方程组的每一个解 (x,y,z) 表示空间中以 (x,y,z) 为坐标的一个点，所有的解代表的点的集合组成方程组的图象. 当 $n=r$ 时图象是一个点，当 $n-r=1$ 时图象是一条直线，当 $n-r=2$ 时图象是一个平面，我们显然认为平面上全体点的集合比直线上全体点的集合更大，并且认为平面是 2 维的，直线是 1 维的. 这个观点能不能推广到 n 为任意值的情况？对任意的 n, 能不能直接将 $n-r$ 称为解集的"维数", 并且认为维数越大的解集越大？

这种观点有道理，但是还有一些问题需要进一步明确：

首先，对于每个方程组，化为最简形式之后的 r 是否唯一？由于解方程组的化简过程并不唯一，会不会发生这样的情况：不同的化简过程得到的 r 值不同？如果有可能不同，用哪一个 $n-r$ 来作为解集合的"维数"?

另一个问题是：什么是"维数"? 如果将解集合中独立参数的个数作为维数，则同一个解集可能有不同的表示方法，从而有不同个数的自由参数，以哪一种表示法中独立参数的个数作为维数呢？

例如，假如 3 元一次方程组的通解是 $(x,y,z)=t_1(1,1,1)+t_2(2,1,5)+t_3(1,-3,13)$, 其中有 3 个独立取值的自由参数，是否意味着解集合的图象是整个 3 维空间？但实际上，由于

$$4(2,1,5)-7(1,1,1)=(1,-3,13),$$

3 个向量 $(1,1,1),(2,1,5),(1,-3,13)$ 共面,这个通解所代表的图象其实与 $(x,y,z)=t_1(1,1,1)+t_2(2,1,5)$ 代表的图象相同，是一个平面，应当是 2 维的.

由此可见，"维数"的概念应当有更确切的定义.

这些问题将在下一章中得到解决.

还有一个问题：能不能不经过方程组的变形，直接根据原方程组判断它是否有解，以及在有解是判断它的解是唯一还是有无穷多？

对于一般的方程组，难以直接判断. 但对于齐次线性方程组，也就是所有方程的常数项都是 0 的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \qquad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0. \end{cases} \tag{1.3.8}$$

我们知道它至少有平凡解 $(x_1,\ldots,x_n)=(0,\ldots,0)$. 我们进一步关心的问题是：除了平凡解以外，它是否还有非平凡解 $(x_1,\ldots,x_n)\neq(0,\ldots,0)$?

假如将齐次线性方程组 (1.3.8) 化成了最简形式 (1.3.7), 在最简形式的方程组中去掉形如 $0=0$ 的恒等式，设剩下的方程个数为 r. 则当 $r<n$ 时，有 $n-r$ 个可以独立取值的自由未知数，让这些未知数取非零值，就得到 (1.3.8) 的非零解，并且在此时方程组有无穷多组解.

不经过解方程组的过程，不知道 r 的具体值，但至少知道 $r \leq m$, m 是方程组 (1.3.8) 中方程的个数．如果 $m<n$, 则不需解方程组就可以知道 $r<n$, 从而知道此时方程组 (1.3.8) 一定有非零解．由此得到：

定理 1.3.1 如果齐次线性方程组的未知数个数大于方程个数，则齐次方程组有非零解，从而有无穷多组解. □

这个定理虽然只是对未知数个数大于方程个数的齐次线性方程组得出的结论，但在以后的学习中将看到，这一结论非常有用.

习题 1.3

1. a, b 取什么值时，下面的方程组有解，并求出其解.

$$\begin{cases} 3x_1+2x_2+ax_3+x_4-3x_5=4 \\ 5x_1+4x_2+3x_3+3x_4-x_5=3 \\ x_1+x_2+3x_3+2x_4+x_5=1 \\ x_2+2x_3+2x_4+6x_5=-3 \\ x_3+bx_4+x_5=1 \end{cases}.$$

2. 讨论当 λ 取什么值时下面的方程组有解：

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

当方程组有解时求出解来，并讨论 λ 取什么值时方程组有唯一解，什么时候有无穷多组解.

3. (1) 求下面的非齐次线性方程组的通解：

$$\begin{cases} x_1+x_2+x_3+x_4+x_5=1, \\ x_1+2x_2+3x_3+4x_4+5x_5=6, \\ x_1-x_3-2x_4-3x_5=-4. \end{cases} \tag{I}$$

(2) 将方程组 (I) 的常数项全部换成 0 得到齐次线性方程组 (II), 求方程组 (II) 的通解. 并将通解写成其中几个特解的线性组合的形式.

(3) 方程组 (I) 的通解能否写成几个特解的线性组合？

(4) 观察方程组 (I) 与 (II) 的通解之间的关系，你发现什么规律？试证明你的结论.

4. 不解方程组，判断下面的方程组是否有非零解：

$$(1)\ \begin{cases} x+y+z=0 \\ 2x+y+5z=0 \end{cases};\qquad (2)\ \begin{cases} x+y+z=0 \\ 2x+y+5z=0 \\ 3x+2y+6z=0 \end{cases}.$$

(提示：注意第 (2) 题的第 3 个方程组是前两个方程的和．)