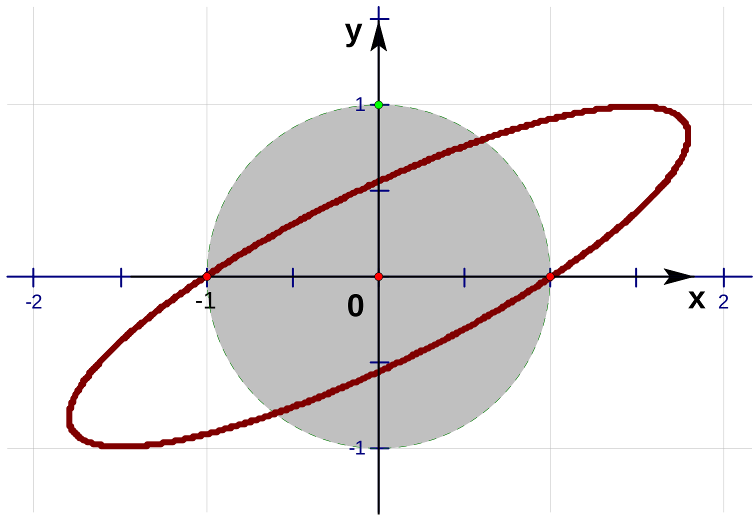
线性代数的几何意义

--------图解线性代数

任广千　胡翠芳 编著



2010.04.20

几何意义名言录

没有任何东西比几何图形更容易印入脑际了，因此用这种方式来表达事物是非常有意义的。 -------笛卡尔

算术符号是文字化的图形，而几何图形则是图像化的公式；没有一个数学家能缺少这些图像化的公式。 --------希尔伯特

“如果代数与几何各自分开发展，那它的进步十分缓慢，而且应用范围也很有限，但若两者互相结合而共同发展，则就会相互加强，并以快速的步伐向着完善化的方向猛进。”

--------拉格朗日

不会几何学就不会正确的思考，而不会正确思考的人不过是行尸走肉。 --------柏拉图

无论是从事数学教学或研究，我是喜欢直观的。学习一条数学定理及其证明，只有当我能把定理的直观含义和证明的直观思路弄明白了，我才认为真正懂了。--------中国当代数学家徐利治

前言

为什么要给出线性代数的几何意义

作为一名工作十多年的电子工程师，作者在想提高自己的专业水平时，深感数学能力的重要。随便打开一篇专著或论文，满纸的微分方程、矩阵扑面而来。竭力迎头而上，每每被打得灰头土脸、晕头转向。我天生就不是搞数学的？我的智力有问题吗？

太失望了，太伤自尊了。转头看看周围的同行，莫不雷同。大多的工程师们靠经验来工作，经验靠时间或试验来积累。数学应用的层次最多就是高中水平。也有硕士博士级的牛人，但也少见把数学工具在工作中应用的得心应手、手到擒来的。

数学工具在科技实践中缺失的严重，导致我们的科技创新能力的严重缺失。普遍现象，绝对的。

返回来想一想，我的智力应该没问题，重点大学都毕业了，能有多严重的问题？所有的工程师们、大学毕业生们的智力也没问题。问题是大家没把数学学好，没有真正掌握它。

（严重声明：数学绝顶高手和天才们不在我说的范围之内，对我等来说，它们是极少数的一小撮的火星人，对它们只能顶礼膜拜，不敢评论。----拜完之后有点小嘀咕：为何钱学森还讲中国没有大师呢？难道数学总得一百分的天才不算大师？）。

为啥没有在四年的大学阶段学好《线代》呢？要知道，学生是通过高考百里挑一录取的，智力应是足够正常的。思来想去，得到几个原因：教材编的大多不好，老师教的大多乏味，学生大多有些偷懒，因为他们大多不知道这些内容有啥用，概念为啥这么叫，定理为啥那样推，老师为啥像刘谦的魔术一样七推八导就证毕了----郁闷多了导致了无语的偷懒。

太多的为啥了。既然错不在学生那就是老师的问题了？其实老师也有委屈：教学大纲要求在几十个学时学会如此多的内容，不填鸭行吗？在如此短的时间内讲完就不错了，哪里还有时间给你释疑解惑。-----韩愈定义的传道授业解惑的师道中的解惑被迫取消了，自己悟道吧。

嘿，错也不全在老师那里。错在哪里？体制的问题一时半会也解决不了，不谈体制的事了。找来找去，只有一个大家都可以责备而且没有人抗议的地方，就是教材不够好。

到大学图书馆（本人主要去深大图书馆）看看，哇塞，一行行、一列列的教材琳琅满目、浩如烟海。名字叫《线性代数》的教材足有一千多册。

打开一本看看，跟十八年前的教材内容一样，疑问还是没得到解决；再打开一本看看，内容还是那个内容，疑问还是那个疑问…。

当浏览到第五百本的时候，皇天不负有心人！终于看到了我那个问题的答案了。长出一口气我又陷入了郁闷之中。要知道，我至少有十打问题要解决呀，上帝。

呵，西方的上帝来拯救我来了，当我浏览到第八百本的时候，一本老外编的教材一下子吸引到我那累的发红的心灵之窗。

我的天，我一阵眩晕，问题至少能解决五打。我抱着一本老厚老厚的海外引进教材看呀想呀，从此以后我专看老外的书，嘿嘿，只有一打的问题了。我想《线性代数》这门学科问题应该不大了，要知道，老外的教材都是引入了当代科技的典型应用案例的，代表了本学科最新的国际潮流的。

大学图书馆的读者很多，朝气蓬勃的现代感大学生在图书馆里做作业。我很羡慕他们这一代：在开放的图书馆里，学生们可以随意的浏览、挑选适合自己的纸资或电子读物。要知道，当年我就读的大学图书馆是闭架的，每每借书要查半天小卡片，查完填好借书单交给工作人员大多得到两个结果：要么书被借完了要么借的书不合适。而且还没有这么多的引进教材参考参考，自学的效率大打折扣。

扯来扯去，千言万语汇成一句话：什么样的《线性代数》学习资料较好，较适合中国学生？我想，本子的物理尺寸要越薄越好，内容要越通俗易懂越好。

书本越薄大家学习的信心越强：小样，这么点厚度还搞不定你，看，信心先有了。

如果只是容量精简了还不行，考试的时候受打击，工作中更受打击。如当年我学的《线性代数》课本是同济编的，内容是精简到家，千锤百炼，没一句废话，超薄。死记硬背，看似搞定了，实际是囫囵吞枣。

如何通俗易懂还不能多说？我一直认为，加上几何意义或者物理意义啥的，一步到位搞定。

这就是本《线性代数的几何意义》的由来。也是这个本子的目标。

目标有了，具体如何编写呢？模仿一下科学大德牛顿的口气：

从线性代数书籍的浩瀚海洋的沙滩上（还没有更高的能力去远洋、去深海处），用一双自己的眼睛，寻找到了一个个闪闪的小珍珠，一片片如玉的小彩贝，然后细细的打磨和擦拭，拂去沙尘，使它们重放光彩，用一根几何意义的锦丝，穿就了这本《线性代数几何意义》的项链，献给热爱思考、痴迷于创造的人们。

呵呵，自不量力，终极目标而已，但意思还是有了。

重要的几何直观意义

在学习中，一旦碰到较抽象难懂的概念或定理，如何搞定？几个办法：一个是看推导过程，推导可以加强你相信它的信心并连通你原有的知识体系。如果推导把你弄昏了，只好弄懂它的几何意义或物理意义啦。

几何意义或者讲几何解释会和人们看到的平面和空间中物体几何外观联系起来，几何上说的通，物理上也就说得通，几何意义和物理意义本质上是一回事（如果你不信物理和几何是一回事，就想想爱因斯坦，想想相对论），因此大家就相信了，就会和大家大脑中的经验和原有知识网络连通，一下子就“懂了”，满心欢喜的，原来是这么一回事。

真理总是简单的和直观的，一位先贤说，不管多么复杂高深的数学理论，总有其直观的背景，不管多么繁难深奥的定理，其证明总有一个简单而直观的中心思想。几何图形能以其生动的直观形象给人留下深刻的印象。可以这样说，在数学中再没有别的什么东西，能比几何图形更容易进入人们的脑海了。

从宏观上看，一种数学理论(包括它的主要概念和方法)往往都有其直观的背景，它们或者是从对某些特殊的事例的观察分析中得到的，或者是直接从几何图形中看出的，或者是从已有的结果类比联想引来的，从几何直观上分析问题的能力，首先是指对于一种数学理论能“洞察其直观背景”。对于它是如何被发现的或如何形成的作出合理的解释或猜测。

一句话，皇皇巨著的理论特别是抽象的数学理论的核心常常可以从几何意义的角度得到解释。

从微观上看，关于某一个具体定理的证明，国外的数学教育家波利亚曾经说：“一个长的证明常常取决于一个中心思想，而这个思想本身却是直观的和简单的”。因此，从几何直观上分析问题的能力，也包括找出证明中的那个关键的简单而直观的思想，也就是象希尔伯特所要求的，能透过概念的严格定义和实际证明中的推演细节，“描绘出证明方法的几何轮廓”。

大师庞加莱和阿达玛关于数学领域的发明创造的观点也认为，数学创造发明的关键在于选择数学观念间的“最佳组合”，从而形成数学上有用的新思想和新概念，而这种选择的基础是“美的直觉”。在这种美的直觉中，也就是在追求某种对称性、和谐性、统一性、简洁性和奇异性当中，以及在某种联想、猜想、假设及非逻辑思维中，几何直观具有头等重要的意义。

事实上，很多数学家都是先利用几何直观猜测到某些结果，然后才补出逻辑上的证明的。这正如我国著名拓扑学家张素诚先生所说的，对数学中的许多问题来说，“灵感”往往来自几何，表达的简洁靠代数，计算的精确靠分析。

嘿嘿，看看上面的数学上的历史牛人的观点，几何形象直观的意义何等重要。其实，大家都知道几何意义的重要，我们在小学和中学的学习阶段，老师常常也讲一些抽象概念所对应的几何意义，为何到了大学我们的大脑就一下子高度抽象起来了？把形象仍得远远的，象瘟疫一样躲着他？目的是训练抽象思维？最终实际结果呢？不可否认，大学毕业后大家确实是抽象了，抽象得只会夸夸其谈讲理论不会干具体活了。 既然你具体的活计不会干那干脆就专搞抽象的理论去嘛，结果也搞不了，为啥？只会做做过的抽象的数学题不会发明创造，没学会真正的抽象，真是越抽象越糊涂。

我觉得，抽象和形象是相辅相成，缺一不可的。由形象而抽象，再由抽象到形象，人的知识结构螺旋架才能旋转而上，达到越来越高的知识峰巅。

如何使用这本书

拼命阐述几何直观在数学学习中的重要意义，但这并不意味着可以否定逻辑推理论证的重要作用。实际上，单纯地依据直观而导致错误的数学例子真是数不胜数。概念或定理的几何直观解释，往往并不等同于原来的概念或定理。运用几何直观可以帮助我们猜想，但猜想并不能代替证明，只有经过一步步严格的逻辑论证以后，才算给出了证明。

形象或直观和抽象本来是一切科学的两面。只是近年来过分强调了抽象思维能力的训练而忽视了几何意义的解释。反过来，我们不能只强调了几何意义而丢掉了计算和推导。因此建议读者：

- 初学者从几何意义入手，轻松而迅速理解和把握线性代数的基本概念和定理几何本质， 建立对线性代数的感性认识，具备了理解复杂及抽象数学的能力。
- 然后，在回到现在的抽象的线性代数的教材，短时间内构筑个人的线性代数的知识体系的“向量空间”，通过适量的习题训练，巩固解决具体问题的动手能力。此时，**具体与抽象一体，理想与现实齐飞**。您，已经成为线性代数的高手和大牛。

注：本文中，几何意义和几何解释的文字意思没有根本区别，一般对于数学概念的对应的几何图形而言称为几何意义，而对运算、变换的过程可对应几何图形的变化过程称为几何解释。

第一章 什么是线性代数？

这一章的内容主要是想对线性代数的大的概念如线性函数、映射和线性变换以及线性代数的发展简史和应用作一简要介绍，本章的目的是让读者知道我们所学的线性代数的实质是什么，到底有什么用。

线性代数是代数学乃至整个数学的一个忒重要的学科，顾名思义，它是研究**线性**问题的**代数**理论。那么什么是代数呢？

代数英文是 Algebra，源于阿拉伯语，其本意是“结合在一起”的意思。也就是说代数的功能是把许多看似不相关的事物“结合在一起”，也就是进行抽象。抽象的目的不是为了显示某些人智商高，而是为了解决问题的方便，为了提高效率，把许多看似不相关的问题化归为一类问题。比如线性代数中的一个重要的抽象概念是线性空间（对所谓的要满足“加法”和“数乘”等八条公理的元素的集合），而其元素被称为向量。也就是说，只要某个集合里的元素满足那么几条公理，元素之间的变化满足这些规律，我们就可以对这个集合（现在可以改名为线性空间了）进行一系列线性化处理和分析，这个陌生的集合的性质和结构特点我们一下子就全知道了，因为宇宙间的所有的线性空间类的集合的性质都一样，地球人都知道（如果地球人都学了线性代数的话）。多么深刻而美妙的结论！这就是代数的一个抽象特性。

注： “代数”这一个词在我国出现较晚，在清代时才传入中国，当时被人们译成“阿尔热巴拉”，直到 1859 年，清代著名的数学家、翻译家李善兰才将它翻译成为“代数学”，一直沿用至今。

既然这个具有特性的集合叫线性空间，顾名思义，当然具有直观的几何意义。线性来源于直线的几何概念，空间来源于二维平面或三维的立体的几何概念。让我们真正幸运的是，所有的五花八门的线性空间（这些线性空间大多隐藏在我们的物理世界中而难以发现，比如隐藏在电子电路世界里面的由电阻、电感或电容组成的电路网络，比如隐藏在高等数学里面的满足微积分运算的数的集合等等）都可以和实数域上的线性空间 R^n 同构。什么意思？就是所有类型的线性空间都和直线、平面、三维立体以及高维正交空间的变换性质一样，所有类型的线性空间里的元素都可以和 R^n 空间的点（向量）相互对应。一句话，R^n 空间就是所有的线性空间的几何意义或几何解释。

实际上，本书对线性代数进行几何意义上的解释正是从向量和二维和三维的实数线性空间 R^2 和 R^3 的角度全面解读的。

那么线性问题又是什么样的问题呢？

在大家的科技实践中，从实际中来的数学问题无非分为两类：一类线性问题，一类非线性问题。线性问题是研究最久、理论最完善的；而非线性问题则可以在一定基础上转化为线性问题求解。因此遇到一个具体的问题，首先判断是线性还是上非线性的；其次若是线性问题如何处理，若是非线性问题如何转化为线性问题。

下面我们通过介绍几个主要的概念来逐渐的把握线性这个核心意思。

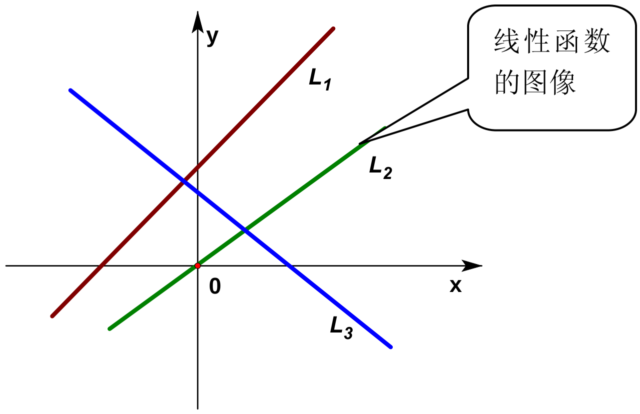
1.1 “线性”的意义

线性代数里面的线性主要的意思就是线性空间里的线性变换。线性变换或线性映射是把中学的线性函数概念进行了重新定义，强调了函数的变量之间的变换的意义。

线性函数的概念

线性函数的概念在初等数学和高等数学中含义不尽相同（高等数学常常把初等数学的关键概念进行推广或进一步抽象化，初等数学的概念就变成了高等数学概念的一个特例）。

在中学的初等数学里，我们知道，函数 $f(x)=kx+b$ （k,b 是不变量），称为一元线性函数，因为在平面直角坐标系中这个函数的图形就是一条直线，就是变量（包括自变量和因变量）之间的关系描述为一条直线，所以把这种函数形象地称为“线性”函数（如图 1.1）；如果 $b=0$，这个函数的外观就变成 $f(x)=kx$ 的形式了，这是一条过原点的直线，如图中直线 L_2。显然，过原点的直线是最简单的线性函数。



在大学的代数里面，为了线性函数的进一步推广（如推广至双线性函数、多线性函数、线性空间、线性泛函...）的远大未来，我们忍痛割“尾”，把一元线性函数 $f(x)=kx+b$ 的 b 割舍掉，成了 $f(x)=kx$ 的形式。

呵呵，简单点说，只有过原点的最简单的直线 $f(x)=kx$ 才被称为一元线性函数。

为什么？

只因为不过原点的直线不满足我们对线性函数的比例性的要求（这又是为什么，本书的后续章节会告诉你，如果你有兴趣继续读下去的话）。

线性函数表现为直线，这只是几何意义。那么所谓“线性”的代数意义是什么呢？实际上，最基本的意义只有两条：可加性和比例性。

1）可加性：即如果函数 $f(x)$ 是线性的，那么有：

$$f(x_1+x_2)=f(x_1)+f(x_2)$$

一句话：和的函数等于函数的和。

2、比例性：也叫做齐次性、数乘性或均匀性，即如果函数 $f(x)$ 是线性的，那么有

$$f(kx)=kf(x) \quad \text{其中 } k \text{ 是常数。}$$

一句话：比例的函数等于函数的比例；或者说自变量缩放，函数也同等比例地缩放。

注：对于函数 $f(x)=ax+b$ 而言不满足此比例性，$f(kx)=akx+b$，$kf(x)=akx+kb$，因此 $f(kx)\neq kf(x)$。严格的讲，$f(x)=ax+b$ 不能再叫线性函数了。

可加性与比例性组合在一块就是“线性”的全部意义了，即有

$$f(k_1x_1+k_2x_2)=k_1f(x_1)+k_2f(x_2)，\text{其中 } k_1,k_2 \text{ 是常数。}$$

一句话：线性组合的函数，等于函数的线性组合。

可加性和比例性的物理意义是什么呢?

线性函数的可加性表明函数所描述的事物具有累加性，所有起因的累加所导致的结果完全等于每个起因独自所引起的结果的累加。可加性看起来简单，似乎没有内涵，其实它界定了所描述的事物是线性还是非线性的。举两个例子：

一个是晶体管放大器，晶体管的电流放大特性分三个区间，截止区、线性区和饱和区。在线性区里面，基极电流是 0.5mA，则集电极电流就是 50mA（设晶体管电流放大倍数是 100）；如果基极电流是 1mA，则集电极电流就是 100mA。进一步地，如果输入的基极电流是 1.5mA(0.5mA 和 1mA 相加)，则集电极电流就是 50mA 和 100mA 之和 150mA。线性区里面的电流放大过程满足可加性。但在其它两个区里面就不满足这个可加性，如饱和区，当基极电流是 50mA 时，设集电极电流是 1000mA（不是 5000mA，因为不在线性区。又设晶体管饱和电流是 1A~1.1A）；当基极电流为 60mA，设集电极的合理电流是 1050mA，好了，输入两个电流的和 110mA，那么集电极电流就只有 1100mA 而不是 1000mA 和 1050mA 之和，因为达到了集体管最大的饱和电流了，电流增加不上去了。饱和区失去了放大电流的累加性。

一个是人力资源的故事例子，呵呵，属于小学一年级级别的脑筋急转弯：王五在旧上海滩的码头上扛货物麻袋，一天能扛 200 袋。好梦不长，几个月后，陈阿真也来扛麻袋了，谁？就是干了虹桥道

馆那件大快人心的事后为了躲避追捕也跑到码头上混迹的陈真，大侠韬光养晦化名陈阿真，一天可以扛 300 袋麻袋。好了，问个问题：王五和阿真两个人一天能抗多少个麻袋？

“500 袋！”天真率性同学们思维很线性，“200 袋加 300 袋就是 500 袋嘛”。

不对（对了就不是脑筋急转弯了），据有好事者统计，俩人第一天一共扛了 600 袋麻袋，第二天一共扛了 400 袋麻袋。怎么回事？原来人力的事情不是线性问题而是非线性的问题：

第一天，　两人摸不透对方的脾气和底牌，为了保住岗位，各施功夫互相竞赛，王五扛了 250 袋，阿真扛了 350 袋，合计 600 袋；

第二天，　两人一想，靠，这样下去还不累死掉，一山不容二虎，给对方捣蛋弄走对方。于是两人便扛麻袋边向对方施展拳脚，内耗了，一天共扛了 400 袋。

总结一下，线性的可加性既是没有互相激励的累加，也是没有互相内耗的累加。一加一就是二，既不大于二也不小于二（对不起哈，一不小心证明了哥德巴赫猜想）。

比例性是啥物理含义呢？比例性又名齐次性说明没有初始值，比如电路，没有输入信号时输出也为零，有几倍的输入量刚好就有几倍的输出量，增量是倍数关系，存量也是倍数关系。

实际上，高等的线性概念正是从最简单的比例函数进行推广的，在大学所学习的线性代数里的线性函数概念被扩展成一个多元线性方程组所表示的一个对应关系。如方程组

$$\begin{cases} y_1 = k_{11}x_1 + ...k_{1n}x_n \\ y_2 = k_{21}x_1 + ...k_{2n}x_n \\ .... \\ y_m = k_{m1}x_1 + ...k_{mn}x_n \end{cases} \quad （k_{11}...k_{mn} 是不变量），$$

是由 m 个 n 元线性函数组成的，而且这 m 个线性函数还是齐次函数，他们全部过原点。看看，高等概念的线性函数是由初等的最简单的过原点的直线扩展来的，如果 $m = n$，那么这个方程组所确定的几何图形也是一条“直线”！而且，这个直线也是过原点的。（$m \neq n$ 的图形是平面或超平面的，平面是多线性的）

注：线性齐次函数形如 $y = k_1x_1 + k_2x_2 + ...k_nx_n$，这个正比例函数的式子中每项里的变量出现的次数都是一次的（没有常数项），整齐划一，故此称为“齐次”的，全称为 **n 元线性齐次函数**。

把 n 元齐次线性方程组称为线性函数有点信心不足，看起来方程组和单个方程差别挺大的。其实，我们也可以把它们写得形式一致：重新定义变量就可以把它改写成初等数学中的线性函数的形式 $f(x) = kx$ 了。这个重新定义的变量就是向量，扩展如下：

1) 初等线性函数的自变量由一个数 x 扩展定义为一个竖排的数组 $\begin{pmatrix} x_1 \\ x_2 \\ .... \\ x_m \end{pmatrix}$，因变量一个数 y 也扩展

定义为一个竖排的数组$\begin{pmatrix} y_1 \\ y_2 \\ .... \\ y_m \end{pmatrix}$，这些 n 元数组和 m 元数组称之为列向量。

2) 初等线性函数的比例系数 k 扩展为由所有的 k_{ij} 构成一个的数的方阵，称之为系数矩阵如下：

$$\begin{bmatrix} k_{11}, k_{12}, ...k_{1n} \\ k_{21}, k_{22}, ...k_{2n} \\ .... \\ k_{m1}, k_{m2}, ...k_{mn} \end{bmatrix}$$

3) 然后我们定义了一种系数矩阵与向量相乘的运算法则（在“矩阵的几何意义”一章中介绍），使我们可以把上述的线性方程组改写为如下形式：

$$\begin{pmatrix} y_1 \\ y_2 \\ .... \\ y_m \end{pmatrix} = \begin{bmatrix} k_{11}, k_{12}, ...k_{1n} \\ k_{21}, k_{22}, ...k_{2n} \\ .... \\ k_{m1}, k_{m2}, ...k_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ .... \\ x_m \end{pmatrix}。$$

4) 上式的形式为进一步简写为：

$$\mathbf{y} = f(\mathbf{x}) = \mathbf{Kx}$$

这里：$\mathbf{y} = f(\mathbf{x}) = \begin{pmatrix} y_1 \\ y_2 \\ .... \\ y_m \end{pmatrix}$，$\mathbf{K} = \begin{bmatrix} k_{11}, k_{12}, ...k_{1n} \\ k_{21}, k_{22}, ...k_{2n} \\ .... \\ k_{m1}, k_{m2}, ...k_{mn} \end{bmatrix}$，$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ .... \\ x_m \end{pmatrix}$。

到了这里，我们终于看到，初等线性函数和高等线性函数的概念终于得到了形式上的统一。

多元线性函数的几何意义

在前面的线性函数的推广中，从一元线性函数一下子推广到了 n 元线性函数组，跨度有点大了。下面补一下中间过程的课，探讨一下从一元线性函数如何推广到 n 元线性函数的。

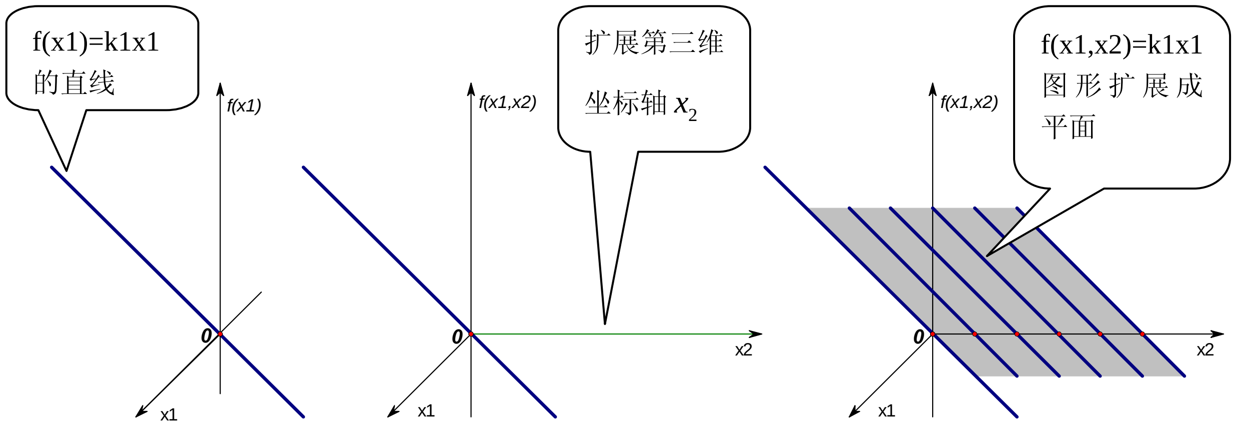
首先我们看看从一元线性函数 $f(x) = kx$ 拓展到二元的线性函数 $f(x_1, x_2) = k_1x_1 + k_2x_2$ 的几何解释。这个几何解释并不是唯一的，目的是让读者认识和理解线性的概念。

拓展的第一步，坐标系由二维扩展到三维：

$f(x) = kx$ 的直线图形是在二维笛卡尔坐标下给出的几何图形，把它放到三维笛卡尔坐标系下，

其函数表达式应写为 $f(x_1,x_2)=k_1x_1$ 或者 $f(x_1,x_2)=k_2x_2$。不失一般性，我们取 $f(x_1,x_2)=k_1x_1$ 为 $f(x)=kx$ 的扩维表达式。

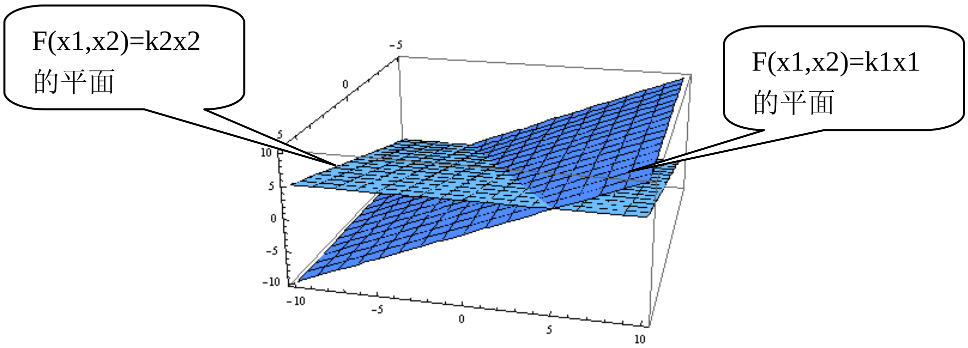
我们知道，$f(x_1,x_2)=k_1x_1$ 的图形是一个过原点的平面。扩维后由一根直线变成了一个平面。这是因为函数 $f(x_1,x_2)=k_1x_1$ 与新生长出来的坐标轴 x_2 没有关系，x_2 可以取任意值；换句话说，x_2 的任意值都在函数 $f(x_1,x_2)=k_1x_1$ 的图像上。进一步说，这是一个过 x_2 坐标轴的平面。

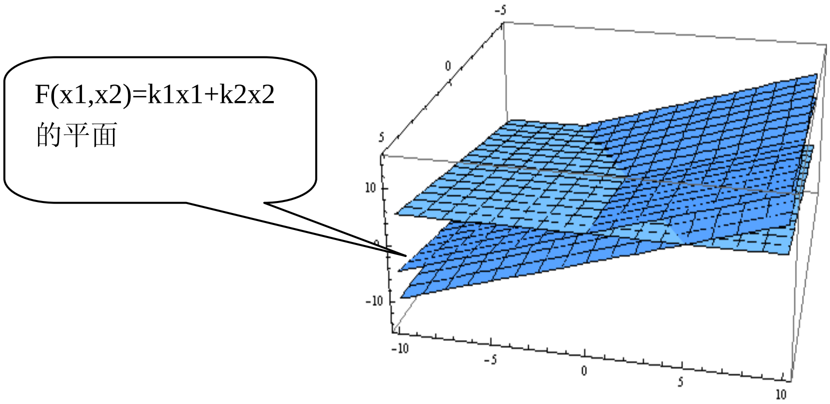


形象的扩展过程可以这样想象：二维平面坐标系里有一根直线图形，这时有 x_2 轴过原点以垂直于坐标系 $x_1 \sim f(x_1)$ 的平面向右方向（右手系）生长出来，然后原来的那条直线 $f(x_1)=k_1x_1$ 沿着坐标轴 x_2 方向向右滑动，无数个平行的直线被 x_2 轴象竹帘子一样串起来，平铺得到了 $f(x_1,x_2)=k_1x_1$ 的平面。这个平面是由无数的直线铺成的，因此，平面也是“线性”的。

拓展的第二步，两个平面加起来：

显然，要得到函数 $f(x_1,x_2)=k_1x_1+k_2x_2$ 的图形，只要把三维坐标系下的两个函数 $f(x_1,x_2)=k_1x_1$ 和 $f(x_1,x_2)=k_2x_2$ 所对应的图形加起来即可得到。一般情形下，两个平面相加仍然是一个平面。如下图示。





因此，线性函数 $f(x_1,x_2)=k_1x_1+k_2x_2$ 的几何图形是一个过原点的平面。

可以想象，由二元线性函数 $f(x_1,x_2)=k_1x_1+k_2x_2$ 继续扩展到三元及 n 元的线性函数 $f(x_1,x_2...,x_n)=k_1x_1+k_2x_2+...+k_nx_n$（坐标系由三维扩展到四维及其 n 维）后，其几何图形仍然是一个“平面”，是一个扩展意义上的平面，常被称为超平面。

如来佛陀是一位伟大的几何学家

---------理解多维空间

由三维扩展到四维的空间的确难以想象，我想给出个人的几点认识供读者参考:

1、 对于笛卡儿坐标系，二维坐标系的两个坐标轴互相正交并构成一个平面空间；三维坐标系的坐标轴互相正交且第三个坐标轴垂直于其余两个坐标轴平面，三个坐标轴构成一个立体空间；则四维坐标系中的四个坐标轴互相正交，第四轴必然与其余的三维立体空间垂直，四个坐标轴构成一个超多面体空间…

2、 四维空间的物理解释就是爱因斯坦的时空理论，三维物理空间之外增加了一个与之垂直的时间轴（垂直或正交的意思应理解为不相关，时间和空间在低于光速的尺度内就是没有关系的两个事物，这就是牛顿的世界）。你，作为一个有生命周期的高级动物实际上是个四维动物，因为你的肉体既占有了一个三维小空间同时又占有了另外一维的时间轴上的一段。

3、 N 维空间的出现实际上是人们在抽象他所观察到的宇宙事物时出现的概念。在一个银河系外的观察者看来，太阳系不过是视界平面上的一个点；当这名观察者快速逼近太阳系时，这个二维平面上的点逐渐变成了三维的太阳系空间，同样，此时的地球在观察者的二维视界平面上也是一个点而已；当观察者来到地球外的大气空间时，地球已是一个三维球体了，而一个人同样在观察者看来是一个点而已...，如果观察者继续体察入微，将会逐步的看到人的身体，身体上的细胞，染色体，原子，原子核等等，这是一个空间套着一个空间的 N 维空间，大的三维空间套着无数个小的三维空间空间。如来佛陀绝对是一位伟大的几何学家，因为 n 多年前他老人家就率先说过，一粒沙子就是一个大千世界。

4、 实际上，在以后的线性代数学习中，坐标轴的正交不是必须的，取消了正交的要求后，我们在平面上就可以画出来大于四维以上的空间来了，你就理解了由 n 个向量张成的 n 个空间的理论，进而想象高维空间的图像也就不是一个困难的事情了。

到此我们明白了多元线性函数的“线性”不能单纯的理解为空间中的一条直线了，根据上面的讨论，把线性函数几何图形想象成一个平面更有代表性。实际上，把 n 个 n 元线性函数组成一个满秩方程组才能表示为一条直线。

线性函数中含有的参数少，涉及的运算简单，仅为加法和乘法，便于运算，是变量数学中最简单的函数；但另一方面，许多复杂的函数都可以在一定范围和精确度下近似地用线性函数来表示，所以线性函数又是变量数学中最重要的函数。

1.2 线性映射或变换的几何意义

线性映射的几何意义

前面说，初等线性函数 $f(x)=kx$ 和高等线性函数 $f(\mathbf{x})=\mathbf{Kx}$ 的表达式一致，因此线性函数的概念形式上是统一的。这种统一在数学的实质意义上也是一致的，就是 函数的“线性”，实质上就是指变量之间的“线性关系”。

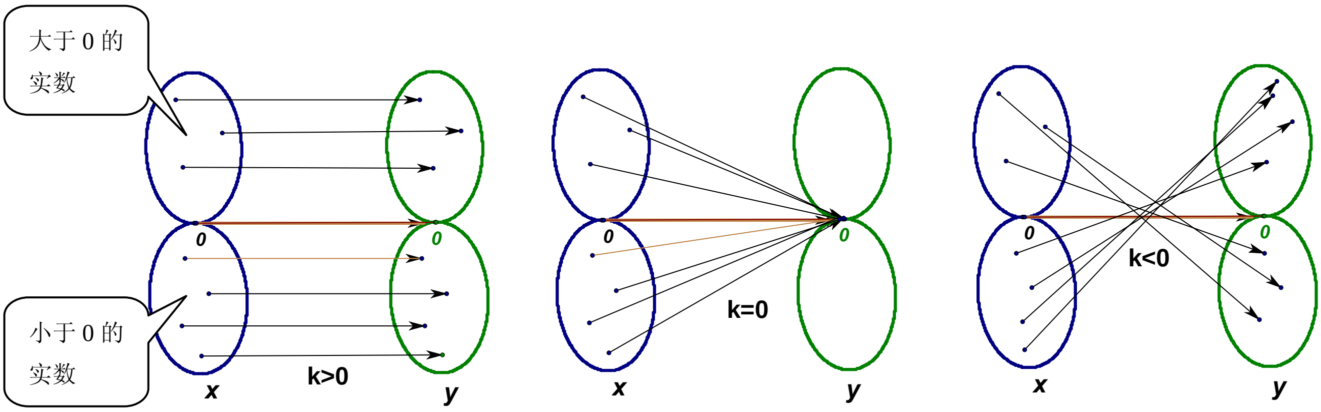
我们再来回味一下这句关于线性函数中心性质的话：线性组合的函数，等于函数的线性组合，详细说来就是，如果自变量从 x 变换为自变量的线性组合 $k_1x_1+k_2x_2$ 时，其函数也从 $f(x)$ 变换为函数的线性组合 $k_1f(x_1)+k_2f(x_2)$ 。因为函数的本意是因变量与自变量之间的对应关系，所以“线性”的本质就是因变量与自变量之间始终保持组合形式不变的一种对应关系，我们把这类特殊的对应关系称之为“线性关系”。因此我们所说的“线性”，实质上就是指变量之间的“线性关系”。

实际上，我们可以引入一种运动的思想，把函数看成一种变换，一种映射，一种从自变量的集合对应变换到因变量的集合的瞬间过程。

这正是线性代数的一个中心思想之一。

对于初等的线性函数 $f(x)=kx$ 而言，我们需要改变中学老师谆谆教导。中学老师说，线性函数的几何图形是所有满足关系式 $y=kx$ 的点 (x,y) 所累积起来的图形。这个静态的图形概念需要改造改造。要在这里加入变换或映射的动作（注意：是动作，一个瞬时的变化动作，只有开始和结果），并突出表达这种变换和投射的关系，我们把表达式 $f(x)=kx$ 改写成 $T:\mathbf{x}\to\mathbf{y},\mathbf{x}\mapsto k\mathbf{x}$， $T:\mathbf{x}\to\mathbf{y}$ 表示为一个从自变量数的集合 $\mathbf{x}$ 到因变量数的集合 $\mathbf{y}$ 的映射， $\mathbf{x}\mapsto k\mathbf{x}$ 表示两个集合里的自变量 x 到因变量 y 之间具体的对应变换关系。

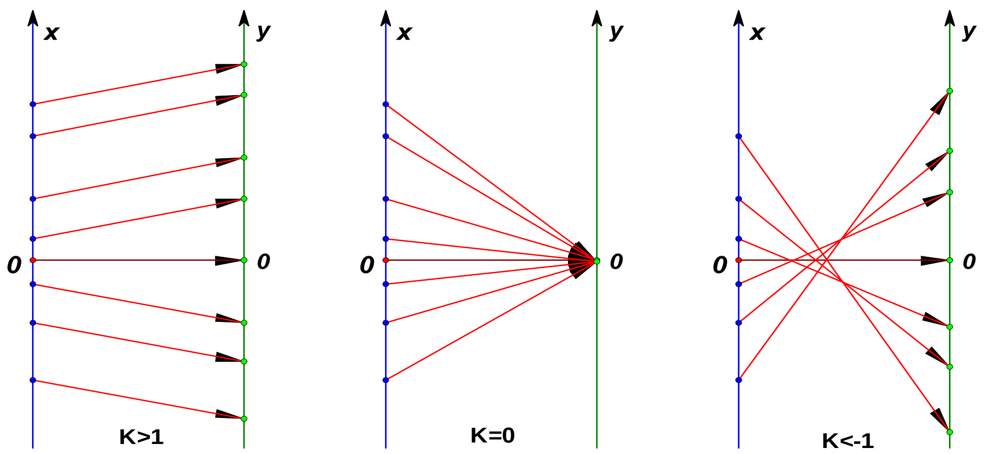
如果我们给出一个映射的集合示意图，则有如下图所示。



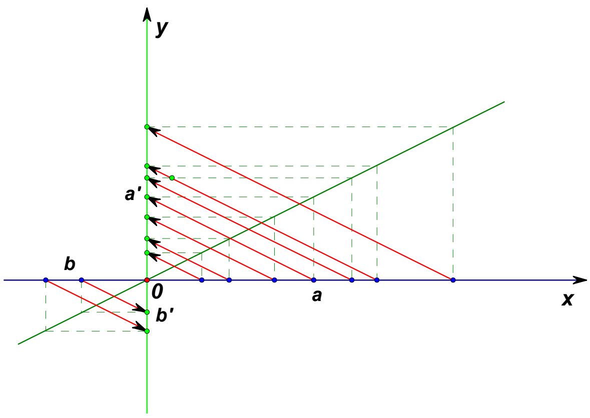
图中给出了一元线性齐次函数 $f(x)=kx$ 当 k 取不同的数时的映射对应关系。在三个图中，有一个共性就是元素 0 必然映射到元素 0。

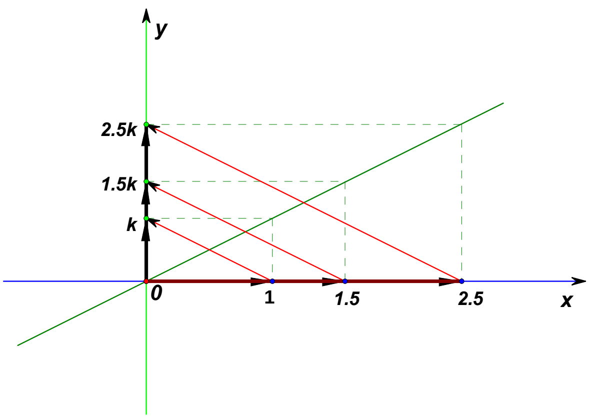
在集合上建立坐标系，用坐标系里的点表示集合里的元素，就可以把映射关系几何化了。

对于一元线性齐次函数 $f(x)=kx$，集合 **x** 和集合 **y** 都是实数。大家知道，一个实数域可以用一根坐标轴就可以表示了。因此集合 **x** 的坐标系就是一根轴，写为 x 轴；集合 **y** 的坐标系也是一根轴，写为 **y** 轴。这样，我们就可以用坐标轴上点之间的映射来替代上图集合的映射表示法。



如果把两个坐标轴的原点进行重合（因为 0 元素必然映射到 0 元素），如果再把两个坐标轴的夹角调整到 $\dfrac{\pi}{2}$ 角，就得到了笛卡尔平面坐标系了（线性代数的里面讲的线性空间坐标系的坐标轴可以是任意非零的夹角）。他 $\mathbf{x}\mapsto k\mathbf{x}$ 用带箭头的线段连接起来，则有下图所示（下图中只画出了 $k>0$ 的映射情况）。





如上图左，x 轴上的点 a 和 b 等等分别映射到 y 轴上的点 a' 和 b' 等等。

如果把点 a 、a'、b 和 b' 分别与原点 0 连起来，就会得到线段 $0a, 0b, 0a', 0b'$。于是线段 $0a$ 映射到线段 $0a'$，线段 $0b$ 映射到线段 $0b'$。到这里我们有了一个暂时的总结：线性映射就是把线段映射到线段。如果我们把线段改称为向量的话，这个总结就是：线性映射就是把向量映射成向量。线性映射把向量变成另外一个向量，或者说把“线”变成“线”，因此得名。

当然，这个线性映射也满足线性的可加性和比例性的性质：可加性就是 x 轴上的两向量的和映射得到的 y 轴向量等于两个 x 轴向量分别映射得到的 y 轴向量的和，比例性就是 x 轴向量的倍数映射到的 y 轴向量等于 x 轴向量映射的 y 轴向量的倍数（上图右给出了例子）。

用一般的数学表达式来描述线性映射的定义就是：

$$T(\alpha+\beta)=T\alpha+T\beta$$
$$T(k\alpha)=kT\alpha$$

其中，α, β 是向量。

注：

你看这里，T 本来表示一种线性映射的动作关系（或函数关系），但在上式中就像一个实数或变量一样参与运算。如 $T(\alpha+\beta)=T\alpha+T\beta$，就像乘法对加法的分配律一样展开运算，因此 T 在这里也叫线性算子。具体的算子比如有微分算子，积分算子，拉普拉斯算子等。

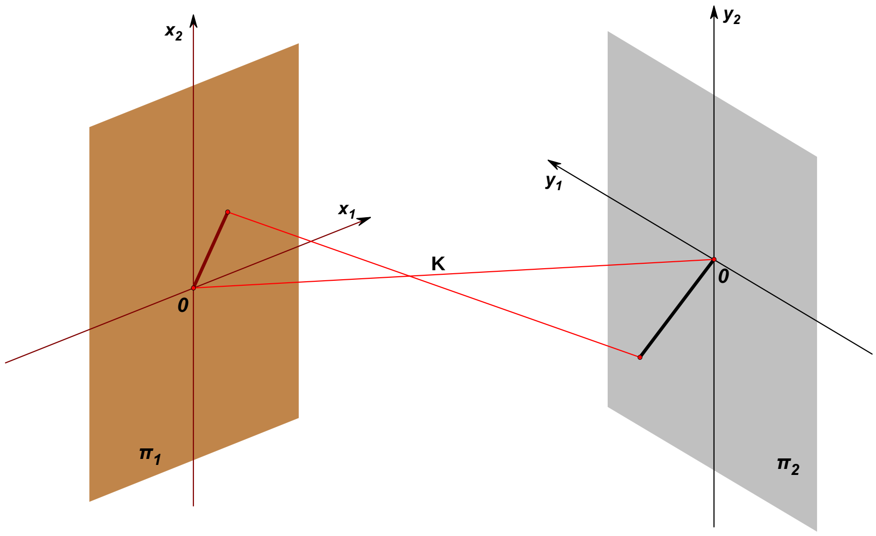
对于高等的线性函数 $f(\mathbf{x})=\mathbf{Kx}$ 而言实际上也有同样的结论。我们把表达式 $f(\mathbf{x})=\mathbf{Kx}$ 改写成 $T:\mathbf{x}\to\mathbf{y}, \mathbf{x}\mapsto\mathbf{kx}$，$T:\mathbf{x}\to\mathbf{y}$ 表示为一个从自变向量的集合 $\mathbf{x}$ 到因变向量的集合 $\mathbf{y}$ 的映射。

为了方便看到 $T:\mathbf{x}\to\mathbf{y}, \mathbf{x}\mapsto\mathbf{kx}$ 的几何解释，我们看看二维的线性函数式：

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

下面我们给出一般情况下的映射图像（细节在后续的章节里有详细描述）。

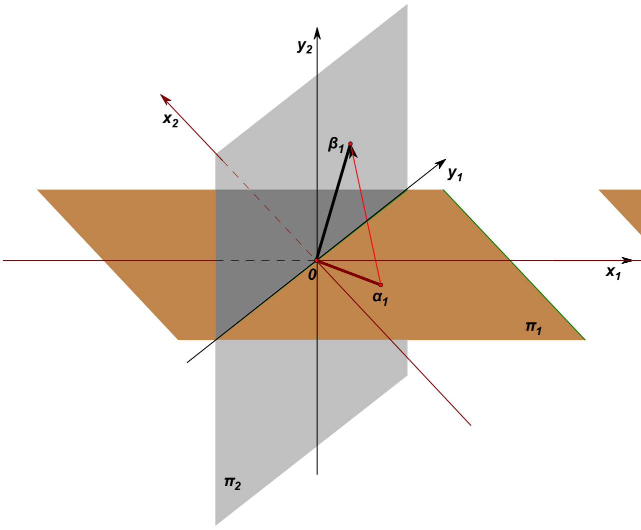
因为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 都是二维向量，所有任意的向量 $\mathbf{x}$ 的集合将构成平面 π_1，在平面 π_1 上构建二维坐标系为 $x_1 0 x_2$；所有任意的向量 $\mathbf{y}$ 的集合构成平面 π_2，在平面 π_2 上也构建二维坐标系为 $y_1 0 y_2$。所以，二维线性函数就构成了两个二维平面之间由矩阵 $\mathbf{K}=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$ 所确定的映射关系，如下图。

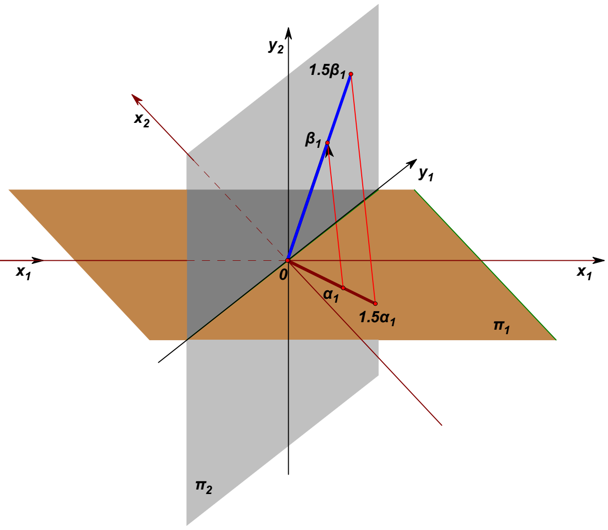


平面 π_1 的原点 0 始终映射到另一个平面 π_2 的原点 0，这是线性映射的最基本要求。

为了更紧密地观察映射之间的关系，我们把平面 π_1 放平，并使两个平面的原点 0 重合，就得到了一个由两个相交平面所构造的三维空间。

下图中，把平面 π_1 上的向量 $\mathbf{x}$ 标注为 $\boldsymbol{\alpha}_i$，把平面 π_2 上的向量 $\mathbf{y}$ 标注为 $\boldsymbol{\beta}_i$（为了和坐标系的标注区别开来）。



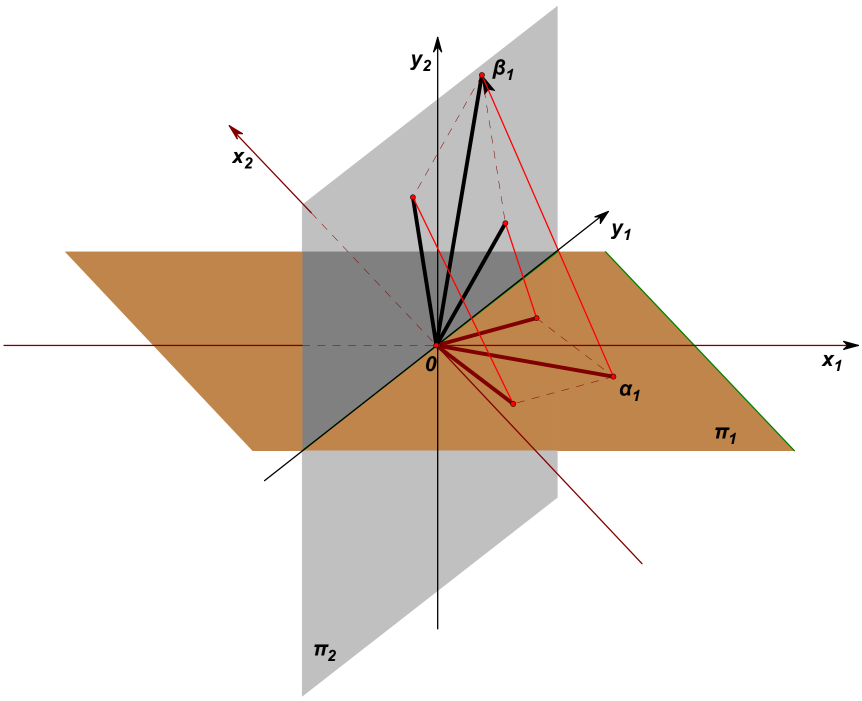


上图左表示了矩阵$\mathbf{K}$把平面π_1上的一个向量$\boldsymbol{\alpha}_1$映射到平面π_2上的向量$\boldsymbol{\beta}_1$，也即线段映射为线段。

图右给出了把一个向量$\boldsymbol{\alpha}_1$比例放大到 1.5 倍后被矩阵$\mathbf{K}$映射到平面π_2上的向量$1.5\boldsymbol{\beta}_1$，这满足线性映射的比例性。实际上，不论矩阵$\mathbf{K}$的元素是什么实数，对于任意的矩阵$\mathbf{K}$，都有这个结论。因为：

$$\begin{aligned}&\boldsymbol{\alpha}_1 \to \mathbf{K}\boldsymbol{\alpha}_1 = \boldsymbol{\beta}_1,\\&1.5\boldsymbol{\alpha}_1 \to \mathbf{K}(1.5\boldsymbol{\alpha}_1) = 1.5\mathbf{K}\boldsymbol{\alpha}_1 = 1.5\boldsymbol{\beta}_1\end{aligned}$$

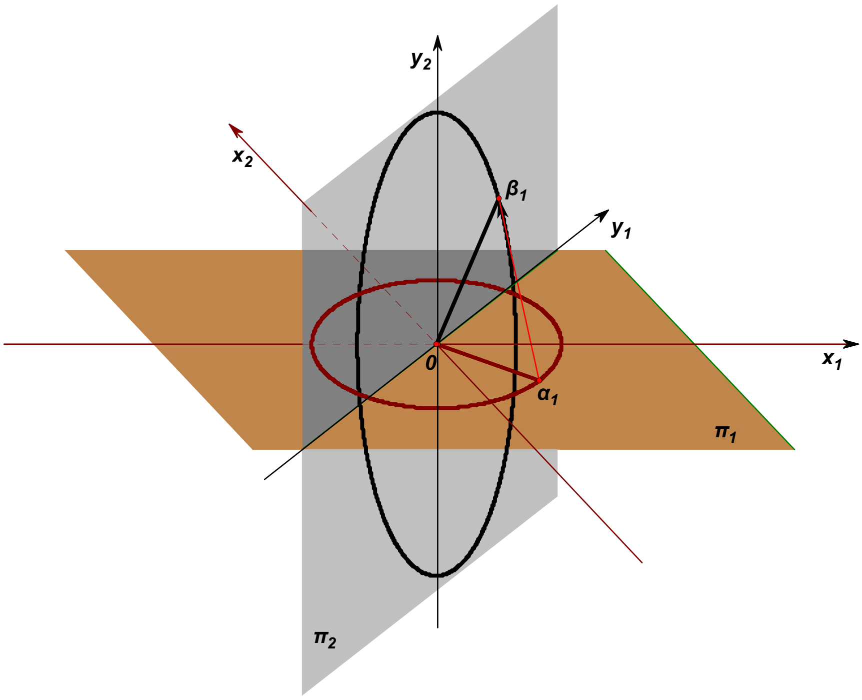
对于线性可加性，我们也给出了图形，如下图。



由上图看出，由于向量的平行四边形加法决定了，一个平面上的平行四边形被矩阵映射成为另一个平面上的平行四边形，这两个平行四边形可能全等，可能相似，大部分情况下既不全等也不相似。

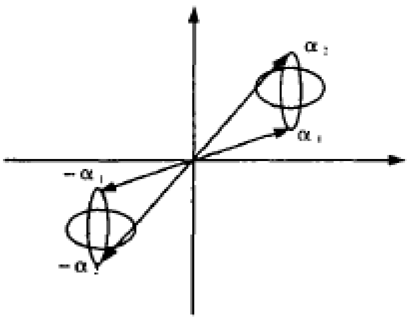
前面我们讨论了一个向量的映射或者是两个向量的和的映射情况。如果由无数等长而异向的向量构成的一个园，那么这个园会由某一个矩阵映射成什么图形呢？

实际上，可以被映射成圆、椭圆或者一根线段，特别情况下被映射成一个点（这个点必然落在原点上）。大多数情况被映射成一个椭圆，如下图。



在这里，把线性函数中 $f(\mathbf{x})=\mathbf{Kx}$ 的变量 $\mathbf{x}$ 看成了一个图形“圆”而不只是一个向量，那么函数值 $f(\mathbf{x})$ 也就成了一个变换后的图形“椭圆”。

把一个线性映射放到二维平面及三维空间中去考察，细心揣摩其几何意义，就不难理解概念的本质。例如，数乘变换 $\sigma(\alpha)=k\alpha$，我们可以把 α 看作一个几何图形，在 $k>1$ 就是对 α 放大 k 倍，在 $0<k<1$ 时就是 α 缩小 k 倍。当 $k=-1$ 时就是把 α 反方向变化。

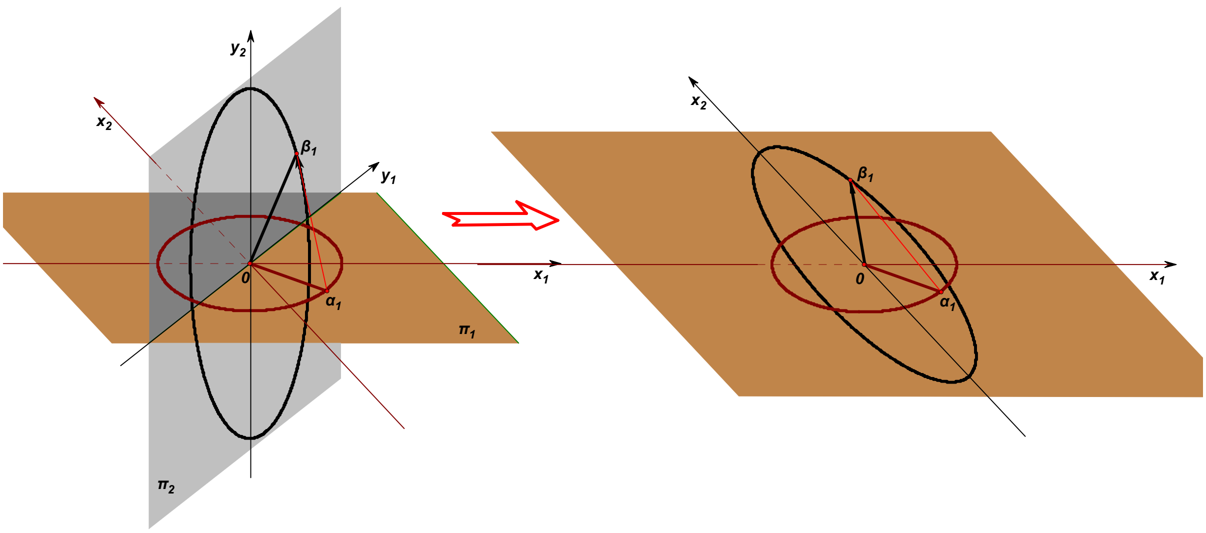


从上面的例子还能够分析出，当 k=-1 时的数乘变换实际上就是关于坐标原点的对称。

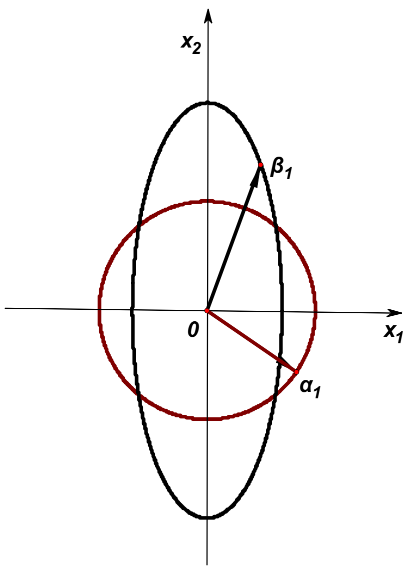
线性变换的几何意义

在大多数的教科书中，线性映射和线性变换被区别为两个概念。如果映射是发生在一个集合上的同一个坐标系中，线性映射就被称为线性变换。线性变换作为线性映射的特例，就是把集合上的两个坐标系合为一个。

如果把二维平面圆的映射整合成变换的例子如下如图所示。



把整合后的图形用直角坐标系画出图形就是:

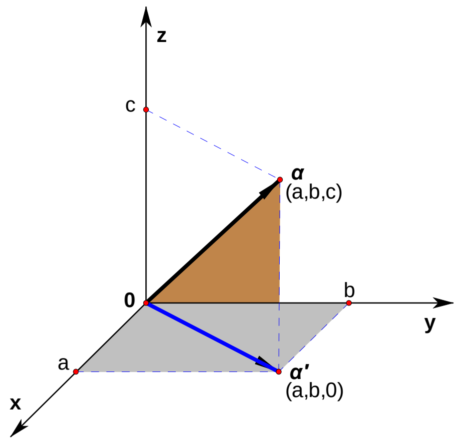


直角坐标系下的图形清楚地显示了一个图形圆被线性变换为一个椭圆。相应的，圆上的一个向量 $\boldsymbol{\alpha}_1$ 映射为椭圆上的向量 $\boldsymbol{\beta}_1$ 。

在线性代数中，我们主要是讨论由矩阵所决定的线性变换的各种特性。下面看两个具体的线性映射的例子：

在平面上所有从原点出发的向量构成的二维线性空间中，把所有向量绕原点作同样角度的旋转是一个变换。不难看出，这时向量α和β的和$\alpha+\beta$旋转所得到的向量$(\alpha+\beta)'$恰好等于α和β旋转所得到的向量α'和β'之和$\alpha'+\beta'$；数 k 与向量α的乘积 kα旋转所得到的向量 kα'恰好等于数 k 与向量α旋转所得到的α'的数乘积 kα'。这就是说，$(\alpha+\beta)'=\alpha'+\beta'$，　$(k\alpha)'=k\alpha'$。

另一个例子：在建立了空间笛卡尔直角坐标系的三维向量空间中，把每一个向量投影在坐标面 x0y 上，也是一个变换（投影变换），这时向量$\alpha=$（a，b，c），在此变换下的象为$\alpha'=$（a，b，0），显然这时也有$(\alpha+\beta)'=\alpha'+\beta'$，　$(k\alpha)'=k\alpha'$。



这两个变换有一个共同的性质：两个向量之和变换后，所得的向量恰好是把这两个向量变换后所得向量之和；数 k 与一个向量数乘后进行变换，所得的向量恰好是数 k 与把此向量变换后所得向量的乘积。此即所谓的可加性和比例性。这种使向量之间加法与数乘法关系都不受影响的变换，它与线性空间的运算相适应，能够反映线性空间中向量的内在联系，是线性空间的重要变换。所以线性变换的定义如下：

数域 F 上线性空间 V 中的变换 T 若满足条件：

$$T(\alpha+\beta)=T\alpha+T\beta(\alpha,\beta\in V)$$
$$T(k\alpha)=kT\alpha(k\in F,\alpha\in V)$$

则称为 V 中的**线性变换**。

同线性映射一样，线性变换把向量变成另外一个向量，或者说把“线”变成“线”。在平面上，线性变换把原点仍变为原点（参考零点没有移动），直线仍然变为直线（没有打弯），平行线仍然是平行线，当然平行四边形仍然是平行四边形。

在工程中常用的差分运算、微分运算及积分运算都属于线性变换，都满足以上的可加和比例性的关系。

1.3 线性代数的内容及发展简史

线性代数是高等代数的一大分支。通过前面的章节的介绍，我们知道，在研究关联着多个因素的量所引起的问题，则需要考察多元函数。如果所研究的关联性是线性的，那么称这个问题为线性问题。一次方程就是研究线性问题的方程，被称为线性方程，讨论线性方程及线性运算的代数就叫做线性代数。

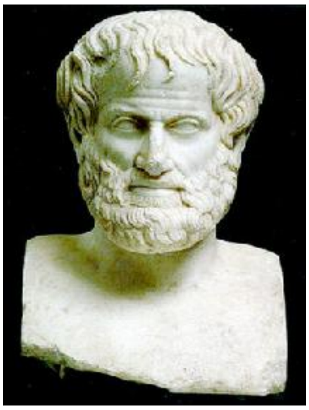
历史上线性代数的第一个问题是关于解线性方程组的问题，而线性方程组理论的发展又促成了作为工具的行列式理论和矩阵论的创立与发展，这些内容已成为我们线性代数教材的主要部分。作为代表“线性”的最基本的概念--向量的概念，从数学的观点来看不过是有序三元数组的一个集合，然而它以力或速度作为直接的物理意义，并且数学上用它能立刻写出物理上所说的事情。向量用于梯度，散度，旋度就更有说服力。同样，行列式和矩阵如导数一样（虽然 dy/dx 在数学上不过是一个符号，表示包括△y/△x 的极限的式子，但导数本身是一个强有力的概念，能使我们直接而创造性地想象物理上发生的事情）。因此，虽然表面上看，行列式和矩阵不过是一种语言或速记，但它的大多数生动的概念能对新的思想领域提供钥匙。然而已经证明这两个概念是数学物理上高度有用的工具。

用一个表格总结一下线性代数的发展历史上做出重要贡献的数学家，如下：

	人物	理论贡献	力量指数
1	关孝和（不详），日本	最早提出行列式概念	★★★
2	柯西（1789-1857），法国	1815 年启用行列式名词，1841 年提出特征方程概念	★★★★
3	西尔维斯特（1814-1897），英国	1850 年启用矩阵名词，1852 年发现惯性定律	★★★
4	凯莱（1821-1895），英国	1855 年引入定义矩阵乘法等运算	★★★★★
5	雅可比（1804-1851），德国	重新发现并证明惯性定律	★★★
6	格拉斯曼（1809-1877），德国	1844 至 1862 年间创建高维线性空间理论	★★★★★
7	外尔斯特拉斯（1815-1897），德国	1868 年完成二次型理论	★★★

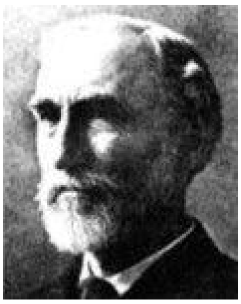
下面我们逐个概念简要介绍其发展历史。

向量



向量又称为矢量，最初被应用于物理学。很多物理量如力、速度、位移以及电场强度、磁感应强度等都是向量。大约公元前 350 年前，古希腊著名学者亚里士多德 （Aristotle）就知道了力可以表示成向量，两个力的组合作用可用著名的平行四边形法则来得到。“向量”一词来自力学、解析几何中的有向线段。最先使用有向线段表示向量的是英国科学家牛顿。 调查表明，一般日常生活中使用的的向量是一种带几何性质的量，除零向量外，总可以画出箭头表示方向。但是在高等数学中还有更广泛的向量。例如，把所有实系数多项式的全体看成一个多项式空间，这里的多项式都可看成一个向量。在这种情况下，要找出起点和终点甚至画出箭头表示方向是办不到的。这种空间中的向量比几何中的向量要广泛得多，可以是任意数学对象或物理对象。这样，就可以指导线性代数方法应用到广阔的自然科学领域中去了。因此，向量空间的概念，已成了数学中最基本的概念和线性代数的中心内容，它的理论和方法在自然科学的各领域中得到了广泛的应用。而向量及其线性运算也为“向量空间”这一抽象的概念提供出了一个具体的模型。

从数学发展史来看，历史上很长一段时间，空间的向量结构并未被数学家们所认识，直到 19 世纪末 20 世纪初，人们才把空间的性质与向量运算联系起来，使向量成为具有一套优良运算通性的数学体系。 向量能够进入数学并得到发展的阶段是 18 世纪末期，挪威测量学家威塞尔首次利用坐标平面上的点来表示复数a＋bi，并利用具有几何意义的复数运算来定义向量的运算．把坐标平面上的点用向量表示出来，并把向量的几何表示用于研究几何问题与三角问题．人们逐步接受了复数，也学会了利用复数来表示和研究平面中的向量，向量就这样平静地进入了数学。 但复数的利用是受限制的，因为它仅能用于表示平面，若有不在同一平面上的力作用于同一物体，则需要寻找所谓三维“复数”以及相应的运算体系。19 世纪中期，英国数学家汉密尔顿（右图）发明了四元数（包括数量部分和向量部分），以代表空间的向量。他的工作为向量代数和向量分析的建立奠定了基础。随后，电磁理论的发现者，英国的数学物理学家麦克思韦尔把四元数的数量部分和向量部分分开处理，从而创造了大量的向量分析。



三维向量分析的开创，以及同四元数的正式分裂，是美国的吉布斯（Gibbs,左图）和海维塞德于 19 世纪 80 年代各自独立完成的。他们提出，一个向量不过是四元数的向量部分，但不独立于任何四元数。他们引进了两种类型的乘法，即数量积和向量积。并把向量代数推广到变向量的向量微积分。从此，向量的方法被引进到分析和解析几何中来，并逐步完善，成为了一套优良的数学工具。

行列式

行列式出现于线性方程组的求解，它最早是一种速记的表达式，现在已经是数学中一种非常有用的工具。行列式是由日本数学家关孝和德国的莱布尼茨和发明的。日本数学家关孝和（右图）1683 年在其著作《解

伏题之法》中第一次提出了行列式的概念与展开算法。同时代的莱布尼兹是欧洲第一个提出行列式概念的人。他在 1693 年 4 月写给洛比达的一封信中使用并给出了行列式，并给出方程组的系数行列式为零的条件。

1750 年，瑞士数学家克莱姆 (G.Cramer,1704-1752) 在其著作《线性代数分析导言》中，对行列式的定义和展开法则给出了比较完整、明确的阐述，并给出了现在我们所称的解线性方程组的克莱姆法则。稍后，法国数学家贝祖 (E.Bezout,1730-1783) 将确定行列式每一项符号的方法进行了系统化，利用系数行列式概念指出了如何判断一个 n 个未知量的 n 个齐次线性方程组有非零解的方法，就是系数行列式等于零是这方程组有非零解的条件。

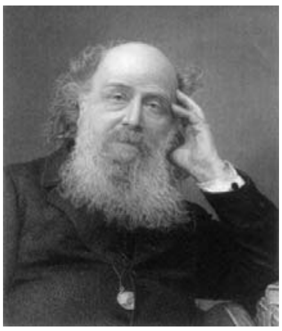
总之，在很长一段时间内，行列式只是作为解线性方程组的一种工具使用，并没有人意识到它可以独立于线性方程组之外，单独形成一门理论加以研究。

在行列式的发展史上，第一个对行列式理论做出连贯的逻辑的阐述，即把行列式理论与线性方程组求解相分离的人，是法国数学家范德蒙 (A-T.Vandermonde,1735-1796) 。范德蒙自幼在父亲的知道下学习音乐，但对数学有浓厚的兴趣，后来终于成为法兰西科学院院士。特别地，他给出了用二阶子式和它们的余子式来展开行列式的法则。就对行列式本身这一点来说，他是这门理论的奠基人。 1772 年，拉普拉斯在一篇论文中证明了范德蒙提出的一些规则，推广了他的展开行列式的方法。

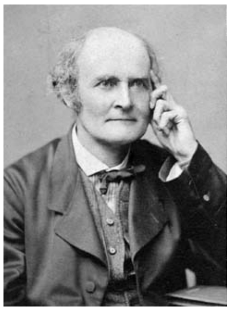
继范德蒙之后，在行列式的理论方面，又一位做出突出贡献的就是另一位法国大数学家柯西(Cauchy)。 1815 年，柯西在一篇论文中给出了行列式的第一个系统的、几乎是近代的处理。其中主要结果之一是行列式的乘法定理。另外，他第一个把行列式的元素排成方阵，采用双足标记法；引进了行列式特征方程的术语；给出了相似行列式概念；改进了拉普拉斯的行列式展开定理并给出了一个证明等。

继柯西之后，在行列式理论方面最多产的人就是德国数学家雅可比 (J.Jacobi,1804-1851) ，他引进了函数行列式，即“雅可比行列式”，指出函数行列式在多重积分的变量替换中的作用，给出了函数行列式的导数公式。雅可比的著名论文《论行列式的形成和性质》标志着行列式系统理论的建成。由于行列式在数学分析、几何学、线性方程组理论、二次型理论等多方面的应用，促使行列式理论自身在 19 世纪也得到了很大发展。整个 19 世纪都有行列式的新结果。除了一般行列式的大量定理之外，还有许多有关特殊行列式的其他定理都相继得到。

矩 阵



矩阵是数学中的一个重要的基本概念，是代数学的一个主要研究对象，也是数学研究和应用的一个重要工具。“矩阵”这个词是由西尔维斯特（James Joseph Sylvester, 1814 - 1897）首先使用的，他是为了将数字的矩形阵列区别于行列式而发明了这个术语。矩阵这个词，它来源于拉丁语，代表一排数。而实际上，矩阵这个课题在诞生之前就已经发展的很好了。从行列式的大量工作中明显的表现出来，为了很多目的，不管行列式的值是否与问题有关，方阵本身都可以研究



和使用，矩阵的许多基本性质也是在行列式的发展中建立起来的。在逻辑上，矩阵的概念应先于行列式的概念，然而在历史上次序正好相反。

英国数学家凯莱 (A.Cayley,1821-1895) 一般被公认为是矩阵论的创立者，因为他首先把矩阵作为一个独立的数学概念提出来，并首先发表了关于这个题目的一系列文章。凯莱同研究线性变换下的不变量相结合，首先引进矩阵以简化记号。 1858 年，他发表了关于这一课题的第一篇论文《矩阵论的研究报告》，系统地阐述了关于矩阵的理论。文中他定义了矩阵的相等、矩阵的运算法则、矩阵的转置以及矩阵的逆等一系列基本概念，指出了矩阵加法的可交换性与可结合性。他用单一的字母 A 来表示矩阵是对矩阵代数发展至关重要的，其公式 $\det(AB)=\det(A)\det(B)$ 为矩阵代数和行列式间提供了一种联系。另外，凯莱还给出了方阵的特征方程和特征根（特征值）以及有关矩阵的一些基本结果。凯莱出生于一个古老而有才能的英国家庭，剑桥大学三一学院大学毕业后留校讲授数学，三年后他转从律师职业，工作卓有成效，并利用业余时间研究数学，发表了大量的数学论文。

1855 年，埃米特 (C.Hermite,1822-1901) 证明了别的数学家发现的一些矩阵类的特征根的特殊性质，如现在称为埃米特矩阵的特征根性质等。后来 ，克莱伯施 (A.Clebsch,1831-1872) 、布克海姆 (A.Buchheim) 等证明了对称矩阵的特征根性质。泰伯 (H.Taber) 引入矩阵的迹的概念并给出了一些有关的结论。

在矩阵论的发展史上，弗罗伯纽斯 (G.Frobenius,1849-1917) 的贡献是不可磨灭的。他讨论了最小多项式问题，引进了矩阵的秩、不变因子和初等因子、正交矩阵、矩阵的相似变换、合同矩阵等概念，以合乎逻辑的形式整理了不变因子和初等因子的理论，并讨论了正交矩阵与合同矩阵的一些重要性质。 1854 年，约当研究了矩阵化为标准型的问题。 1892 年，梅茨勒 (H.Metzler) 引进了矩阵的超越函数概念并将其写成矩阵的幂级数的形式。傅立叶、西尔和庞加莱的著作中还讨论了无限阶矩阵问题，这主要是适用方程发展的需要而开始的。

矩阵本身所具有的性质依赖于元素的性质，矩阵由最初作为一种工具经过两个多世纪的发展，现在已成为独立的一门数学分支——矩阵论。而矩阵论又可分为矩阵方程论、矩阵分解论和广义逆矩阵论等矩阵的现代理论。矩阵及其理论现已广泛地应用于现代科技的各个领域。

矩阵的发展是与线性变换密切相连的。到 19 世纪它还仅占线性变换理论形成中有限的空间。现代向量空间的定义是由皮亚诺（Peano）于 1888 年提出的，皮亚诺以公理的方式定义了有限维或无限维向量空间。二次世界大战后随着现代数字计算机的发展，矩阵又有了新的含义，特别是在矩阵的数值分析等方面。 由于计算机的飞速发展和广泛应用，许多实际问题可以通过离散化的数值计算得到定量的解决。于是作为处理离散问题的线性代数，成为从事科学研究和工程设计的科技人员必备的数学基础。

线性方程组

线性方程组的解法，早在中国古代的数学著作《九章算术 方程》章中已作了比较完整的论述。在中国人的手稿中出现了解释如何消去变元的方法求解带有三个未知量的三方程系统，其中所述方法实质上相当于现代的对方程组的增广矩阵施行初等行变换从而消去未知量的方法，即高斯消元法。在西方，线性方程组的研究是在 17 世纪后期由莱布尼茨开创的。他曾研究含两个未知量的三个线性方程组组成的方程组。麦克劳林在 18 世纪上半叶研究了具有二、三、四个未知量的线性方程组，得到了现在称为克莱姆法则的结果。克莱姆不久也发表了这个法则。 18 世纪下半叶，法国数学家贝祖对线性方程组理论进行了一系列研究，证明了 n 个 n 元齐次线性方程组有非零解的条件是系数行列式等于零。

19 世纪，英国数学家史密斯 (H.Smith) 和道奇森 (C-L.Dodgson，见图) 继续研究线性方程组理论，前者引进了方程组的增广矩阵和非增广矩阵的概念，后者证明了 n个未知数m 个方程的方程组相容的充要条件是系数矩阵和增广矩阵的秩相同。这正是现代方程组理论中的重要结果之一。

大量的科学技术问题，最终往往归结为解线性方程组。因此在线性方程组的数值解法得到发展的同时，线性方程组解的结构等理论性工作也取得了令人满意的进展。现在，线性方程组的数值解法在计算数学中占有重要地位。

二次型

二次型也称为“二次形式”。二次型的系统研究是从 18 世纪开始的，它起源于对二次曲线和二次曲面的分类问题的讨论。将二次曲线和二次曲面的方程变形，选有主轴方向的轴作为坐标轴以简化方程的形状，这个问题是在 18 世纪引进的。柯西在其著作中给出结论：当方程是标准型时，二次曲面用二次项的符号来进行分类。然而，那时并不太清楚，在化简成标准型时，为何总是得到同样数目的正项和负项。西尔维斯特回答了这个问题，他给出了 个变数的二次型的惯性定律，但没有证明。这个定律后被雅可比重新发现和证明。 1801 年，高斯在《算术研究》中引进了二次型的正定、负定、半正定和半负定等术语。

1.4 线性代数有什么用？

线性代数有什么用？这是每一个圈养在象牙塔里，在灌输式教学模式下的“被学习”的学生刚刚开始思考时的第一个问题。我稍微仔细的整理了一下学习线代的理由，竟然也罗列了不少，不知道能不能说服你：

1、 如果你想顺利地拿到学位，线性代数的学分对你有帮助；

2、 如果你想继续深造，考研，必须学好线代。因为它是必考的数学科目，也是研究生科目《矩阵论》、《泛函分析》的基础。例如，泛函分析的起点就是无穷多个未知量的无穷多线性方程组理论。

3、 如果你想提高自己的科研能力，不被现代科技发展潮流所抛弃，也必须学好，因为瑞典的L.

戈丁说过，没有掌握线代的人简直就是文盲。他在自己的数学名著《数学概观》中说：

要是没有线性代数，任何数学和初等教程都讲不下去。按照现行的国际标准，线性代数是通过公理化来表述的。它是第二代数学模型，其根源来自于欧几里得几何、解析几何以及线性方程组理论。…，如果不熟悉线性代数的概念，像线性性质、向量、线性空间、矩阵等等，要去学习自然科学，现在看来就和文盲差不多，甚至可能学习社会科学也是如此。

4、 如果毕业后想找个好工作，也必须学好线代：

- 想搞数学，当个数学家（我靠，这个还需要列出来，谁不知道线代是数学）。恭喜你，你的职业未来将是最光明的。如果到美国打工的话你可以找到最好的职业（参考本节后附的一份小资料）。
- 想搞电子工程，好，电路分析、线性信号系统分析、数字滤波器分析设计等需要线代，因为线代就是研究线性网络的主要工具；进行IC集成电路设计时，对付数百万个集体管的仿真软件就需要依赖线性方程组的方法；想搞光电及射频工程，好，电磁场、光波导分析都是向量场的分析，比如光调制器分析研制需要张量矩阵，手机信号处理等等也离不开矩阵运算。
- 想搞软件工程，好，3D游戏的数学基础就是以图形的矩阵运算为基础；当然，如果你只想玩3D游戏可以不必掌握线代；想搞图像处理，大量的图像数据处理更离不开矩阵这个强大的工具，《阿凡达》中大量的后期电脑制作没有线代的数学工具简直难以想象。
- 想搞经济研究。好，知道列昂惕夫（Wassily Leontief）吗？哈佛大学教授，1949年用计算机计算出了由美国统计局的25万条经济数据所组成的42个未知数的42个方程的方程组，他打开了研究经济数学模型的新时代的大门。这些模型通常都是线性的，也就是说，它们是用线性方程组来描述的，被称为列昂惕夫“投入-产出”模型。列昂惕夫因此获得了1973年的诺贝尔经济学奖。
- 相当领导，好，要会运筹学，运筹学的一个重要议题是线性规划。许多重要的管理决策是在线性规划模型的基础上做出的。线性规划的知识就是线代的知识啊。比如，航空运输业就使用线性规划来调度航班，监视飞行及机场的维护运作等；又如，你作为一个大商场的老板，线性规划可以帮助你合理的安排各种商品的进货，以达到最大利润。
- 对于其他工程领域，没有用不上线代的地方。如搞建筑工程，那么奥运场馆鸟巢的受力分析需要线代的工具；石油勘探，勘探设备获得的大量数据所满足的几千个方程组需要你的线代知识来解决；飞行器设计，就要研究飞机表面的气流的过程包含反复求解大型的线性方程组，在这个求解的过程中，有两个矩阵运算的技巧：对稀疏矩阵进行分块处理和进行LU分解； 作餐饮业，对于构造一份有营养的减肥食谱也需要解线性方程组；知道有限元方法吗？这个工程分析中十分有效的有限元方法，其基础就是求解线性方程组。知道马尔科夫链吗？这个“链子”神通广大，在许多学科如生物学、商业、化学、工程学及物理学等领域中被用来做数学模型，实际上马尔科夫链是由一个随机变量矩阵所决定的一个概率向量序列，看看，矩阵、向量又出现了。
- 另外，矩阵的特征值和特征向量可以用在研究物理、化学领域的微分方程、连续的或离散的动力系统中，甚至数学生态学家用以在预测原始森林遭到何种程度的砍伐会造成猫头鹰的种群灭亡；大名鼎鼎的最小二乘算法广泛应用在各个工程领域里被用来把实验中得到的大量测量数据来拟合到一个理想的直线或曲线上，最小二乘拟合算法实质就是超定线性方程组的求解；二次型常常出现在线性代数在工程（标准设计及优化）和信号处

理（输出的噪声功率）的应用中，他们也常常出现在物理学（例如势能和动能）、微分几何（例如曲面的法曲率）、经济学（例如效用函数）和统计学（例如置信椭圆体）中，某些这类应用实例的数学背景很容易转化为对对称矩阵的研究。

嘿嘿（脸红），说实在的，我也没有足够经验讲清楚线代在各个工程领域中的应用，只能大概人云亦云地讲述以上线代的一些基本应用。因为你如果要真正的讲清楚 线代的一个应用，就必须充分了解所要应用的领域内的知识，最好有实际的工程应用的经验在里面；况且线性代数在各个工程领域中的应用真是太多了，要知道当今成为一个工程通才只是一个传说。

总结一下，线性代数的应用领域几乎可以涵盖所有的工程技术领域。如果想知道更详细的应用材料，建议看一下《线性代数及应用》，这是美国David C. Lay 教授写的迄今最现代的流行教材。国内的教材可以看看《线性代数实践及MATLAB入门》，这是西电科大陈怀琛教授写的最实用的新教材。

附录: **来自金融危机时期的学好数学的利好消息**

2009 年元月 26 日华尔街日报刊登一篇关于职业优劣评比的报导，标题为 'Doing the Math to Find the Good Jobs'，该文引述 Les Krantz 根据美国劳工统计局与人口普查局 (U.S. Bureau of Labor Statistics and the Census Bureau) 的资料所做的一项整理研究，依据五项指标：工作环境，所得，职业前景，体力要求和压力，针对 200 种职业进行综合评比排名。

(阅读原文http://online.wsj.com/article/SB123119236117055127.html)

美国的职业评比排名结果可能出乎大多数读者意料之外，现将排名最前和最后的十种工作抄录于下 (数字代表排名)：

前十名职业：

1. 数学家 (Mathematician)
2. 保险统计师 (Actuary)
3. 统计学家 (Statistician)
4. 生物学家 (Biologist)
5. 软件工程师 (Software engineer)
6. 计算机系统分析师 (Computer systems analyst)
7. 历史学家 (Historian)
8. 社会学家 (Sociologist)
9. 工业设计师 (Industrial designer)
10. 会计师 (Accountant)

后十名职业：

200. 伐木工 (Lumberjack)

199. 酪农 (Dairy farmer)

198. 出租车司机 (Taxi driver)

197. 船员 (Seaman)

196. 屋顶工 (Roofer)

195. 清洁队员 (Garbage collector)

194. 焊工 (Welder)

193. 码头工 (Roustabout)

192. 钢架工 (Ironworker)

191. 建筑工 (Construction worker)

数学家之所以排名领先的部分原因是他们的工作环境相对舒适，不会接触有毒物质，不用提重物或弯腰爬行，而且收入颇丰。据估计，美国数学家的平均年薪达 94,160 美元。

文中提及 19 年前珍妮弗.寇特因为考虑低工作压力而选择研习数学，现年 38 岁的她，年收入高于前述平均薪资。她参与一个以数学为基础的计算机程序设计团队，过去所开发的程序曾经应用于电影<黑客任务 (The Matrix)>和<骇速快手 (Speed Racer)>。她在家里和同事以网络联系，绝少超时工作，没有什么工作压力。珍妮弗说：“解决问题牵涉许多思考，我发现那会让人冷静下来”。

以下章节待续。。。