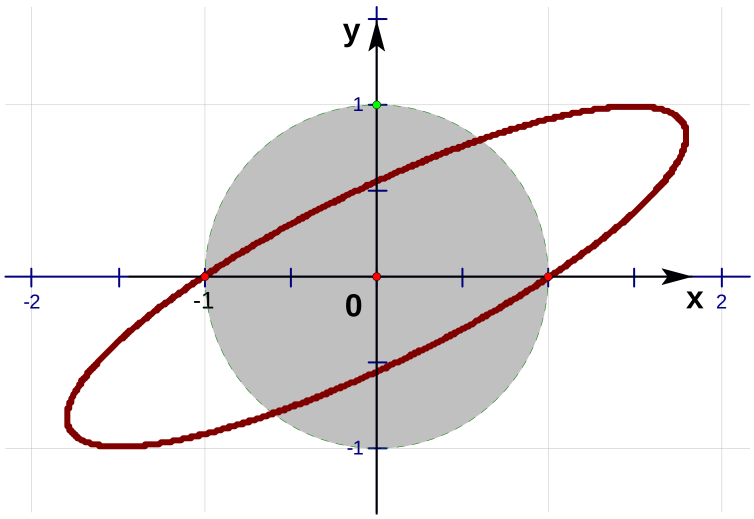
# 线性代数的几何意义

## ————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。

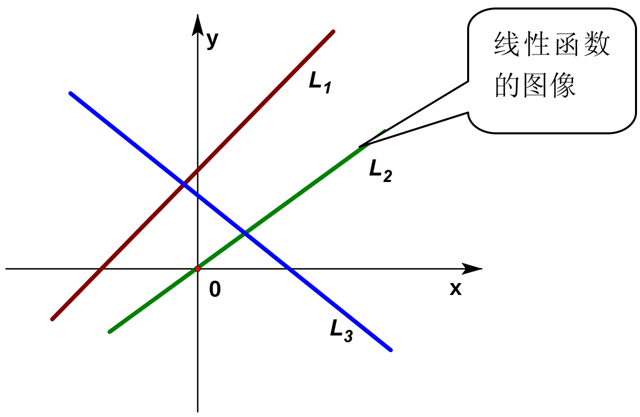
## 1.1 “线性”的意义

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

$$f(kx)=kf(x)$$ 其中 $k$ 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

$$f(k_1x_1+k_2x_2)=k_1f(x_1)+k_2f(x_2)$$，其中 $k_1,k_2$ 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

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

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$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{pmatrix}$，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：

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

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：

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

4) 上式的形式为进一步简写为：

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

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

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。

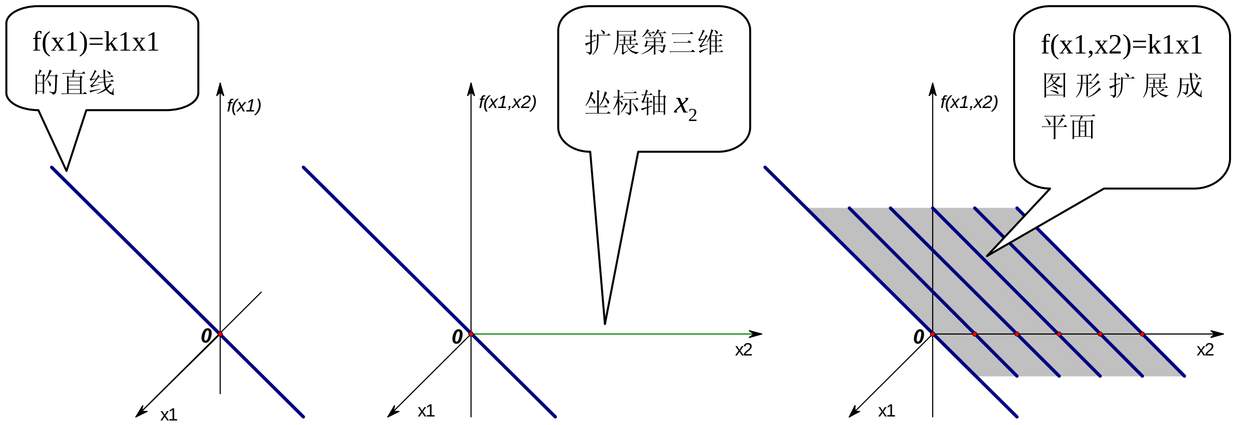
首先我们看看从一元线性函数 $f(x) = kx$ 拓展到二元的线性函数 $f(x_1, x_2) = k_1x_1 + k_2x_2$ 的几何解释。这个几何解释并不是唯一的，目的是让读者认识和理解线性的概念。

**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。

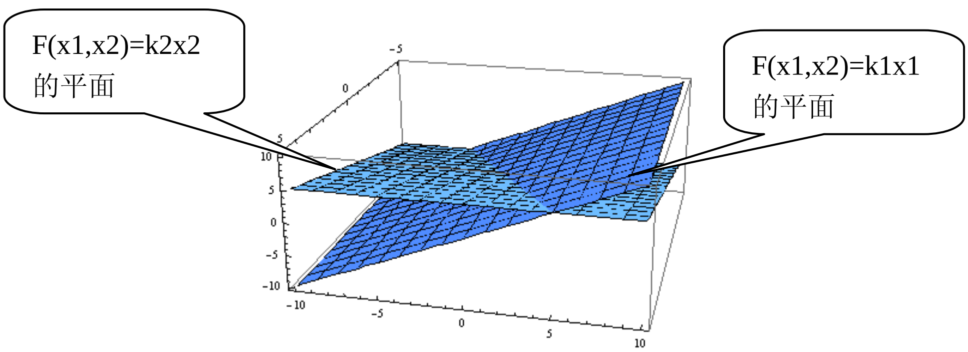
我们知道，$f(x_1,x_2)=k_1x_1$ 的图形是一个过原点的平面。扩维后由一根直线变成了一个平面。这是因为函数 $f(x_1,x_2)=k_1x_1$ 与新生长出来的坐标轴 $x_2$ 没有关系，$x_2$ 可以取任意值；换句话说，$x_2$ 的任意值都在函数 $f(x_1,x_2)=k_1x_1$ 的图像上。进一步说，这是一个过 $x_2$ 坐标轴的平面。

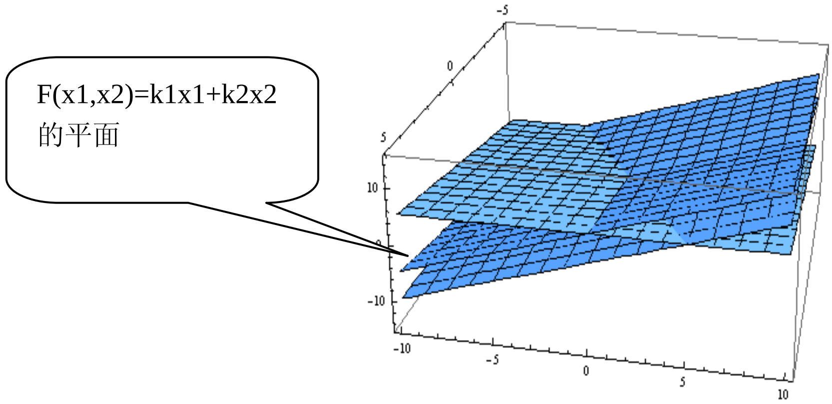



形象的扩展过程可以这样想象：二维平面坐标系里有一根直线图形，这时有 $x_2$ 轴过原点以垂直于坐标系 $x_1\sim f(x_1)$ 的平面向右方向（右手系）生长出来，然后原来的那条直线 $f(x_1)=k_1x_1$ 沿着坐标轴 $x_2$ 方向向右滑动，无数个平行的直线被 $x_2$ 轴象竹帘子一样串起来，平铺得到了 $f(x_1,x_2)=k_1x_1$ 的平面。这个平面是由无数的直线铺成的，因此，平面也是“线性”的。

**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”。

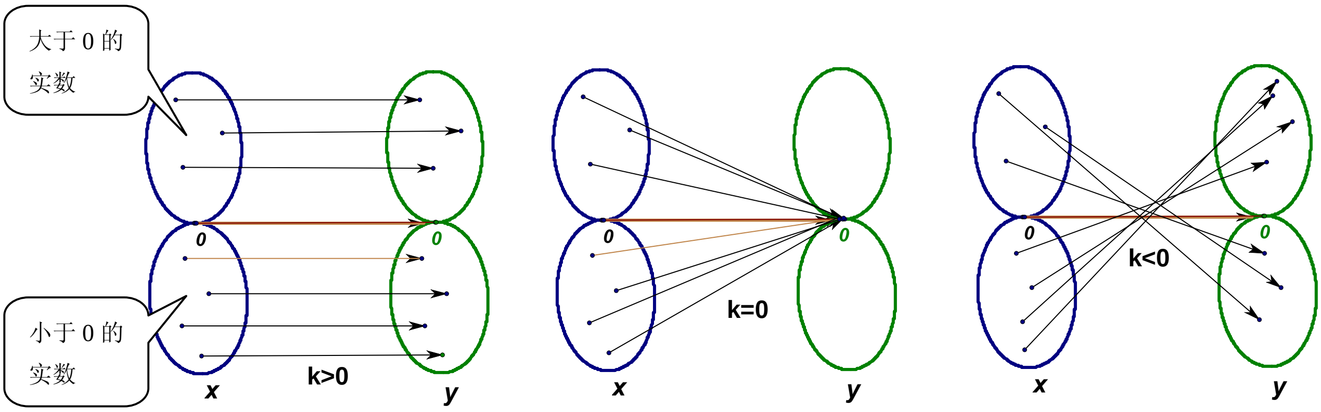
我们再来回味一下这句关于线性函数中心性质的话：线性组合的函数，等于函数的线性组合，详细说来就是，如果自变量从 $x$ 变换为自变量的线性组合 $k_1x_1+k_2x_2$ 时，其函数也从 $f(x)$ 变换为函数的线性组合 $k_1f(x_1)+k_2f(x_2)$ 。因为函数的本意是因变量与自变量之间的对应关系，所以“线性”的本质就是因变量与自变量之间始终保持组合形式不变的一种对应关系，我们把这类特殊的对应关系称之为“线性关系”。因此我们所说的“线性”，实质上就是指变量之间的“线性关系”。

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。

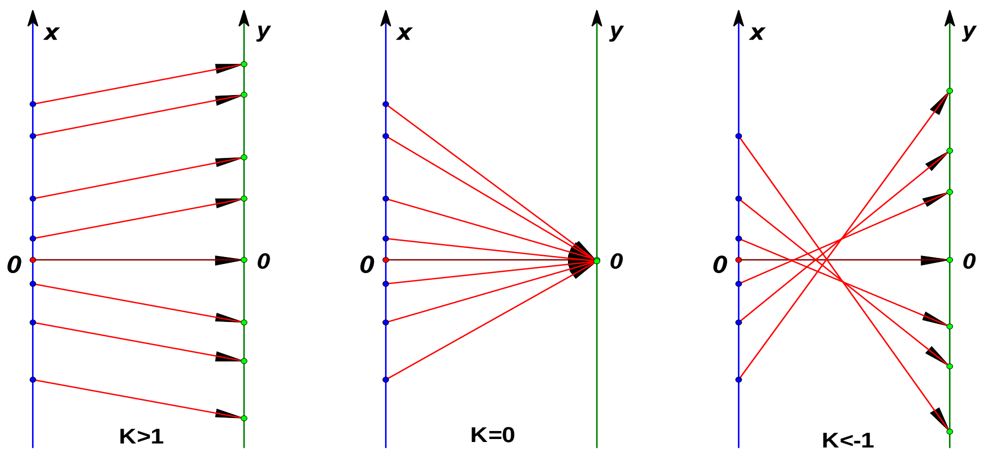
如果我们给出一个映射的集合示意图，则有如下图所示。



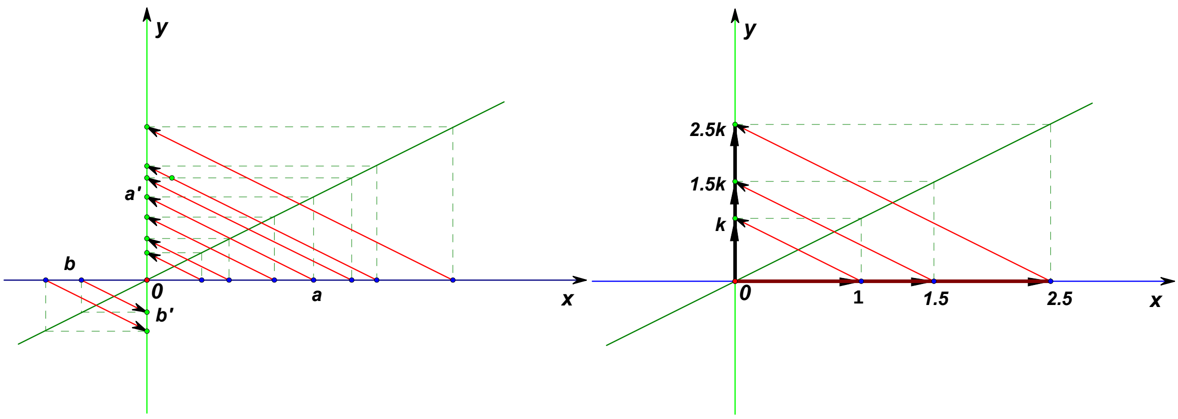

图中给出了一元线性齐次函数 $f(x)=kx$ 当 $k$ 取不同的数时的映射对应关系。在三个图中，有一个共性就是元素 0 必然映射到元素 0。

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 $\mathbf{x}$ 和集合 $\mathbf{y}$ 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 $\mathbf{x}$ 的坐标系就是一根轴，写为 $x$ 轴；集合 $\mathbf{y}$ 的坐标系也是一根轴，写为 $\mathbf{y}$ 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， $\alpha, \beta$ 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。

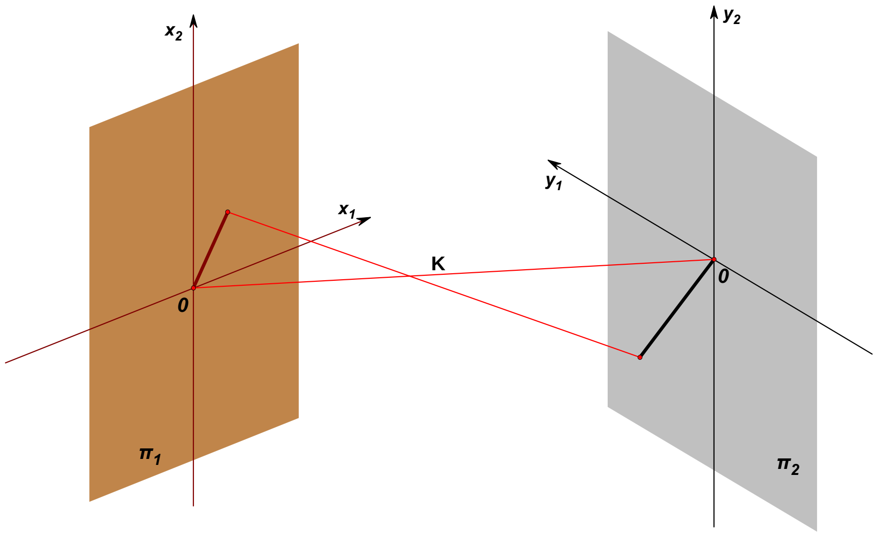
对于高等的线性函数 $f(\mathbf{x}) = \mathbf{Kx}$ 而言实际上也有同样的结论。我们把表达式 $f(\mathbf{x}) = \mathbf{Kx}$ 改写成 $T: \mathbf{x} \to \mathbf{y}, \mathbf{x} \mapsto \mathbf{kx}$ ， $T: \mathbf{x} \to \mathbf{y}$ 表示为一个从自变向量的集合 $\mathbf{x}$ 到因变向量的集合 $\mathbf{y}$ 的映射。

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）。

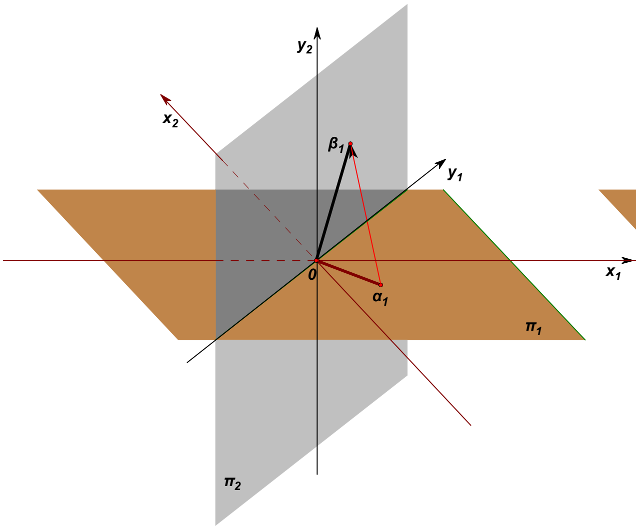
因为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 都是二维向量，所有任意的向量 $\mathbf{x}$ 的集合将构成平面 $\pi_1$，在平面 $\pi_1$ 上构建二维坐标系为 $x_1 0 x_2$；所有任意的向量 $\mathbf{y}$ 的集合构成平面 $\pi_2$，在平面 $\pi_2$ 上也构建二维坐标系为 $y_1 0 y_2$。所以，二维线性函数就构成了两个二维平面之间由矩阵 $\mathbf{K}=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$ 所确定的映射关系，如下图。

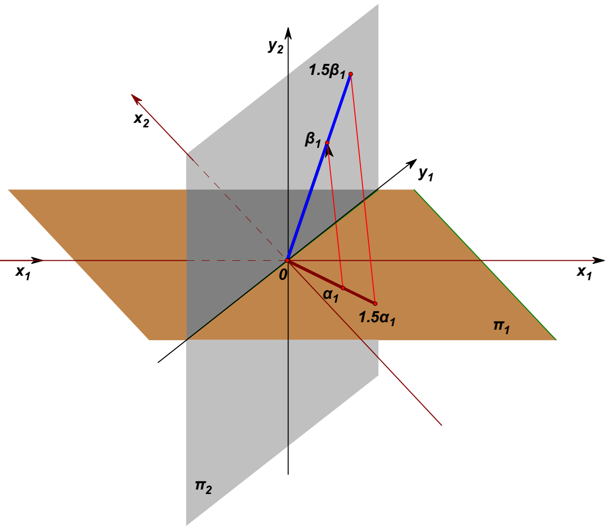



平面 $\pi_1$ 的原点 0 始终映射到另一个平面 $\pi_2$ 的原点 0，这是线性映射的最基本要求。

为了更紧密地观察映射之间的关系，我们把平面 $\pi_1$ 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 $\pi_1$ 上的向量 $\mathbf{x}$ 标注为 $\boldsymbol{\alpha}_i$，把平面 $\pi_2$ 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）。



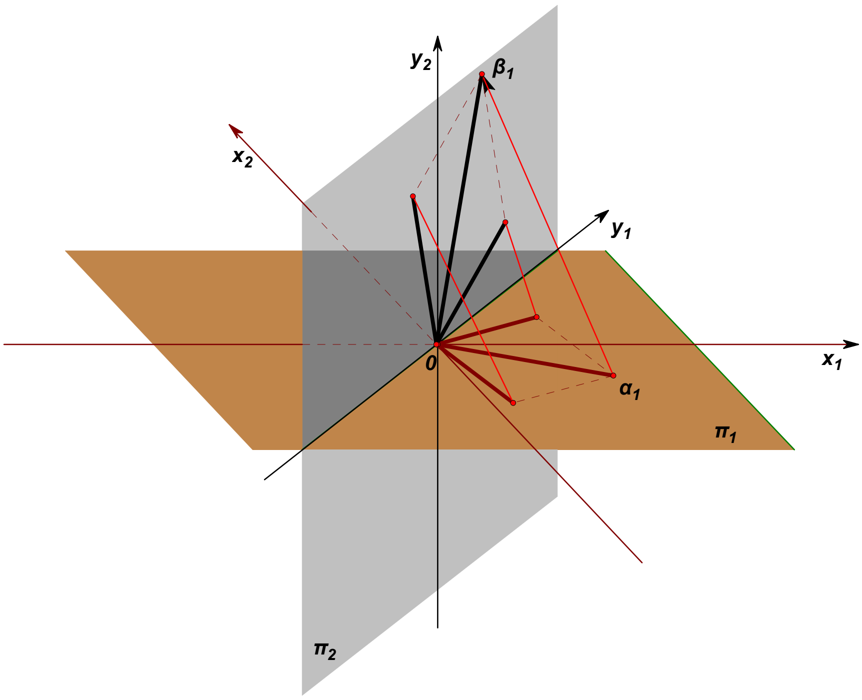




上图左表示了矩阵 $\mathbf{K}$ 把平面 $\pi_1$ 上的一个向量 $\boldsymbol{\alpha}_1$ 映射到平面 $\pi_2$ 上的向量 $\boldsymbol{\beta}_1$，也即线段映射为线段。

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 $\pi_2$ 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：

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

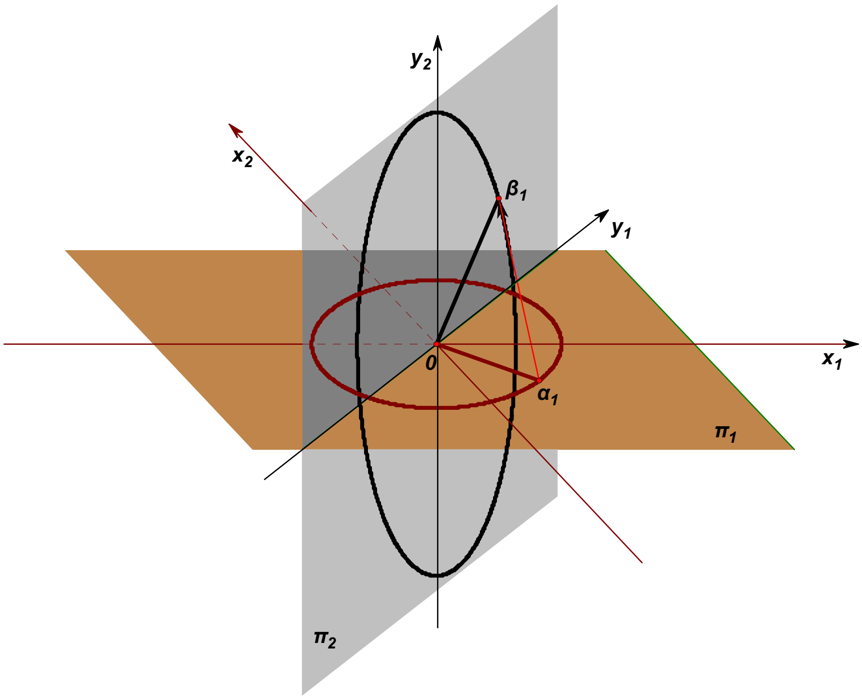
对于线性可加性，我们也给出了图形，如下图。



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。

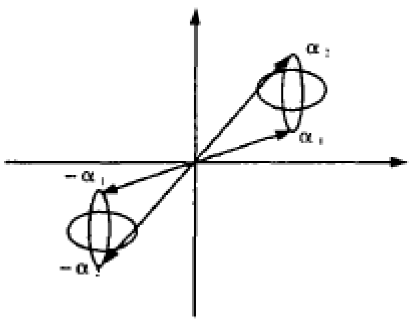
前面我们讨论了一个向量的映射或者是两个向量的和的映射情况。如果由无数等长而异向的向量构成的一个园，那么这个园会由某一个矩阵映射成什么图形呢？

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 $\alpha$ 看作一个几何图形，在 $k>1$ 就是对 $\alpha$ 放大 $k$ 倍，在 $0<k<1$ 时就是 $\alpha$ 缩小 $k$ 倍。当 $k=-1$ 时就是把 $\alpha$ 反方向变化。

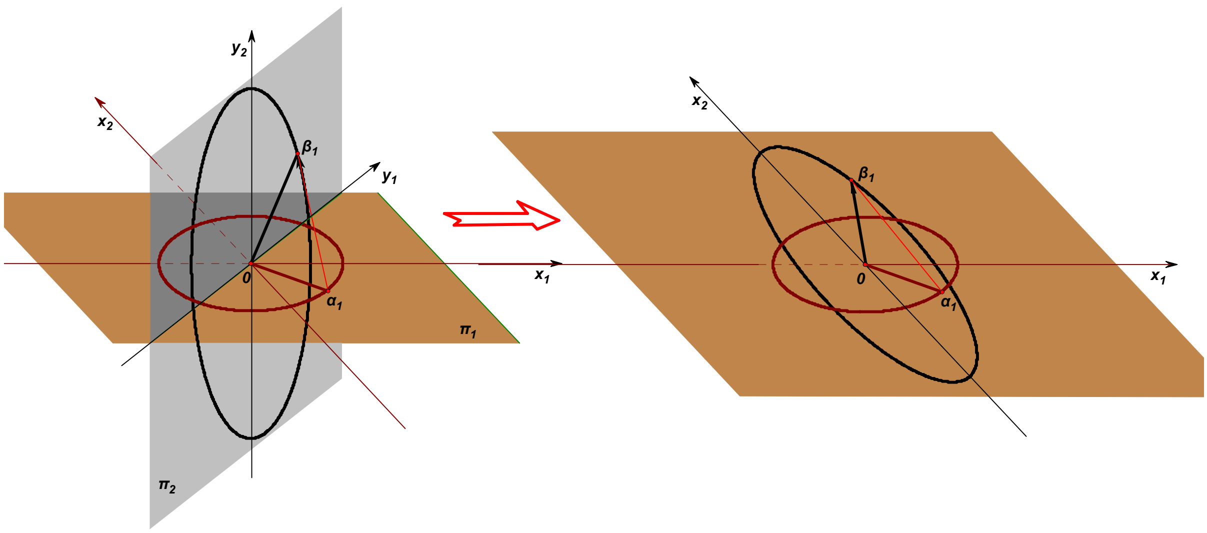


从上面的例子还能够分析出，当 k=-1 时的数乘变换实际上就是关于坐标原点的对称。

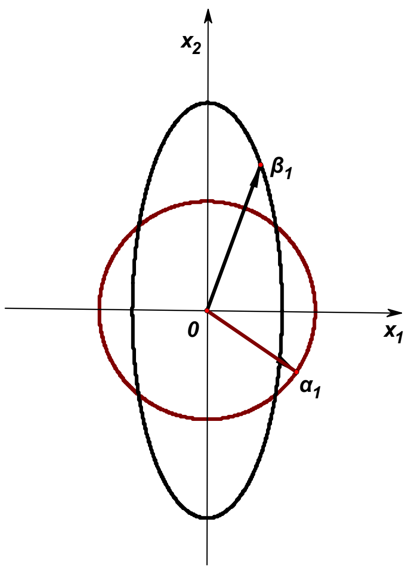
## 线性变换的几何意义

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:

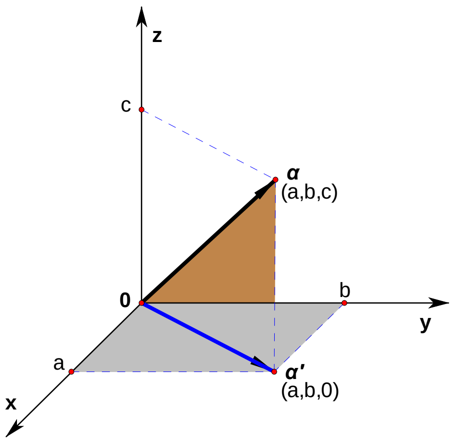



直角坐标系下的图形清楚地显示了一个图形圆被线性变换为一个椭圆。相应的，圆上的一个向量 $\boldsymbol{\alpha}_1$ 映射为椭圆上的向量 $\boldsymbol{\beta}_1$ 。

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$\alpha$和$\beta$的和$\alpha+\beta$旋转所得到的向量$(\alpha+\beta)'$恰好等于$\alpha$和$\beta$旋转所得到的向量$\alpha'$和$\beta'$之和$\alpha'+\beta'$；数 k 与向量$\alpha$的乘积 k$\alpha$旋转所得到的向量 k$\alpha'$恰好等于数 k 与向量$\alpha$旋转所得到的$\alpha'$的数乘积 k$\alpha'$。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。

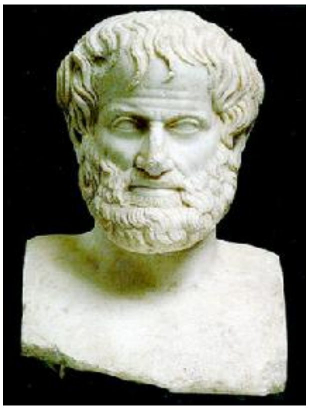
历史上线性代数的第一个问题是关于解线性方程组的问题，而线性方程组理论的发展又促成了作为工具的行列式理论和矩阵论的创立与发展，这些内容已成为我们线性代数教材的主要部分。作为代表“线性”的最基本的概念--向量的概念，从数学的观点来看不过是有序三元数组的一个集合，然而它以力或速度作为直接的物理意义，并且数学上用它能立刻写出物理上所说的事情。向量用于梯度，散度，旋度就更有说服力。同样，行列式和矩阵如导数一样（虽然 dy/dx 在数学上不过是一个符号，表示包括△y/△x 的极限的式子，但导数本身是一个强有力的概念，能使我们直接而创造性地想象物理上发生的事情）。因此，虽然表面上看，行列式和矩阵不过是一种语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

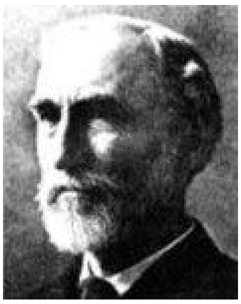
下面我们逐个概念简要介绍其发展历史。

## 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。

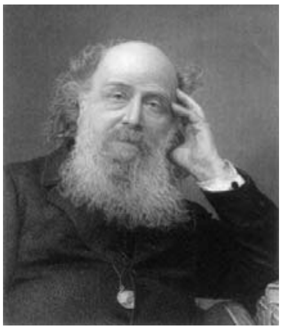
总之，在很长一段时间内，行列式只是作为解线性方程组的一种工具使用，并没有人意识到它可以独立于线性方程组之外，单独形成一门理论加以研究。

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。

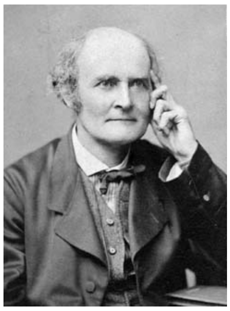
继范德蒙之后，在行列式的理论方面，又一位做出突出贡献的就是另一位法国大数学家柯西(Cauchy)。 1815 年，柯西在一篇论文中给出了行列式的第一个系统的、几乎是近代的处理。其中主要结果之一是行列式的乘法定理。另外，他第一个把行列式的元素排成方阵，采用双足标记法；引进了行列式特征方程的术语；给出了相似行列式概念；改进了拉普拉斯的行列式展开定理并给出了一个证明等。

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”。

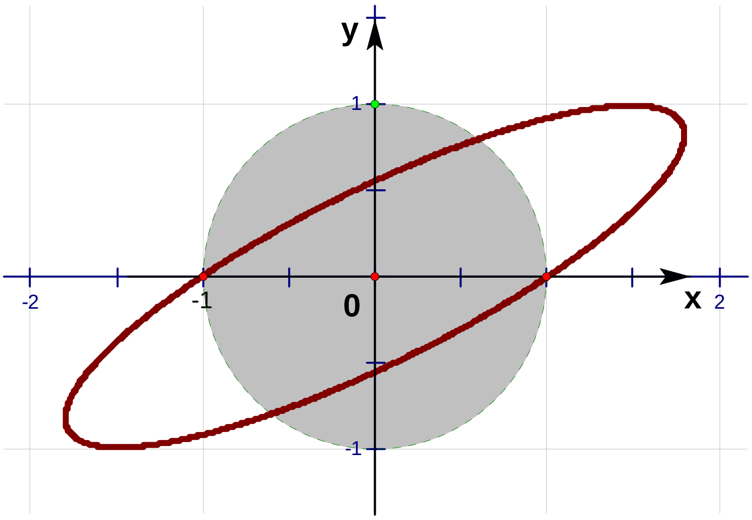
以下章节待续。。。

---图解线性代数---

# 线性代数的几何意义

# 之（2）

任广千　胡翠芳 编著




2010.04.25

# 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

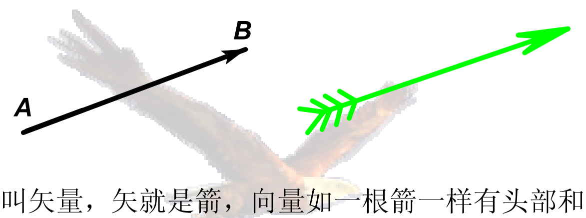
# 第二章　向量的基本几何意义

向量的概念始终贯穿当代科学的主要内容中，也始终贯穿线性代数的主要内容中，所以我们不妨回顾回顾这个概念的几何意义，以期更清晰地理解线性代数的几何本质。

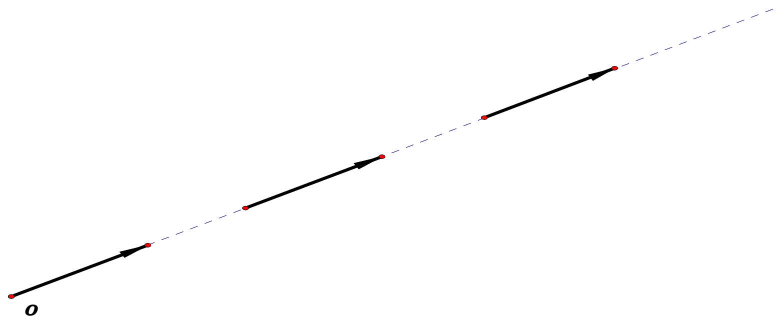
## 2.1　向量概念的几何意义

### 自由向量的概念

向量（Vector）和标量的概念是发明四元数的爱尔兰数学家W。R。哈密尔顿给出的。向量是一个既有大小又有方向的量，这个量本身就是个几何的概念。我们常常把它与标量（只有大小的量）相区别。抓住向量的**大小**和**方向**这两个特征，一般用一个有向线段来表示一个向量(显然，向量本身就是一个几何图形)，记为 $\overrightarrow{AB}$ 或者 $\alpha$ 。如下图：

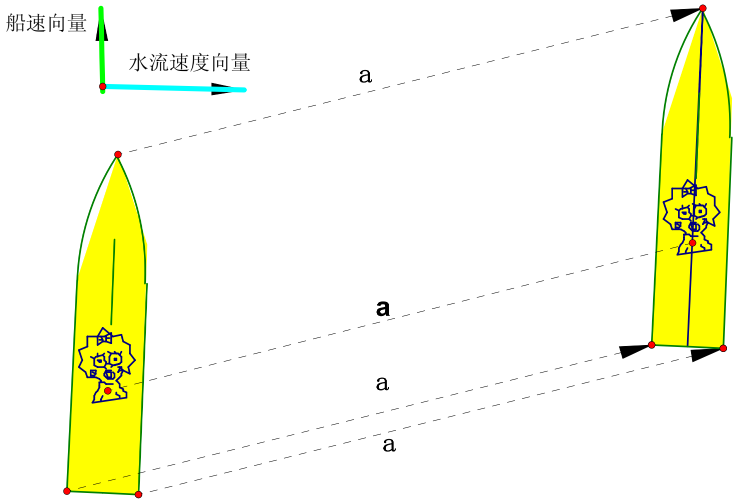



在物理学中，也把向量叫矢量，矢就是箭，向量如一根箭一样有头部和尾部，箭在空间自由的飞行中箭杆的长度不会变，这一点与向量相同；但箭在重力的作用下会改变方向，但一个确定的向量不允许改变方向，一个向量改变了方向就变成了另外一个向量了。所以向量的“飞行”称为平移，这种在一条直线上平移的向量称为自由向量（物理学中常称为滑动向量）。




沿着直线飞行的箭簇在每一时刻所表示的无数向量归属于同一个向量，这些无数的向量实际上是平行的向量。另外还有不在一条直线上的平行而相等的向量，如下的例子：

考察一个刚体的平行移动。当刚体从一个位置平行移动到另一个位置时（比如说这个刚体是麦吉小姐过河坐的小船，小船从河流的一边驶向对岸），刚体上各质点在同一时间段内有相同的位移，各点所画出的位移向量a有相同的大小和方向，他们每一个都反映了刚体位移的情况，因此刚体的平移运动可以用这些向量中的任一个来表示。基于这样的原因，凡是两个向量大小相等、方向相同的，我们就说这两个向量是相等的。因此，一个向量在保持长度和方向不变的条件下可以自由平移。如有必要，也可以将几个向量平移到同一个出发点或者坐标原点。




从上面的例子，我们感悟到自由向量为何可以是自由的。实际上，就是因为向量没有确定的位置，它们不依赖于任何坐标系而存在。因此从逻辑上看，无数的向量可能有相同的表述，所有的这些向量都互相平行，相等，并具有相同的量值和方向。

## 向量的数学表示

向量的数学表示一般是用小写的黑体字母 **a**、**b**、**c** 等表示。当手写时因为黑体的粗笔画书写不方便，因此常在字母上面加上箭头来与其它字母区别，如 $\vec{a}$、$\vec{b}$、$\vec{c}$。

以上的表示不便计算，如何对向量象数字一样进行运算呢？

因为在数学学科中，向量被处理为自由向量，为了与解析技术所用的坐标联系起来，我们把空间中所有的向量的尾部都拉到坐标原点，这样 N 维点空间可以与 N 维向量空间建立一一对应关系：N 维点空间中点(0，0，0…0)取作原点，那么每一个点都可以让一个向量和它对应，这个向量就是从坐标原点出发到这个点为止的向量。

注：

向量被看作线性空间或向量空间中的一个元素。但向量与点不同，向量表示的是两点之间的位移而不是空间中的物理位置；向量还可以确定方向，而一个点就不能。

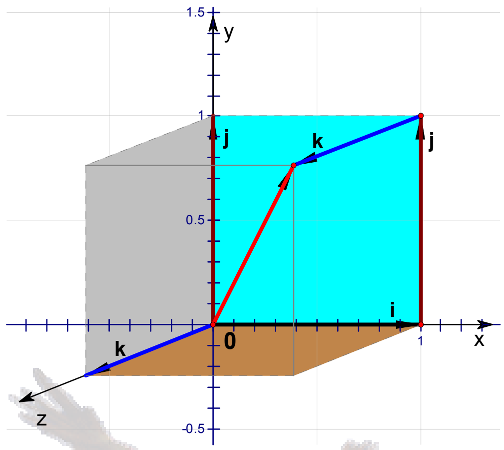
其实，一旦我们确定好一个坐标系，一个向量就与一个点相对应，而点用所谓坐标的有序数组表示的，因此我们就也可以把向量用有序数组表示。有了有序数组就可以运算了。使用有序数组或者解析式表述的向量是把以原点为起点的向量末端的坐标值表示，并把坐标值用圆括号括起来，如 $\mathbf{v}=(x,y,z)$。在这里这个有序数组 $(x,y,z)$ 称之为向量。

在二维平面上，由原点引出的向量用两个有序实数表示；在三维空间中，由三个有序数表示三维向量。那么 n 维向量就可以由以上二维和三维向量的定义推广得到。虽然 n 维向量的几何意义难以想象，但其现实意义我们还是可以把握的。比如，在三维空间中，我们只要知道一个球的球心位置和半径的大小就可以确定这个球面。把球心坐标和半径值写成有序数组，我们就得到了一个四维向量。

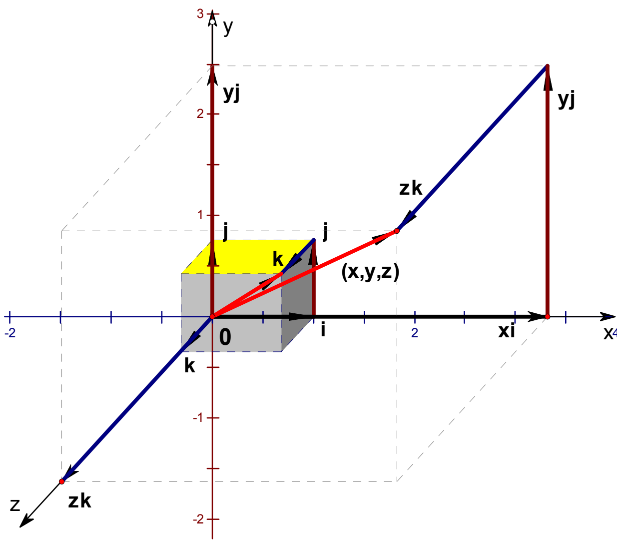
一个向量可以被分解为三个单位坐标向量的线性表示（实际上这个概念很重要，在今后的向量的运算和矩阵运算理解中起着关键作用）。例如向量 $(1, 1, 1)$ 分解如下：

$$(1, 1, 1) = (1, 0, 0) + (0, 1, 0) + (0, 0, 1) = \mathbf{i} + \mathbf{j} + \mathbf{k}$$

如下图，把单位坐标向量 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 分别首尾连接相加，就得到了 $(1, 1, 1)$ 的图像。




那么，任意一个向量 $\mathbf{v} = (x, y, z)$ 就可以表示为 $\mathbf{v} = (x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$，即单位坐标向量的线性表示。显然，分别对单位坐标向量进行缩放 $x, y, z$ 倍然后相加，就得到了这个向量 $(x, y, z)$ 的图像。和上图相似，我们就可以得到了如下的任意一个向量的分解图像。



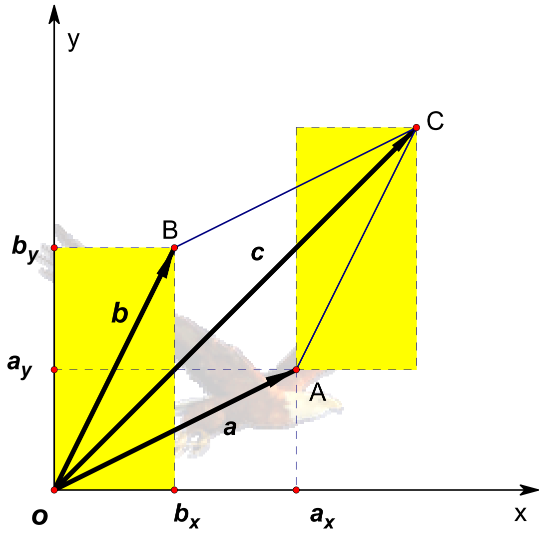

向量的运算有加法、减法和乘法，乘法有三种，但没有除法。下面我们分别介绍这些运算的细节。

## 2.2 向量的加法的几何及物理意义

设两个向量 **a** 和 **b**，它们的二维分量解析式为 $\mathbf{a}=(a_x,a_y)$， $\mathbf{b}=(b_x,b_y)$；三维分量的解析表达式为 $\mathbf{a}=(a_x,a_y,a_z)$， $\mathbf{b}=(b_x,b_y,b_z)$。则我们定义这两个向量的加法为

$\mathbf{a}+\mathbf{b}=(a_x+b_x,a_y+b_y)$，或者 $\mathbf{a}+\mathbf{b}=(a_x+b_x,a_y+b_y,a_z+b_z)$。向量加法的定义看起来很简单，就是两个向量的各分量分别对应相加形成了和向量的分量。

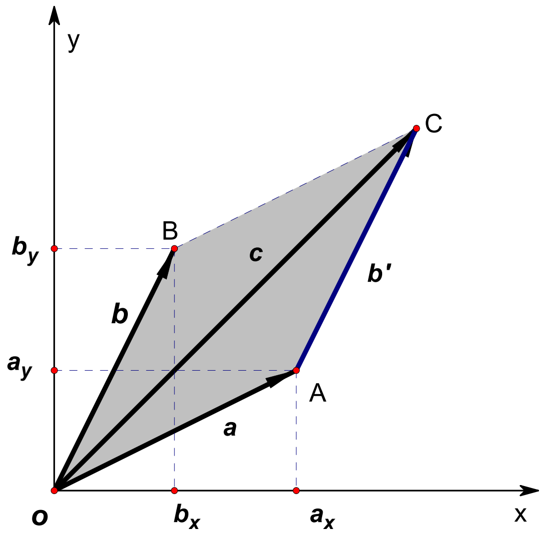
那么 $\mathbf{c}=\mathbf{a}+\mathbf{b}$ 的几何意义是什么呢？请看下面二维向量的图解。




这个图形可以这样解读，表示向量 **b** 的分量的矩形被放到表示向量 **a** 的分量的矩形上面，**a** 向量矩形的尾端 **A** 连接上 **b** 向量分量矩形的头端 **A**。叠加后矩形的顶端 C 就是和向量的尾端。

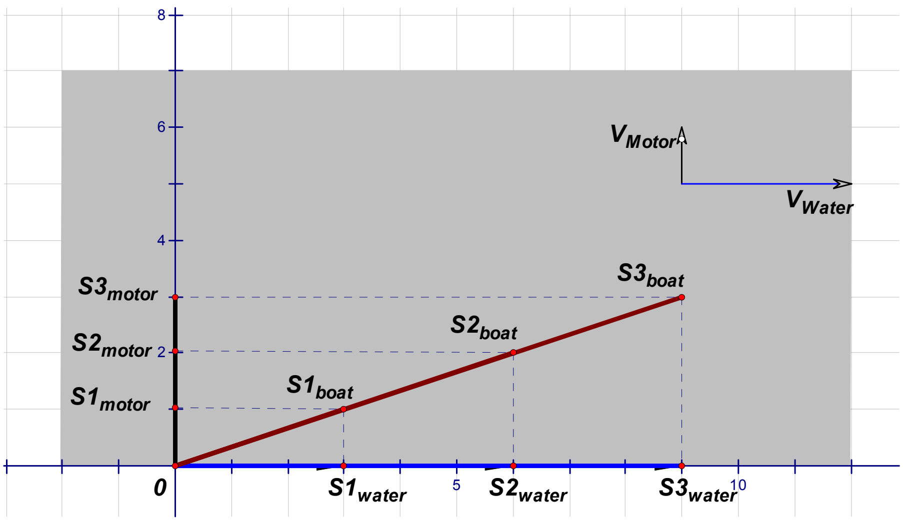
连接 BC，AC 后，就是**平行四边形的法则**的几何解释。

当然，如果把向量 **b** 平移（平行移动）到 AC 的位置，与向量 **a** 的尾部相接，就是三角形的法则。

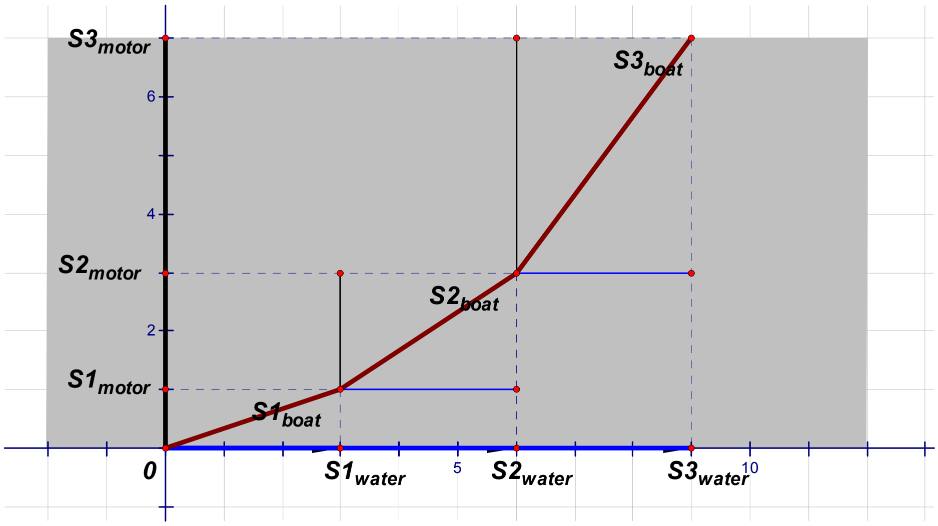



向量的所谓三角形或平行四边形法则不是人们凭空想当然的数学规定，而是从物理世界中抽象出来的向量运算法则。比如我们前面提到的船只过河的例子，船头指向的方向是船的马力驱动得到的位移为 $S_{Motor}$（不考虑水流影响），水流的方向是水的冲击力对船造成的位移 $S_{Water}$（不考虑船的马力影响），那么，实际情况是船的真正位移 $S_{Boat}$ 是一条斜线，这条斜线就是 $S_{Motor}$ 和 $S_{Water}$ 的合成。它们的合成关系就是平行四边形的关系。

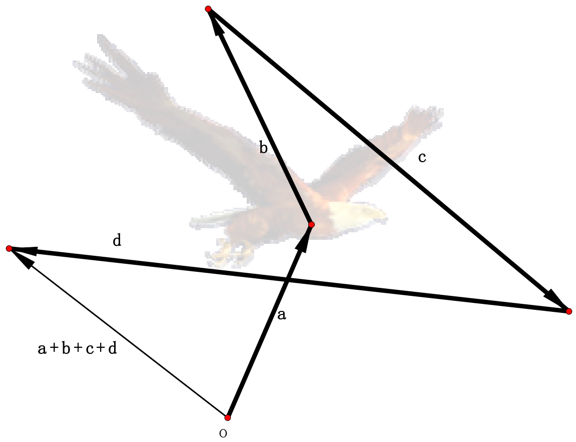
如果水的流速和船的马力不变，其中三个时刻（任意）的位移的合成图图下：

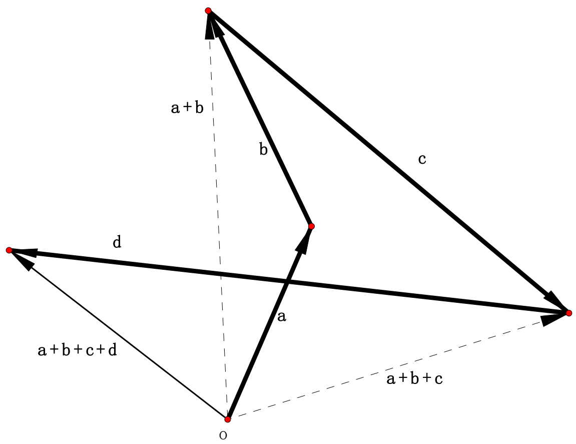



如如果水的流速不变，但在第二时刻和第三时刻船的马力逐步变大，那么三个时刻（任意）的位移的合成图图下：




两个向量的加法叫做三角形法或者平行四边形法，那么多个向量的加法同样也满足这些法则，并可以由三角形法则得到多个向量的多边形法则。下边我们画出多个向量的加法和减法的图例。



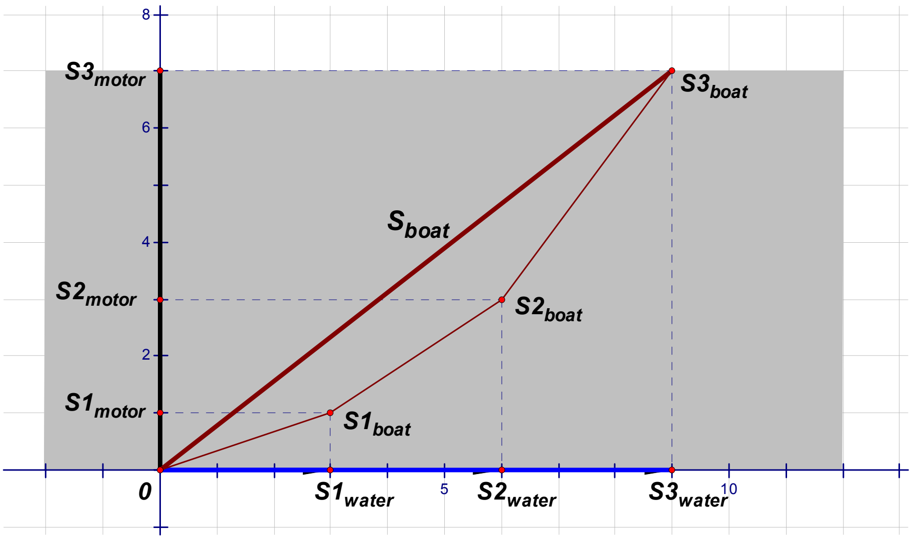




上图左是把 **a**，**b**，**c**，**d** 四个向量按照三角形法则相加的图例，在图中，我们把 **a**，**b**，**c**，**d** 四个向量依次首尾相接，直到画完所有向量，最后只是把第一个向量的尾部 o 指向最后一个向量的首部 P 画出的向量就是 4 个向量的和。这个画法可以称作多向量的多边形加法法则。

多边形法则很容易从三角形法则推导出来，上图右中虚线向量即是应用三角形的向量法则画出

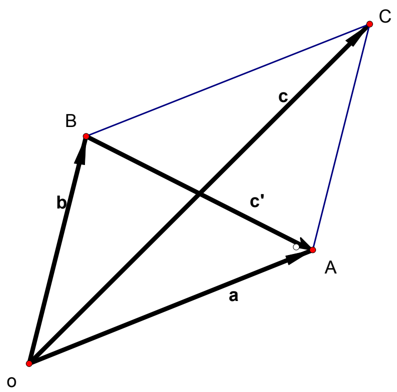
了中间向量逐次相加的结果，最后推出了左图的多边形法则。

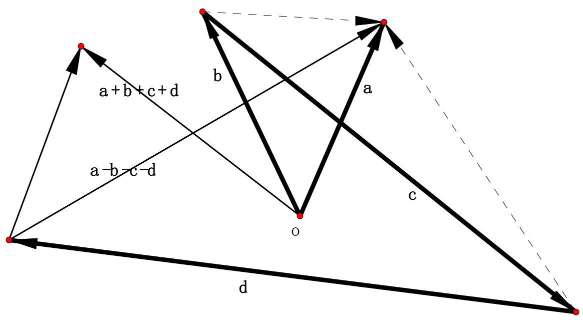
多变形法则体现在船只过河的例子就是把船的每时刻的位移进行合成，就得到如下图所示：




多向量加法的数学本质，实际上是这些向量在坐标轴上（以 0 点为坐标原点的坐标系）的投影（或坐标分量）的合成（相加或相减）后的结果。向量的更高一级的运算如点积、叉积的定义也是这个数学本质的体现。

关于减法，实际上是加法的特殊形式，是加法的逆运算。向量减法，我们可以用定义加法的方式定义减法，例如定义 $\mathbf{c}=\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b})$，只要把被减向量 b 反向后再与向量 a 相加即可。实际上在平行四边形法则中，和向量和差向量构成平行四边形的两个对角线。




三维向量的加法。

# 2.3 向量的内积的几何和物理意义

## 向量的内积的几何解释

向量的内积也叫数量积、标积、点积（点积的名称来自于内积的计算符号$\mathbf{a}\cdot\mathbf{b}$）等，都是一个意思，就是内积的结果是个数量或者标量。内积的定义有两个，下面我们把它们列举出来并探讨一下它们的关系。

$$\mathbf{a}\cdot\mathbf{b}=ab\cos\theta \text{ ........................。。 1}$$

$$\mathbf{a}\cdot\mathbf{b}=a_xb_x+a_yb_y+a_zb_z \text{ ...........。 2}$$

公式 1 是说，向量 **a** 和 **b** 的长度之积再乘以它们之间的夹角的余弦；

公式 2 的意思是向量 **a** 和 **b** 的坐标分量分别对应乘积的和。

定义内积有很多好处，除了物理上的直接应用外，至少我们可以应用这个定义（公式 2）去计算一个向量的长度（在已知它的坐标时）。比如我们求向量 **a** 的长度：

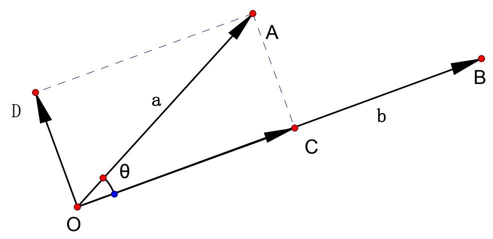
$$a=\sqrt{\mathbf{a}\cdot\mathbf{a}}=\sqrt{a_xa_x+a_ya_y+a_za_z}=\sqrt{a_x^2+a_y^2+a_z^2}\text{ 。}$$

这两个公式有关系吗？当然有：

假设我们选一个这样的坐标系，x 轴沿向量 **a** 的方向，那么 $a_x=a$，$a_y=a_z=0$，则公式

$\mathbf{a}\cdot\mathbf{b}=a_xb_x+a_yb_y+a_zb_z=ab_x$，

$ab_x$ 就是 **a** 的长度乘以 **b** 在 **a** 方向上的分量，这个分量就是 **b** 在 **a** 上的投影，因此公式 $\mathbf{a}\cdot\mathbf{b}=ab\cos\theta$ 得证（如果它对一个坐标系成立，则对所有的坐标系都成立）。

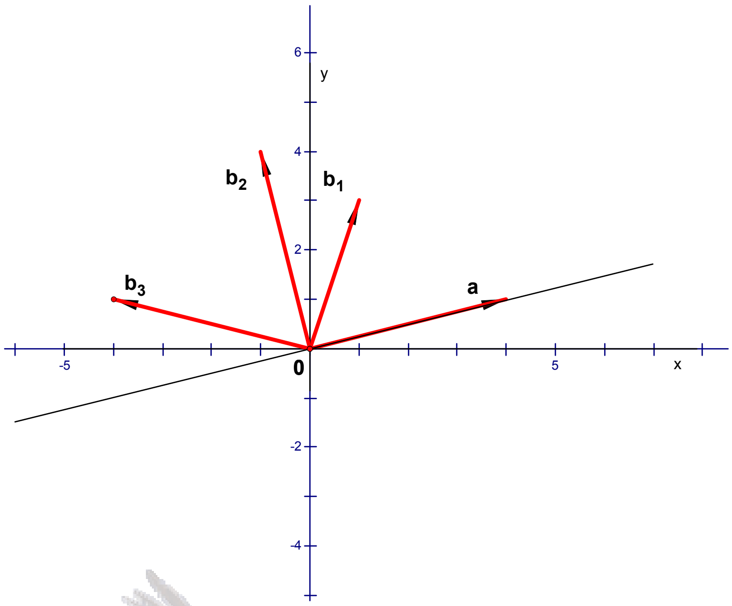



因此，向量内积的几何解释就是一个向量在另一个向量上的投影的积，也就是同方向的积。

特别的，如果一个向量如 a 是某个坐标轴的单位坐标向量，那么，两个向量的内积$\mathbf{a}\cdot\mathbf{b}$就是向量 b 在此坐标轴上的坐标值。这个结论非常重要，这是傅立叶分析的理论基础。

另外，对两个向量内积的投影的几何意义可以得到其他的几何解释，这些解释在应用上就显得比较直观。比如，从内积数值上我们可以看出两个向量的在方向上的接近程度。当内积值为正值时，两个向量大致指向相同的方向（方向夹角小于 90 度）；当内积值为负值时，两个向量大致指向相反

的方向（方向角大于 90 度）；当内积值为 0 时，两个向量互相垂直（这个性质经常在向量几何中作为判断直线与直线是否垂直）。笼统说来，内积值越大，两个向量的在方向上的就越接近，内积值越小，两个向量的在方向上的就越相反。




上图中，向量 a 与向量 b1 方向相近，内积为正；a 与 b2 方向垂直，内积为 0；a 与 b3 方向大致相反，内积为负。

## 向量的内积的物理解释

向量内积的物理应用或者说物理意义很多，生活中我们也需要内积计算。比如我上周购买的食物的价格向量是 P=（蔬菜 2 元/斤，大米 1。5 元/斤，猪肉 5 元/斤，啤酒 3 元/瓶），消耗的数量向量为 d=（3。5 斤，5 斤，2 斤，3 瓶）；那么我上周的饮食消费就是向量 P 和 d 的内积：

$$\mathbf{p}\cdot\mathbf{d} = (2,\ 1.5,\ 5,\ 3)\cdot(3.5,\ 5,\ 2,\ 3) = 7 + 7.5 + 10 + 9 = 33.5\text{元。}$$

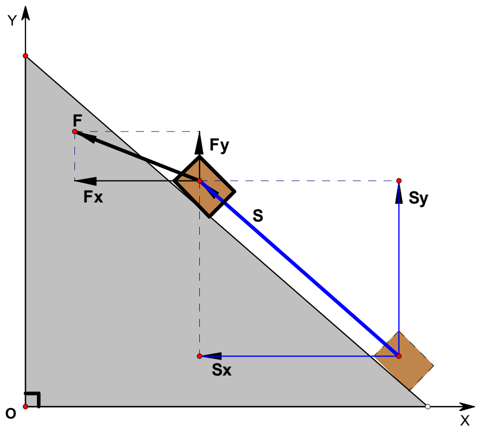
另外内积的一个经典例子就是当一个物体从某处被拉到另一处的所做的功，下面我们把这个做功的图画出来来印证以上内积两大公式的一致性。

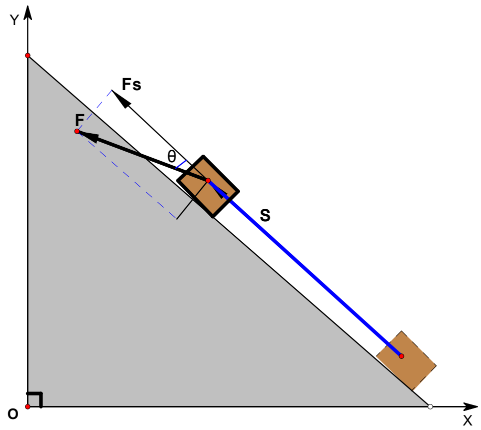
我们假设是在一个斜坡上用力 **F** 斜上拉一个物体，位移为 **S**（没有重力的情况下）。那么这个力 **F** 所作的功为（分量的分解见图左）：

$$W = \mathbf{F}x\mathbf{S}x + \mathbf{F}y\mathbf{S}y$$

另外，我们也可以把力 **F** 沿着 **S** 的方向和垂直 **S** 的方向（按照图右所示）进行分解，那么这个力 **F** 所作的功又可表示为

$$W = \mathbf{F}s\mathbf{S} = \mathbf{FS}\cos\theta$$



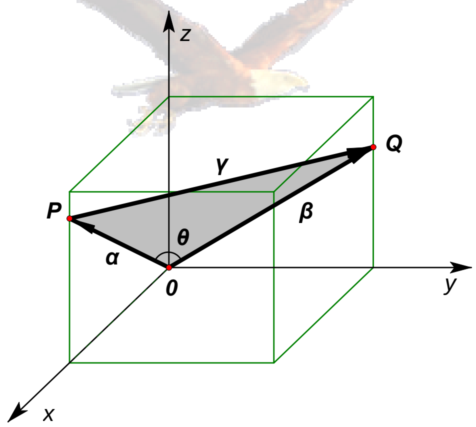




由此，我们从物理原理上印证了内积两大公式的一致性。

## 向量的内积两个定义的关系的数学推导

下面我们对内积的定义进行推导来帮助大家确信这种关系：

设 O，P，Q 为空间的三点，令 $\alpha=\overline{OP}$，$\beta=\overline{OQ}$，$\gamma=\overline{PQ}$，α，β 夹角为 θ，如图。




内积概念直观图

由余弦定理知，$|\gamma|^2=|\alpha|^2+|\beta|^2-2|\alpha||\beta|\cos\theta$ （1）

再设

$$\alpha=(x_1,y_1,z_1)$$
$$\beta=(x_2,y_2,z_2)$$
$$\gamma=(x_1\text{-}x_2,y_1\text{-}y_2,z_1\text{-}z_2)$$

代入上（1）式得：

$$\left(x_1-x_2\right)^2+\left(y_1-y_2\right)^2+\left(z_1-z_2\right)^2=x_1{}^2+y_1{}^2+z_1{}^2+x_2{}^2+y_2{}^2+z_2{}^2-2|\alpha||\beta|\cos\theta$$

即 $x_1x_2+y_1y_2+z_1z_2=|\alpha||\beta|\cos\theta$ ......（2）

这样，对向量 α，β 就有唯一确定的实数 $|\alpha||\beta|\cos\theta$ 与之对应。即得到以 α，β 为自变量的二元函数，记作（α，β），称作 α，β 的内积。

## 向量的内积与正交变换

**定理 1** 设 $\sigma$ 是 $V$ (欧式空间)的一个变换。若对任意向量 $\mathbf{a},\mathbf{b}\in V$，都有

$$|\sigma(\mathbf{a})+\sigma(\mathbf{b})|=|\mathbf{a}+\mathbf{b}|$$

则 $\sigma$ 是 $V$ 的一个正交变换。

这个定理我们不证明，其几何意义如何理解呢？几何意义是“保持以 $V$ 中任意两个向量为邻边的平行四边形的对角线之长不变”。图示如下：

**定理 2** 设 σ 是 V 的一个变换。如果 σ 既是保长度变换又是保夹角变换，那么 σ 必为正交变换。。证明如下：

设 $\xi,\eta\in \mathrm{V}$，当 $\xi\neq 0\neq\eta$ 时，由

$$|\sigma(\xi)|=|\xi|,|\sigma(\eta)|=|\eta|,$$

$$\frac{\langle\sigma(\xi),\sigma(\eta)\rangle}{|\sigma(\xi)||\sigma(\eta)|}=\frac{\langle\xi,\eta\rangle}{|\xi||\eta|},$$

得

$$\langle\sigma(\xi),\sigma(\eta)\rangle=\langle\xi,\eta\rangle$$

即保持了变换的内积不变，因而是线性的，或正交变换。

如果 V 的一个变换 σ 既是保长度变换又是保夹角变换(即 σ 保持 V 中任二非零向量间的夹角不变)，那么 σ 就应该保对角线长，从而 σ 是一个正交变换，以上证明事实正是这样。

正交（orthogonal）是直观概念中垂直的推广。作为一个形容词，只有在一个确定的内积空间中才有意义。若内积空间中两向量的内积为 0，则称它们是正交的。如果能够定义向量间的夹角，则正交可以直观的理解为垂直。物理中：运动的独立性，也可以用正交来解释。

# 2.4 向量的外积的几何和物理意义

## 叉积的定义及其几何解释

向量的外积（Cross product）也译作叉积（同点积类似，此名称也来自于外积的计算符号 **a** x **b**），因为叉积会产生。新的一维，两个向量确定了一个二维的平面，叉积又会产生垂直于这个平面的向量，因此**外积的概念只能应用于 3 维和三维以上的向量空间。**

叉积的定义也有两个，下面我们把它们列举出来并探讨一下它们的关系。

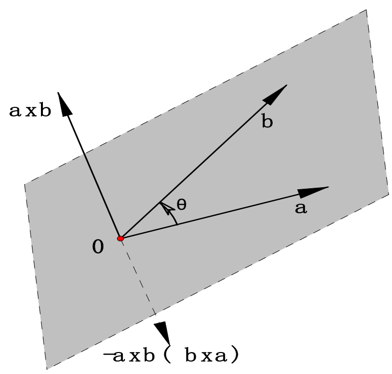
设三维空间中的两个向量为 $\mathbf{a}=(a_x,a_y,a_z)$，$\mathbf{b}=(b_x,b_y,b_z)$，则

$$\mathbf{a}\times\mathbf{b}=(a_yb_z-a_zb_y,a_zb_x-a_xb_z,a_xb_y-a_yb_x)\ldots\ldots1$$

$\mathbf{a}\times\mathbf{b}\;=ab\sin\theta N$（其中 N 是垂直于 **a** 和 **b** 展成的平面的单位向量）……2

公式 1 是用向量的三维坐标值表述的解析式。这个公式表面含义是叉积 c 的 x 轴的分量是 $a_yb_z-a_zb_y$，y 轴的分量是 $a_zb_x-a_xb_z$，z 轴的分量是 $a_xb_y-a_yb_x$。显然，叉积 c 的 x 方向的分量是向量 a 和 b 在 yoz 平面上的分量计算出来；类似的，c 的 y 方向的分量是向量 a 和 b 在 xoz 平面上的分量计算出来，c 的 z 方向的分量是向量 a 和 b 在 xoy 平面上的分量计算出来的。实际上，叉积的这三个分向量分别又是三个叉积向量。为什么是这样的呢？后面的物理意义解析会告诉你原因。

公式 2 是叉积几何意义的定义式。$\mathbf{a}\times\mathbf{b}$ 为一个新生成的向量，这个向量垂直于 **a** 和 **b** 展成的平面（图中的虚线平行四边形，由线段 0**a** 和 0**b** 所确定的平面）；同样向量 $\mathbf{b}\times\mathbf{a}$ 也垂直这个平面，但方向与 $\mathbf{a}\times\mathbf{b}$ 所指的方向相反，即 $\mathbf{b}\times\mathbf{a}=-\mathbf{a}\times\mathbf{b}$。




我们可以用右手法则来帮助记忆这个定义，右手的大拇指指向向量 **a** x **b** 的方向，则弯曲的四指则指向向量 **a** 和 **b** 叉乘的顺序：从向量 **a** 沿着 **a** 和 **b** 间较小夹角转向向量 **b**。反过来，如果已知两个向量进行叉乘，那么用右手法则则可以知道这两个向量叉乘出来的向量的方向。

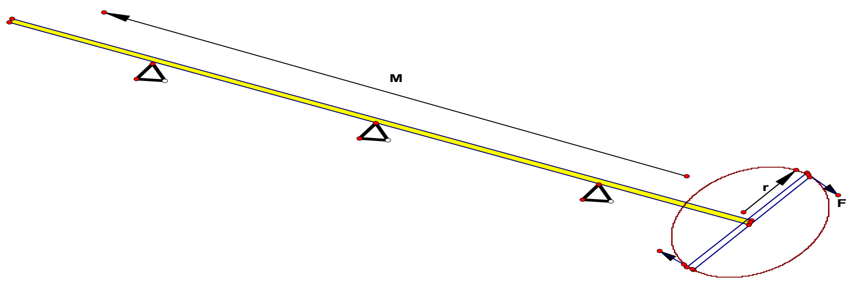
## 叉积的物理意义

叉乘的定义看起来有点怪，大家可能感觉到，叉积向量好像不是太真实似的，特别是方向定义的显然是人为的。实际上叉积这种向量与前面介绍的向量确实不同，所以在物理学中又被称为赝向量或轴向量。

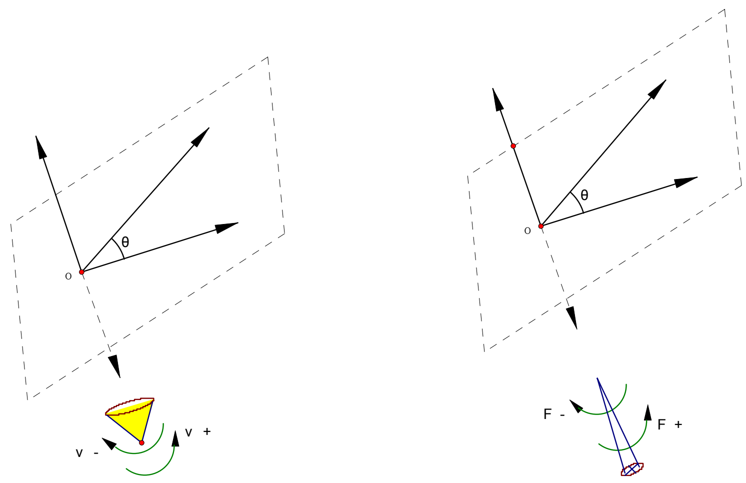
但这个定义也是从物理应用方面得来的。举一个例子：知道陀螺的原理吗？高速旋转的陀螺会定向。陀螺所定义的方向就是矢径向量$\vec{r}$和线速度$\vec{v}$叉乘结果角速度$\vec{\omega}$方向。类似的一个例子是螺钉，螺钉只要左右向旋转即可在螺孔中前进或者后退。用螺丝刀把这棵螺钉按照 F+的方向右旋，那么旋转时的扭力向量$\vec{F}$和矢径向量$\vec{r}$这两个叉乘的结果即是力矩$\vec{M}$的方向，这棵螺钉就会沿着力矩$\vec{M}$在螺母孔内前进，反方向就会改变叉积的方向进而退出螺孔（右螺旋螺钉）。也就是力矩或叉乘向量的方向就是螺钉的螺旋前进的方向，这个方向垂直于螺丝刀口和扭力的方向，也就是垂直于被叉积的两个向量的方向。

力矩就是向量的叉积。还有点疑惧？好，弄个夸张一点的。我们把螺钉的原理稍微改变一下：假如有一个 100 米长的细钢棒（好长），钢棒架在几个支架上，钢棒一端装有摇臂，当有人用摇臂扭转钢棒时，这个扭转的力（就是力矩）会沿着这个长长的钢棒一直延伸到钢棒的尾端，并且整个钢棒上都有扭转的力存在，无论我们碰触钢棒的任何部位都会感知到这个力矩的存在。这个扭转的力是多大呢？如果摇动的人用力越大，摇臂越长，这个力矩就会越大，我们就越难用手抓停它。

呵呵，力矩沿着 100 米长的、与摇臂和摇动的力垂直的方向，无处不在！




力矩的方向



另外一个例子就是我们经常骑的自行车，车子静止的时候我们在车上会摔下来，一旦骑行起来车子就会平稳而不会左右倾倒，这也是叉积的功劳（与陀螺的原理相同）。

下面我们看一看叉积解析式的物理意义的分解。同样，我们也举扭矩的例子。这里我们再次把叉积解析公式重新列在这里：

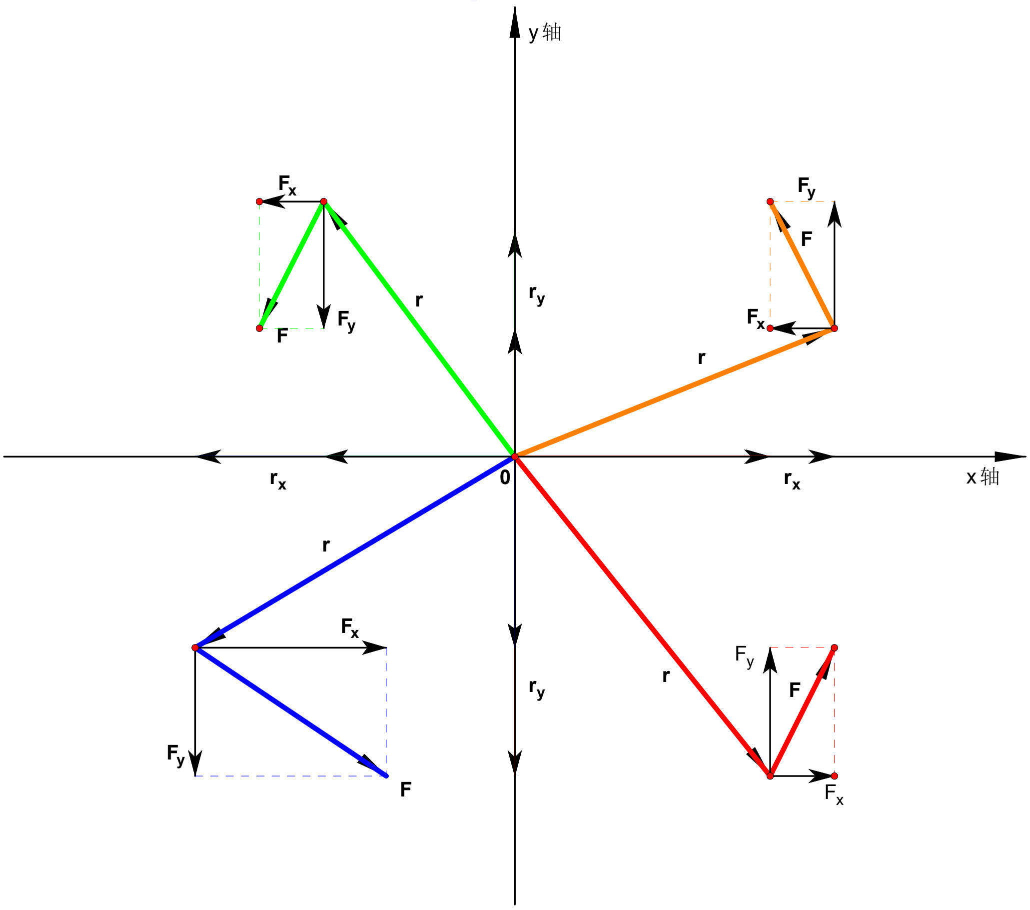
$$\mathbf{c}=\mathbf{a}\text{x }\mathbf{b}=(a_yb_z-a_zb_y, a_zb_x-a_xb_z, a_xb_y-a_yb_x)$$

前边讲过， 一个向量可以分解为沿着 x、y、z 轴的分向量，或者讲一个三维向量可以看作是三个分别与坐标轴同向向量之和。即：

$$\mathbf{v}=(x,y,z)=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$$

在这里我们同样可以认为（$a_yb_z-a_zb_y$）$\mathbf{i}$ 是 x 方向的向量，（$a_zb_x-a_xb_z$）$\mathbf{j}$ 是 y 方向的向量，（$a_xb_y-a_yb_x$）$\mathbf{k}$ 是 z 方向的向量。

前面讲叉积的这三个分向量分别又是三个叉积，从何讲起？下面我们以叉积的 z 轴分量的（$a_xb_y-a_yb_x$）$\mathbf{k}$ 来比对物理上力矩的概念。假设，我们的书面为 x0y 平面，z 轴垂直书面并指向我们，以这个三维坐标系的原点 0 为转轴，在力 $\mathbf{F}$ 的作用下逆时针转动，转轴向量为 $\mathbf{r}$。$\mathbf{r}\times\mathbf{F}$ 的方向指向我们。下面来看看这个图解：



图中在 x0y 坐标系的四个象限都画出了力矩的例子，并且把叉乘的两个向量 **r** 和 **F** 都进行了分解。**F** 分解为 $\mathbf{F_x}$ 和 $\mathbf{F}_y$，**r** 也分解 $\mathbf{r_x}$ 和 $\mathbf{r_y}$。分解后我们同样对分解的分向量 $\mathbf{r_x}$、$\mathbf{r_y}$、$\mathbf{F_x}$ 和 $\mathbf{F}_y$ 进行求力矩的叉积 $\mathbf{M} \ =\mathbf{r}_x\times\mathbf{F}_y+\mathbf{r}_y\times\mathbf{F}_x$（其它的分向量因为方向相同或者相反：$\mathbf{r}_x\times\mathbf{F}_x=\mathbf{r}_y\times\mathbf{F}_y=0$，因此忽略不写了）。在初中我们就知道力矩等于力乘以力臂，力臂与扭力垂直。因此，我们分别得出第一象限的力矩是 $\mathbf{M_1}=\mathbf{r}_x\times\mathbf{F}_y+\mathbf{r}_y\times\mathbf{F}_x=xF_y\mathbf{k}+y(-F_x)\mathbf{k}=(xF_y-yF_x)\mathbf{k}$。这里，**k** 是指向 z 坐标方向的单位向量。

类似地，其它三个象限的力矩分别是：

$$\mathbf{M}_2=\mathbf{r}_x\times\mathbf{F}_y+\mathbf{r}_y\times\mathbf{F}_x=(-x)(-F_y)\mathbf{k}+y(-F_x)\mathbf{k}=(xF_y-yF_x)\mathbf{k}$$
$$\mathbf{M}_3=\mathbf{r}_x\times\mathbf{F}_y+\mathbf{r}_y\times\mathbf{F}_x=(-x)(-F_y)\mathbf{k}+(-y)F_x\mathbf{k}=(xF_y-yF_x)\mathbf{k}$$
$$\mathbf{M}_4=\mathbf{r}_x\times\mathbf{F}_y+\mathbf{r}_y\times\mathbf{F}_x=xF_y\mathbf{k}+(-y)F_x\mathbf{k}=(xF_y-yF_x)\mathbf{k}$$

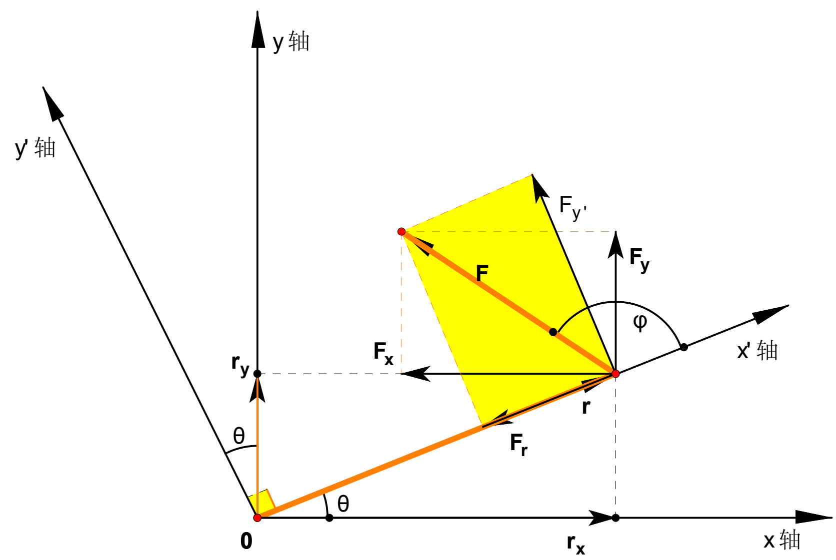
我们看，四个象限的力矩表达式相同，都是 $(xF_y-yF_x)\mathbf{k}$！

在这个向量 $(xF_y-yF_x)\mathbf{k}$ 中，$(x,y)$ 是矢径向量 **r** 的坐标，$(F_x,F_y)$ 是扭力向量 **F** 的坐标；在 xoy 的平面上，向量 **r** 和向量 **F** 的叉积的大小等于 $xF_y-yF_x$，方向 **k** 指向 z 轴方向。

显然，我们把向量 **r** 和向量 **F** 改写成通用向量 **a** 和 **b**，这个结果就变成了（$a_xb_y-a_yb_x$）**k**。

当然，我们同样可以推论出 x 轴、y 轴方向的叉积表达式如前所述。

实际上，我们把坐标系重新选择或者把坐标系右旋一个 θ 角度，可以得到叉积定义的另一个公式 **a** x **b** $=ab\sin\theta N$。



如图所示，两个向量 **r** 和 **F** 在 xoy 坐标系中 **F** 分解为 $\mathbf{F_x}$ 和 $\mathbf{F}_y$，**r** 也分解 $\mathbf{r_x}$ 和 $\mathbf{r_y}$，如果把 xoy 坐标轴右旋一个 θ 角，变为 x'0y'坐标系，刚好使 x'轴与向量 **r** 重合。

显然，在新的坐标系下，**r** 不必分解分量了；**F** 只需在为分解为 $\mathbf{F}_{x'}$ 和 $\mathbf{F}_{y'}$，则新坐标系下的叉积（z 方向的力矩 **M**）

$$\mathbf{M}_z = \mathbf{r}\times\mathbf{F} = \mathbf{r}\times(\mathbf{F}_{x'}+\mathbf{F}_{y'}) = \mathbf{r}\times\mathbf{F}_{x'}+\mathbf{r}\times\mathbf{F}_{y'} = \mathbf{r}\times\mathbf{F}_{y'}$$

又因为 x' 轴与 **r** 重合，且 **F** 与 **r** 的夹角为 φ，因此上式继续等于：

$$\mathbf{M}_z = \mathbf{r}\times\mathbf{F}_{x'} = \mathbf{r}\times\mathbf{F}_r = -rF\cos(\pi-\phi)\mathbf{k} = rF\sin(\phi)\mathbf{k}$$

如果我们把向量 **r** 和 **F** 改写成通用向量 **a** 和 **b**，且不是强调在 xyz 三维坐标下，那么 z 向的单位向量 **K** 可以写成与叉乘的两个向量垂直且满足右手系的 N。至此得到叉乘的第二个公式：**a x b** $=ab\sin\theta N$。

## 2.5 向量混合运算的几何意义

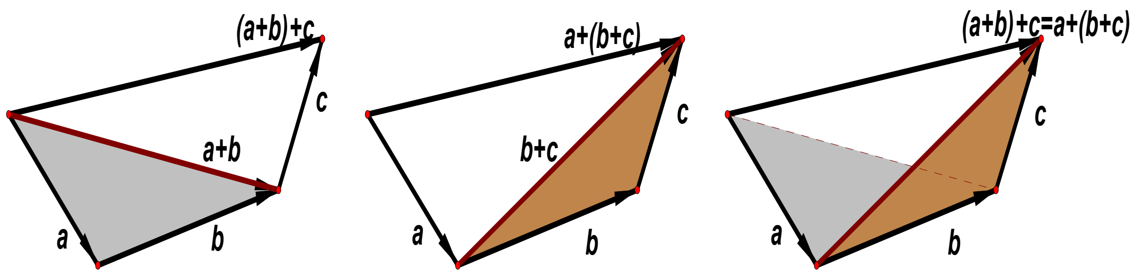
我们所讨论的向量的混合运算包括向量加法和乘法的混合运算，仔细的研究一些混合运算的几何意义有助于理解向量的几何本质。

### 向量加法的结合律的几何解释

三个向量加法的结合律为

$$(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$$

其图解是显然的，第一幅图给出了 $(\mathbf{a}+\mathbf{b})+\mathbf{c}$ 的加法图解，第二幅图给出了 $\mathbf{a}+(\mathbf{b}+\mathbf{c})$ 的图解，第三幅图把两者叠放在一起，显示了两个加法有相同的结果。图形如下：

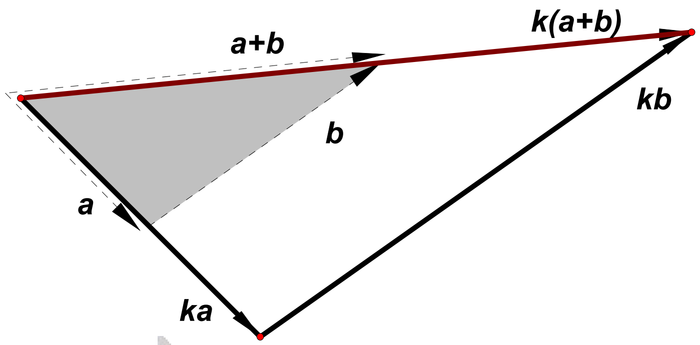


## 向量数乘的分配律的几何解释

数乘两个向量和的分配率为

$$k(\mathbf{a}+\mathbf{b}) = k\mathbf{b}+k\mathbf{c}$$

其图解是显然的，图中设数量 k 大于 1，$\mathbf{a}+\mathbf{b}$ 的加法三角形和 $k\mathbf{a}+k\mathbf{b}$ 的加法三角形是两个相似三角形，因而得到图形如下：

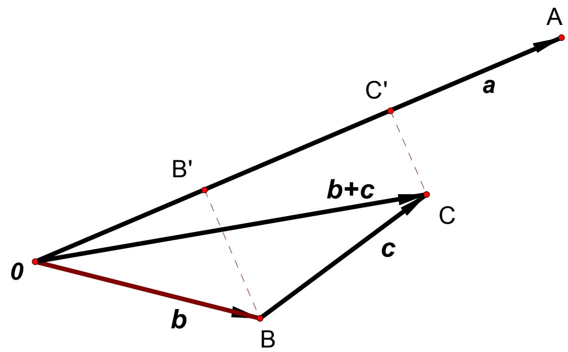



## 向量点积的分配律的几何解释

向量点积的分配律为

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}$$

如下图，$0B'$ 为向量 $\mathbf{b}$ 在向量 $\mathbf{a}$ 上的投影，$B'C'$ 为向量 $\mathbf{c}$ 在向量 $\mathbf{a}$ 上的投影，$0C'$ 为向量 $\mathbf{b}+\mathbf{c}$ 在向量 $\mathbf{a}$ 上的投影。




由图有：

$$\mathbf{a}\cdot\mathbf{b} = 0A\cdot 0B' \ , \qquad \mathbf{a}\cdot\mathbf{c} = 0A\cdot B'C'$$

于是得：

$$\mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}=0A\cdot 0B'+0A\cdot B'C'=0A\cdot(0B'+B'C')=0A\cdot 0C'$$

而又有 $\mathbf{a}\cdot(\mathbf{b}+\mathbf{c})=0A\cdot 0C'$，因此有分配律成立。

## 向量叉积的分配律的几何解释 1

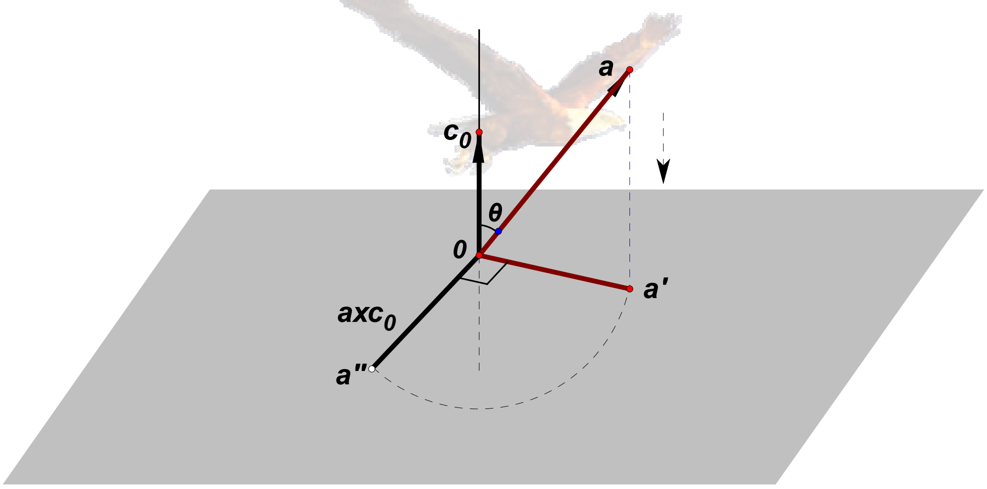
向量的分配律表述如下的等式：

$$(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$$

对于理解这个分配律的几何解释，我们可以有两个图解的解释。

先说一个较熟悉的几何解释。

在详述几何解释之前，我们先介绍一个引理的几何意义。就是对于单位向量 $\mathbf{c}^0$，其与另一个向量 $\mathbf{a}$ 的叉积 $\mathbf{a}\times\mathbf{c}^0$ 的几何图形如何作图求出。




如上图所示，先过向量单位向量 $\mathbf{c}^0$ 的起始 0 点构造一个垂直于它的平面，然后把向量 $\mathbf{a}$ 投影到这个平面上得到向量 $\mathbf{a}'$，接着把这个向量 $\mathbf{a}'$ 在平面上绕 0 点顺时针旋转 $\pi/2$ 角度。容易知道，旋转后的向量 $\mathbf{a}''$ 同时垂直于 $\mathbf{a}$ 和 $\mathbf{c}^0$。下面我们将说明 $\mathbf{a}''$ 恰好等于 $\mathbf{a}\times\mathbf{c}^0$。

因为 $\mathbf{a}$ 和 $\mathbf{c}^0$ 的叉积向量的长度有

$$|\mathbf{a}\times\mathbf{c}^0|=|\mathbf{a}||\mathbf{c}^0|\sin\theta=|\mathbf{a}|\sin\theta=|\mathbf{a}'|$$

也就是说，任意一个向量 $\mathbf{a}$ 和单位向量 $\mathbf{c}^0$ 叉积的长度 $|\mathbf{a}\times\mathbf{c}^0|$ 等于向量 $\mathbf{a}$ 向平面（垂直于 $\mathbf{c}^0$）

的投影向量长度$|\mathbf{a}'|$。而$|\mathbf{a}''|$又是$|\mathbf{a}'|$旋转得到的。因此$|\mathbf{a}''|=|\mathbf{a}'|$。

另外，前面我们说了$\mathbf{a}''$同时垂直于$\mathbf{a}$和$\mathbf{c}^0$，综合得知 $\mathbf{a}''=\mathbf{a}\times\mathbf{c}^0$。

引理说完了。下面开始说正体了。

这里有三个向量$\mathbf{a}$、$\mathbf{b}$和$\mathbf{a}+\mathbf{b}$，那么这三个向量分别与单位向量$\mathbf{c}^0$叉积的向量图形可以模仿上述的过程得到。如下图所示我们得到了三个叉积向量 $\mathbf{a}''=\mathbf{a}\times\mathbf{c}^0$、$\mathbf{b}''=\mathbf{b}\times\mathbf{c}^0$ 和 $(\mathbf{a}+\mathbf{b})''=(\mathbf{a}+\mathbf{b})\times\mathbf{c}^0$。

另外，请注意，三个向量因为有相加的关系，构成了一个平行四边形。这个平行四边形在投影下不会改变其边的平行或连接的关系。因此投影后的三个向量$\mathbf{a}'$、$\mathbf{b}'$和$(\mathbf{a}+\mathbf{b})'$在平面上仍然构成一个新的平行四边形。因此符合平行四边形法则的加法规律，故有$\mathbf{a}'+\mathbf{b}'=(\mathbf{a}+\mathbf{b})'$。

接着来，呵呵。投影平行四边形整体顺时针旋转$90^0$后仍然是个平行四边形，因此三个向量$\mathbf{a}''$、$\mathbf{b}''$和$(\mathbf{a}+\mathbf{b})''$仍然符合平行四边形法则，因此有$\mathbf{a}''+\mathbf{b}''=(\mathbf{a}+\mathbf{b})''$。

把 $\mathbf{a}''=\mathbf{a}\times\mathbf{c}^0$，$\mathbf{b}''=\mathbf{b}\times\mathbf{c}^0$， $(\mathbf{a}+\mathbf{b})''=(\mathbf{a}+\mathbf{b})\times\mathbf{c}^0$代到上述等式，得：

$$\mathbf{a}\times\mathbf{c}^0+\mathbf{b}\times\mathbf{c}^0=(\mathbf{a}+\mathbf{b})\times\mathbf{c}^0$$

好，到此问题基本解决。追后临门一脚是，把单位向量$\mathbf{c}^0$伸缩为一般的向量$\mathbf{c}=|\mathbf{c}|\mathbf{c}^0$。即对等式两边同乘以一个常数$|\mathbf{c}|$，得

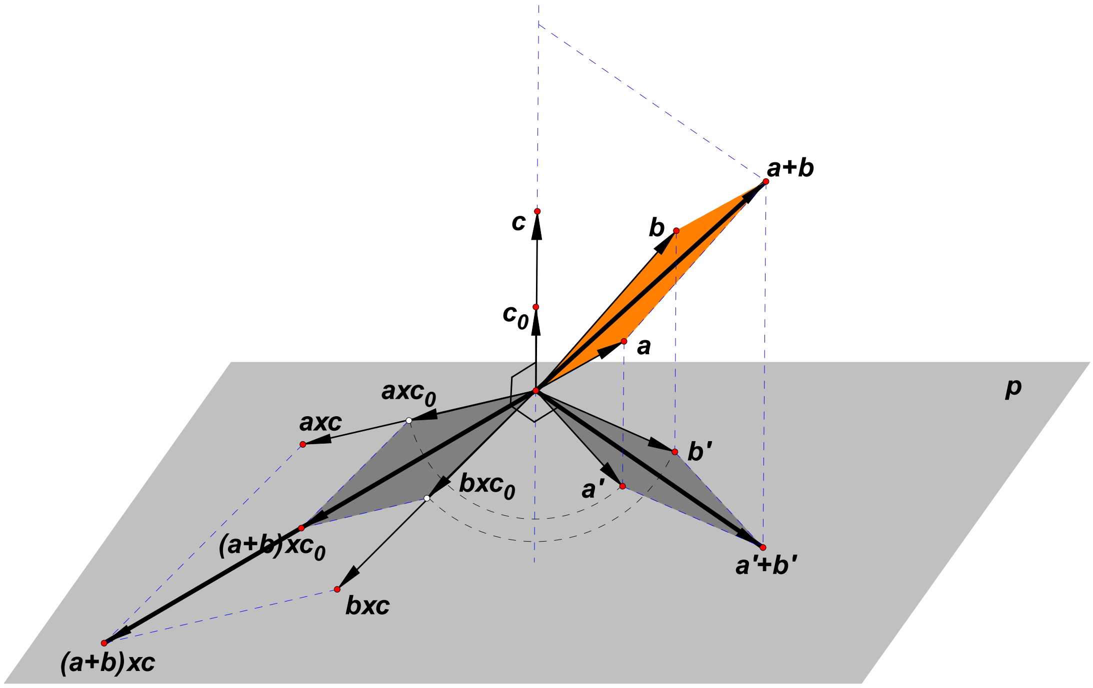
$$\mathbf{a}\times\mathbf{c}^0|\mathbf{c}|+\mathbf{b}\times\mathbf{c}^0|\mathbf{c}|=(\mathbf{a}+\mathbf{b})\times\mathbf{c}^0|\mathbf{c}|$$

即

$$\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}=(\mathbf{a}+\mathbf{b})\times\mathbf{c}$$

倒换等式两边的项即结论：

$$(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$$




总结一下：

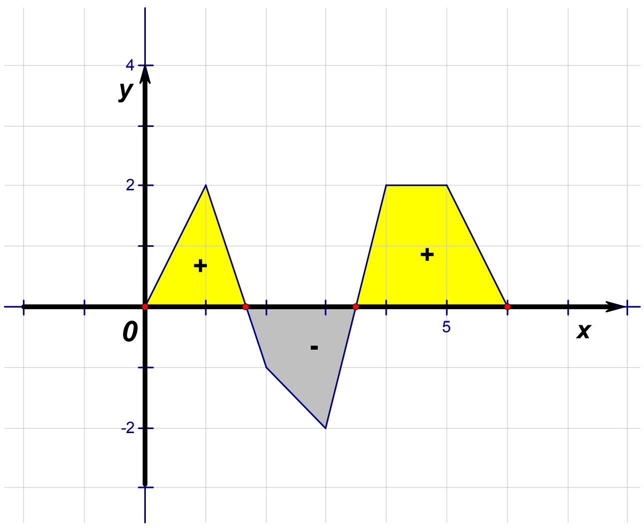
上述过程表述为向量的“一投一转，再加一伸”。

## 向量叉积的分配律的几何解释 2

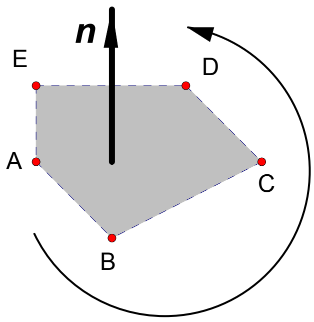
“一投一转，再加一伸”，叉积分配率整得忒麻烦？再来一个简单点的几何解释。

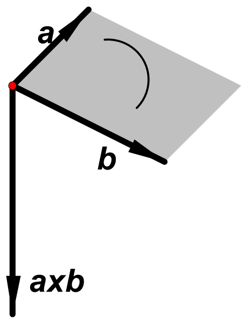
这个简洁的叉积分配图解需要有方向的面积的概念。我们首先介绍一下有向面积。

有方向线段被用来作为向量的图形。那么面积也是有方向的。在微积分对 x-y 平面上的曲线与 x 座标轴围成的面积积分中，X 座标轴上方的面积积分值为正，轴下方的面积积分值为负。这实际上也是面积有方向的表现。



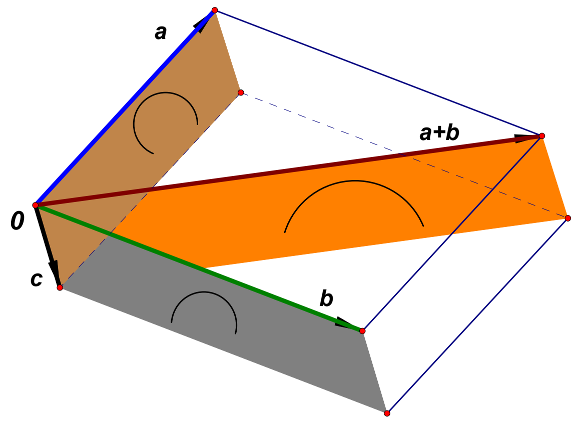
一个有边界的平面（如图左），它的大小有它的面积决定，它的方向由它在空间的法线的方向来确定，因此有向面积也可以用向量的办法来完全刻画：向量的方向就是法线的方向，向量的长度正比于它的面积。按照右手规则，如果面积周线的回转方向是 ABCDE，那么法线的正向，也就是代表这面积的方向就是向上的。根据这种说法，我们可以这样说，向量不但可以表示一个有方向的线段，而且也可以表示一个有方向的面积。反过来讲，**一个有向线段（一定长度的箭头）被用来作为向量的几何图形（这是几乎所有数学书的做法），而一个有方向的面积也可以表示一个向量**。为了方便，我们可以称前者为线向量，称后者为面向量。

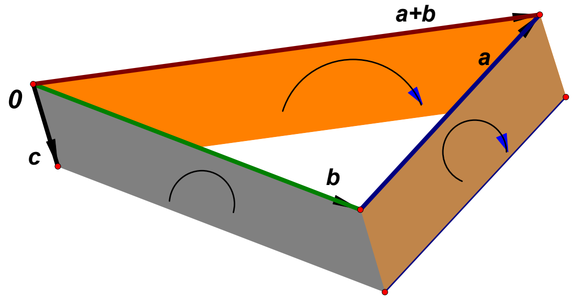


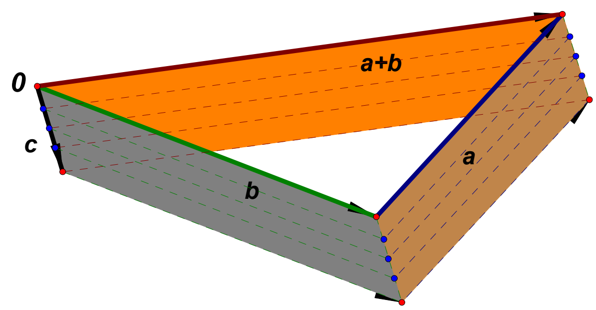

向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的叉积就是一个有向面积的例子（如图右），以向量 $\mathbf{a}$ 和 $\mathbf{b}$ 为边的平行四边形的有向面积是用 $\mathbf{a}\times\mathbf{b}$ 来表示的，因为所有向量都被处理成线向量，因此 $\mathbf{a}\times\mathbf{b}$ 也被刻画成有向线段的图形，这个有向线段垂直于被代表的面，线段长度等于 $\mathbf{a}$ 和 $\mathbf{b}$ 构成的平行四边形的面积。在这里，线向量和面向量混叠在一起。下面的叙述中，我们割掉了线向量这条小尾巴，只留下了面向量-------一个具有旋转方向的平行四边形面。

前面讲过，两个线向量 $\mathbf{a}$ 和 $\mathbf{b}$ 相加得到一个线向量 $\mathbf{a}+\mathbf{b}$，这个过程满足平行四边形法则和三角形法则。那么，线向量 $\mathbf{a}$ 和 $\mathbf{b}$ 分别和第三个线向量 $\mathbf{c}$ 叉积依次得到了两个面向量 $\mathbf{a}\times\mathbf{c}$ 和 $\mathbf{b}\times\mathbf{c}$。实际上，面向量的加法运算同样满足平行四边形法则和三角形法则。向量叉积的分配律 $\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}=(\mathbf{a}+\mathbf{b})\times\mathbf{c}$ 的图示如下图所示。




面向量的加法法则的证明可以从封闭面的和向量为零着手。我们这里不再证明了。实际上，我们有更形象的图证来理解这个法则。比如对于三角形法则（右图），我们可以想象有向量 $\mathbf{c}$ 的长度那么多的线向量 $\mathbf{a}+\mathbf{b}$ 的三角形叠加在一起形成面向量，叠加的方向沿着向量 $\mathbf{c}$ 的方向进行，如下图。



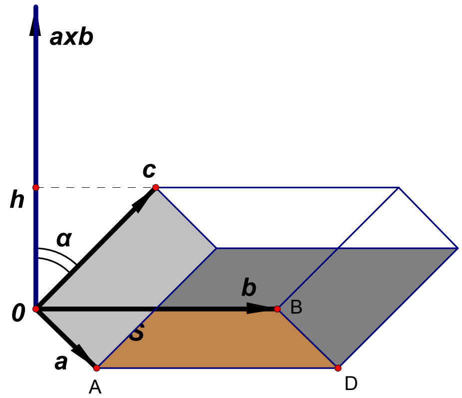

## 向量的混合积的几何解释

三个向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$，如果先作两向量 $\mathbf{a}, \mathbf{b}$ 的点积 $\mathbf{a}\cdot\mathbf{b}$，因为它是数量，所以再与第三个向量 $\mathbf{c}$ 相乘的结果 $(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$ 表述一个新向量，它是向量 $\mathbf{c}$ 的伸缩，与向量 $\mathbf{c}$ 平行。

如果先作两向量 $\mathbf{a}, \mathbf{b}$ 的叉积 $\mathbf{a}\times\mathbf{b}$，这个所得的向量与第三个向量 $\mathbf{c}$ 再作点积 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 或者叉积 $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$，前者表示数量叫做三向量的混合积或三重数积；后者表示向量，叫做三重矢积。下面我们仅对向量的混合积的几何意义进行讨论。

三向量的混合积有关系 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}=(\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}$，且其中 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=|\mathbf{a}\times\mathbf{b}||\mathbf{c}|\cos\alpha$，$(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}=|\mathbf{b}\times\mathbf{c}||\mathbf{a}|\cos\beta$，$(\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}=|\mathbf{c}\times\mathbf{a}||\mathbf{b}|\cos\gamma$。混合积是这样的一个数，他的绝对值表示以向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 为棱的平行六面体的体积。如果向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 以顺序组成右手系，那么积的符号是正的；如果组成左手系，积就是负的。

我们知道，向量积 $\mathbf{a}\times\mathbf{b}$ 是一个向量，它的模在数值上等于以向量 $\mathbf{a}, \mathbf{b}$ 为边的平行四边形 $0ADB$ 的平面，它的方向垂直于这个平行四边形的平面，且当向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 以顺序组成右手系时，即向量 $\mathbf{a}\times\mathbf{b}$ 与向量 $\mathbf{c}$ 是朝着此平面的同一侧（如图）；且当向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 以顺序组成左手系时，即向量 $\mathbf{a}\times\mathbf{b}$ 与向量 $\mathbf{c}$ 是朝着此平面的另一侧。所以若向量 $\mathbf{a}\times\mathbf{b}$ 与向量 $\mathbf{c}$ 之间的夹角为 $\alpha$，则当向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 以顺序组成右手系时，夹角 $\alpha$ 为锐角；当向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 以顺序组成左手系时，夹角 $\alpha$ 为钝角。




又因为 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=|\mathbf{a}\times\mathbf{b}||\mathbf{c}|\cos\alpha$，则有当向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 组成右手系时，$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 为正值；当向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 组成左手系时，$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 为正值。

因此，以向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的底（平行四边形 $0ADB$）的面积 $S$ 在数值上等于 $|\mathbf{a}\times\mathbf{b}|$，它的高 $h$ 等于向量 $\mathbf{c}$ 在向量 $\mathbf{a}\times\mathbf{b}$ 上的投影，即 $h=|\mathbf{c}|\cos\alpha$，所以平行六面体的体积等于

$$V=Sh=|\mathbf{a}\times\mathbf{b}||\mathbf{c}|\cos\alpha\text{。}$$

## 2.6 向量的积和张量之间的关系

从前面的看来，向量的内积定义和外积定义确有意义，但对于我们玩惯实数乘法且还没有得到高等数学训练的人来说，还是有点拿捏不住。为什么不能像两个多项式的乘法一样定义两个向量的乘法呢？

完全可以这样相乘。

### 二维向量的内积、外积和张量

先看二维空间中的两个向量 $\mathbf{a}=(a_x,a_y)$，$\mathbf{b}=(b_x,b_y)$ 把它们改写成带有 $x,y$ 坐标轴的单位向量的形式，就是 $\mathbf{a}=a_x\mathbf{i}+a_y\mathbf{j}$，$\mathbf{b}=b_x\mathbf{i}+b_y\mathbf{j}$，我们对向量 $\mathbf{a}$ 和 $\mathbf{b}$ 就象普通的多项式乘法分配律一样展开运算，得到如下：

$$\mathbf{ab}=(a_x\mathbf{i}+a_y\mathbf{j})(b_x\mathbf{i}+b_y\mathbf{j})\ =\ \underline{a_xb_x\mathbf{ii}+a_yb_y\mathbf{jj}}+\underline{(a_xb_y\mathbf{ij}+a_yb_x\mathbf{ji})}$$

这里有个关键的问题，就是如何定义坐标轴的单位向量 $\mathbf{i},\mathbf{j}$ 之间的运算？我们发现，不同的规

定，就会得到不同的结论：

- 当定义 $\mathbf{i},\mathbf{j}$ 之间的运算为内积运算时，即 $\mathbf{ii}=\mathbf{jj}=1,\ \mathbf{ij}=\mathbf{ji}=0$ 时上式化简为：

$$\mathbf{ab}=\underline{a_x b_x + a_y b_y}=\mathbf{a}\cdot\mathbf{b}$$

这正是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 作内积运算的定义式。

- 当定义 $\mathbf{i},\mathbf{j}$ 之间的运算为外积运算时，即 $\mathbf{ii}=\mathbf{jj}=0,\ \mathbf{ij}=-\mathbf{ji}=\mathbf{k}$ 时上式化简为：

$$\mathbf{ab}=\ a_x b_y \mathbf{ij}+a_y b_x \mathbf{ji}=(a_x b_y - a_y b_x)\mathbf{k}=\mathbf{a}\times\mathbf{b}$$

这正是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 作外积运算的定义式，在 二维向量空间外又生成了一个第三维向量。

- 当定义 $\mathbf{i},\mathbf{j}$ 之间的运算只是作为一个顺序的记号时，即 $\mathbf{ij}=1,\ \mathbf{ii}=\mathbf{jj}=0,\ \mathbf{ji}=-1$ 时上式化简为：

$$\mathbf{ab}=\ a_x b_y \mathbf{ij}+a_y b_x \mathbf{ji}=a_x b_y - a_y b_x=\begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix}=\begin{vmatrix} \mathbf{a} \\ \mathbf{b} \end{vmatrix}$$

这正是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 作行向量构成的方阵的行列式的运算的定义式。

- 当定义 $\mathbf{i}=1$，$\mathbf{j}=\sqrt{-1}$，与复数进行对应时，上式化简为：

$$\mathbf{ab}=\underline{a_x b_x - a_y b_y}+\underline{(a_x b_y + a_y b_x)\mathbf{j}}$$

这正是复数 $\mathbf{a}$ 和 $\mathbf{b}$ 作乘法运算的定义式。

总结一下：

本等式 $\mathbf{ab}=(a_x\mathbf{i}+a_y\mathbf{j})(b_x\mathbf{i}+b_y\mathbf{j})\ =\ \underline{a_x b_x \mathbf{ii}+a_y b_y \mathbf{jj}}+\underline{(a_x b_y \mathbf{ij}+a_y b_x \mathbf{ji})}$ 的第一部分包含了向量 $\mathbf{a}$ 和 $\mathbf{b}$ 内积的结果，第二部分包含了向量 $\mathbf{a}$ 和 $\mathbf{b}$ 外积的结果或者是行列式的结果，即：

$$\mathbf{ab}=(\mathbf{a}\cdot\mathbf{b})+(\mathbf{a}\times\mathbf{b})$$

从这个结论看来，向量的内积运算、外积运算覆盖了二维向量及其复数的所有乘积模式的结果。

**实际上， 象上述的多项式一样的向量乘法叫做向量的直积，向量的直积是向量之间最简单的一种乘法运算，其结果是张量（向量是一阶张量的一种），所以也叫做向量（矢量）的张量积，俗称并矢。**

因此，采用较高等一点的说法就是，向量的张量积包含了向量的内积和外积的结果。

## 三维向量的内积、外积和张量

再看看三维空间中的两个向量为 $\mathbf{a}=a_x\mathbf{i}+a_y\mathbf{j}+a_z\mathbf{k}$，$\mathbf{b}=b_x\mathbf{i}+b_y\mathbf{j}+b_z\mathbf{k}$ ， 同样，我们对向量 $\mathbf{a}$ 和 $\mathbf{b}$ 同样展开直积运算，得到如下的张量：

$$\begin{aligned}\mathbf{ab}&=(a_x\mathbf{i}+a_y\mathbf{j}+a_z\mathbf{k})(b_x\mathbf{i}+b_y\mathbf{j}+b_z\mathbf{k})\\&=\underline{a_xb_x\mathbf{ii}+a_yb_y\mathbf{jj}+a_zb_z\mathbf{kk}}+\underline{(a_xb_y\mathbf{ij}+a_yb_x\mathbf{ji})+(a_xb_z\mathbf{ik}+a_zb_x\mathbf{ki})+(a_yb_z\mathbf{jk}+a_zb_y\mathbf{kj})}\end{aligned}$$

有点复杂，有没有看出点规律？

与二维向量处理的方式类似，我们定义坐标轴的单位向量 $\mathbf{i},\mathbf{j},\mathbf{k}$ 之间的不同运算，得到了以下不同的结论：

- 当定义 $\mathbf{i},\mathbf{j},\mathbf{k}$ 之间的运算为内积运算时，即 $\mathbf{ii}=\mathbf{jj}=\mathbf{kk}=1$，$\mathbf{ij}=\mathbf{ji}=\mathbf{ik}=\mathbf{ki}=\mathbf{jk}=\mathbf{kj}=0$ 时上式化简为：

$$\mathbf{ab}=a_xb_x+a_yb_y+a_zb_z=\mathbf{a}\cdot\mathbf{b}$$

这正是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 作内积运算的定义式。

- 当定义 $\mathbf{i},\mathbf{j},\mathbf{k}$ 之间的运算为外积运算时，即 $\mathbf{ii}=\mathbf{jj}=\mathbf{kk}=0$， $\mathbf{ij}=-\mathbf{ji}=\mathbf{k}$， $\mathbf{jk}=-\mathbf{kj}=\mathbf{i}$，

  $\mathbf{ki}=-\mathbf{ik}=\mathbf{j}$ 时上式化简为：

$$\begin{aligned}\mathbf{ab}&=(a_xb_y\mathbf{ij}+a_yb_x\mathbf{ji})+(a_xb_z\mathbf{ik}+a_zb_x\mathbf{ki})+(a_yb_z\mathbf{jk}+a_zb_y\mathbf{kj})\\&=(a_xb_y-a_yb_x)\mathbf{k}+(a_zb_x-a_xb_z)\mathbf{j}+(a_yb_z-a_zb_y)\mathbf{i}\\&=\mathbf{a}\times\mathbf{b}\end{aligned}$$

这正是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 作外积运算的定义式，在三维向量空间内又生成了第三个同维向量。

总结一下：

本等式，

$$\begin{aligned}\mathbf{ab}&=(a_x\mathbf{i}+a_y\mathbf{j}+a_z\mathbf{k})(b_x\mathbf{i}+b_y\mathbf{j}+b_z\mathbf{k})\\&=\underline{a_xb_x\mathbf{ii}+a_yb_y\mathbf{jj}+a_zb_z\mathbf{kk}}+\underline{(a_xb_y\mathbf{ij}+a_yb_x\mathbf{ji})+(a_xb_z\mathbf{ik}+a_zb_x\mathbf{ki})+(a_yb_z\mathbf{jk}+a_zb_y\mathbf{kj})}\end{aligned}$$

其第一部分包含了向量 $\mathbf{a}$ 和 $\mathbf{b}$ 内积的结果，第二部分包含了向量 $\mathbf{a}$ 和 $\mathbf{b}$ 外积的结果或者是行列式的结果，即：

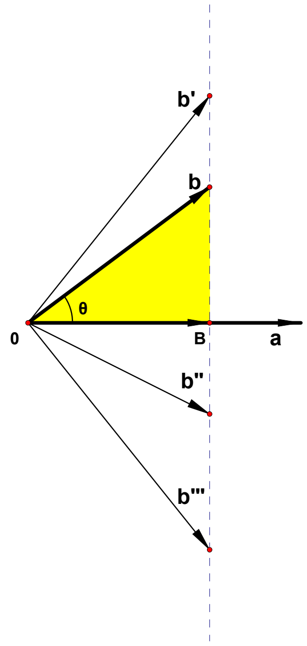
$$\mathbf{ab}=(\mathbf{a}\cdot\mathbf{b})+(\mathbf{a}\times\mathbf{b})$$

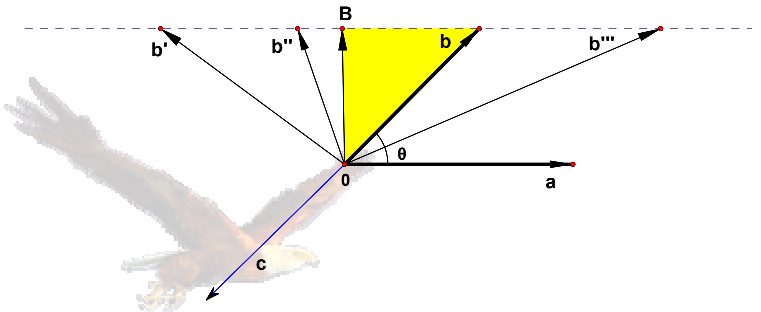
从这个结论看来，向量的内积运算、外积运算覆盖了两个三维向量的所有乘积模式的结果，或者说，向量的张量积包含了向量的内积和外积的结果。

## 2.7 向量有没有除法?

向量的乘法有两种：点积和叉积。一般讲除法是乘法的逆运算。那么向量的除法是点积的除法呢还是叉积的除法？看来比较麻烦，所以大家较少谈论向量的除法。这里我们看看如何麻烦的？先约定一下符号：两个向量 $\mathbf{a}$ 和 $\mathbf{b}$，分别有 $\mathbf{a}\cdot\mathbf{b}=c$ 和 $\mathbf{a}\times\mathbf{b}=\mathbf{c}$；已知 $\mathbf{a}$ 和乘积，如何求 $\mathbf{b}$。

能不能得到向量 $\mathbf{b}$，请看下图。




左图中，我们根据点积的公式 $\mathbf{a}\cdot\mathbf{b}=ab\cos\theta$，立刻得到：

$$\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\mathbf{b}'=\mathbf{a}\cdot\mathbf{b}''=\mathbf{a}\cdot\mathbf{b}'''\ldots=c$$

无数个向量 $\mathbf{b}$、$\mathbf{b}'$、$\mathbf{b}''$、$\mathbf{b}'''$…皆能得到同样的点积值。这个意思是说，点积没有除法。

再看右图，我们根据叉积的公式 $\mathbf{a}$ x $\mathbf{b}$ $=ab\sin\theta N$，立刻得到：

$$\mathbf{a}\times\mathbf{b}=\mathbf{a}\times\mathbf{b}'=\mathbf{a}\times\mathbf{b}''=\mathbf{a}\times\mathbf{b}'''\ldots=\mathbf{c}$$

无数个向量 $\mathbf{b}$、$\mathbf{b}'$、$\mathbf{b}''$、$\mathbf{b}'''$…同样皆能得到同样的叉积值。这个意思是说，叉积也没有除法。

结果真让人失望！

莫急，如果我们把点积和叉积联立解方程组，倒可以解出除法的表达式出来。解方程：

$$\begin{cases}\mathbf{a}\cdot\mathbf{b}=c\\ \mathbf{a}\times\mathbf{b}=\mathbf{c}\end{cases}$$

对于$\mathbf{a}\times\mathbf{b}=\mathbf{c}$，两边左叉乘$\mathbf{a}$，并用二重向量叉积公式$\mathbf{a}\times(\mathbf{b}\times\mathbf{c})=\mathbf{b}\cdot(\mathbf{a}\cdot\mathbf{c})-\mathbf{c}\cdot(\mathbf{a}\cdot\mathbf{b})$得

$$\mathbf{a}\times(\mathbf{a}\times\mathbf{b})=\mathbf{a}\cdot(\mathbf{a}\cdot\mathbf{b})-\mathbf{b}\cdot(\mathbf{a}\cdot\mathbf{a})$$

把方程组的两个等式带入，即得$\mathbf{a}\times\mathbf{c}=\mathbf{a}\cdot c-\mathbf{b}\cdot(\mathbf{a}\cdot\mathbf{a})$，整理得到

$$\mathbf{b}=\frac{c\mathbf{a}-\mathbf{a}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{a}}$$

这就是向量的除法。

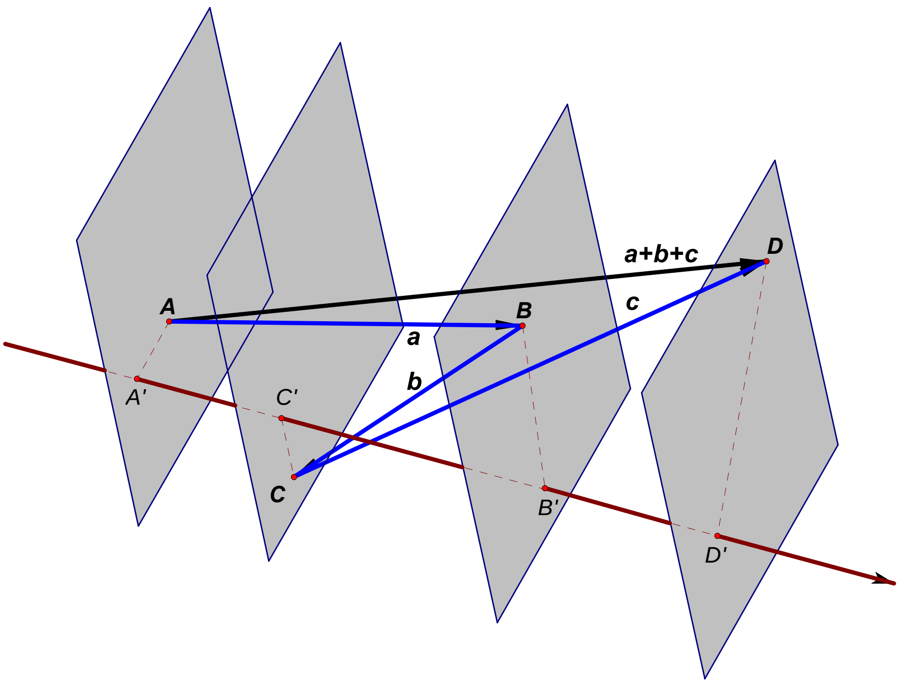
# 2.8 向量的投影和的几何解释

## 多个向量在任意轴上的投影和

向量的加法的推广之一就是：

多个或有限个向量的和在任意轴上投影等于各个向量在同一轴上投影的和。

下面我们给出其几何图形的图解。




图中，向量$\mathbf{a}$ $\mathbf{b}$ $\mathbf{c}$及其和向量$\mathbf{a}+\mathbf{b}+\mathbf{c}$的图形分别是：

$\mathbf{a}=\overrightarrow{AB}$，$\mathbf{b}=\overrightarrow{BC}$，$\mathbf{c}=\overrightarrow{CD}$，$\mathbf{a}+\mathbf{b}+\mathbf{c}=\overrightarrow{AD}$，各向量的端点 A、B、C、D 在轴上的投影分别是$A'$、$B'$、$C'$、$D'$。

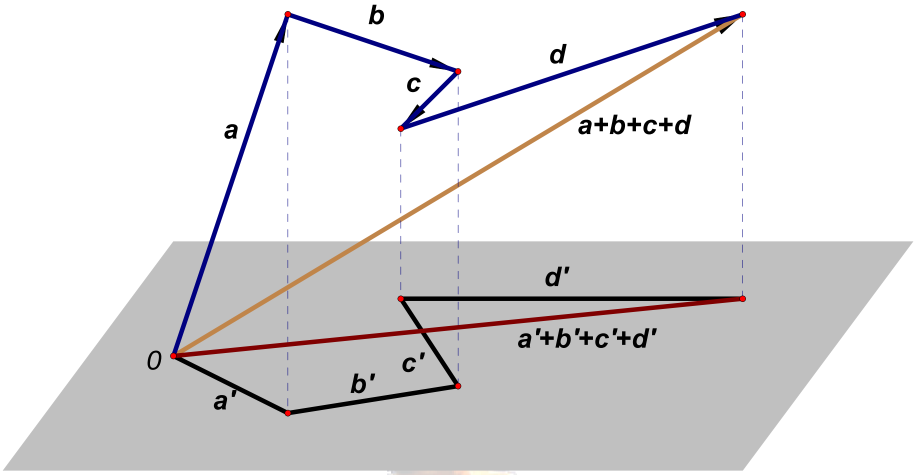
图示是显然的，$A'B'+B'C'+C'D'=A'D'$，命题得解。

### 多个向量在任意平面上的投影和

向量的加法的推广之二就是：

多个或有限个向量的和在任意平面上投影等于各个向量在同一平面上投影的和。

下面我们给出其几何图形的图解。




与向量在轴上的投影类似。结论在图解中一目了然。这个结论是由投影变换的性质所决定。

## 2.9 变向量的几何意义

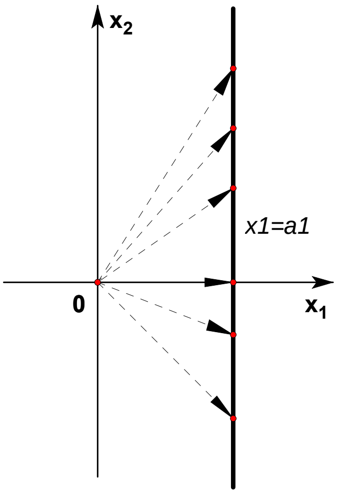
对于用数组$(a_1, a_2, ... a_n)$表示的向量$\mathbf{a}$，如果数组中的元素部分或者全部是变量，那么这个变向量在 n 维坐标系下表示的几何图形是什么呢？

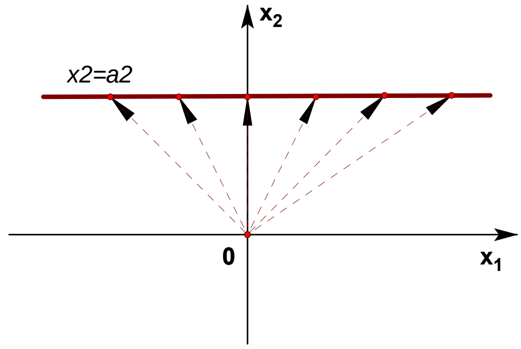
### 二维变向量的几何图形

在二维平面上，二维向量的两个分量全部为可变的未知量，记为$(x_1, x_2)$，那么变向量$(x_1, x_2)$可以表示平面上任意一个向量（无数个）。如果$x_1, x_2$取遍整个实数范围，则可以涵盖了整个向量平面，因此$(x_1, x_2)$的几何图形表示为一个平面。

如果我们固定$(x_1, x_2)$一个分量，如$(a_1, x_2)$，其中$a_1$表示某一确定的实数，$x_2$表示为确定的变元。那么变向量$(a_1, x_2)$表示的是无数个向量，这些向量的末端全部在直线$x_1 = a_1$上（如图2。X)。

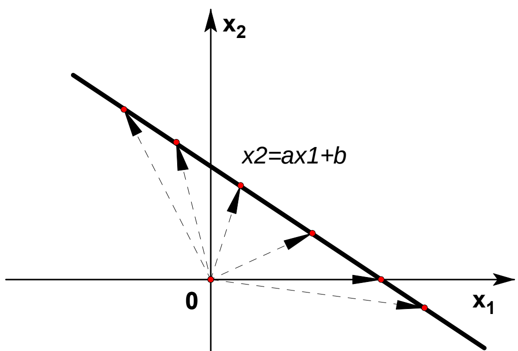
类似的，变向量$(x_1,a_2)$表示的是直线$x_2=a_2$，无数个向量的末端全部在直线$x_2=a_2$上。

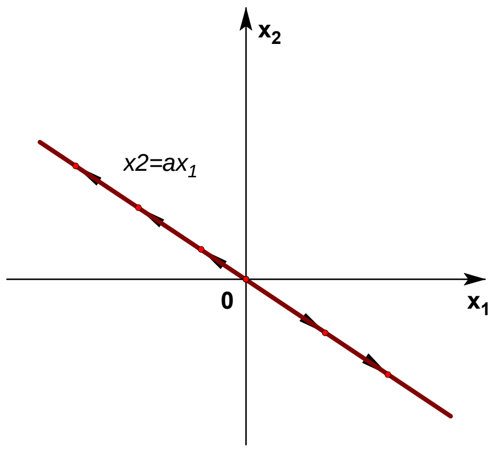


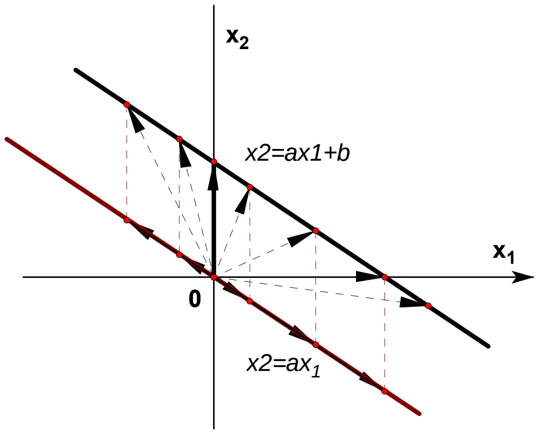

进一步，如果向量的分量之间有线性关系的话，比如$x_2=ax_1+b$，向量可以表示为$(x_1,ax_1+b)$的形式。那么变向量$(x_1,ax_1+b)$就可以表示一个直线，所有向量的末端在直线$x_2=ax_1+b$上。

当然，如果$b=0$，即$(x_1,ax_1)$就表示为一条过原点的直线，这时候所有的向量的始端和末端都处于直线上。




方程$x_2=ax_1+b$和$x_2=ax_1+b$是平行线，$b$是$x_2$轴上的截距。我们用变向量来表示这对平行线的关系就是$(x_1,ax_1+b)=(x_1,ax_1)+(0,b)$。用向量的观点解释就是，变向量$(x_1,ax_1+b)$是变向量$(x_1,ax_1)$及常向量$(0,b)$的和。从图形上解释就是直线$(x_1,ax_1+b)$是由过原点的直线$(x_1,ax_1)$沿向量$(0,b)$平移$b$得到的。其几何图性给出如下：




## 三维变向量的几何图形

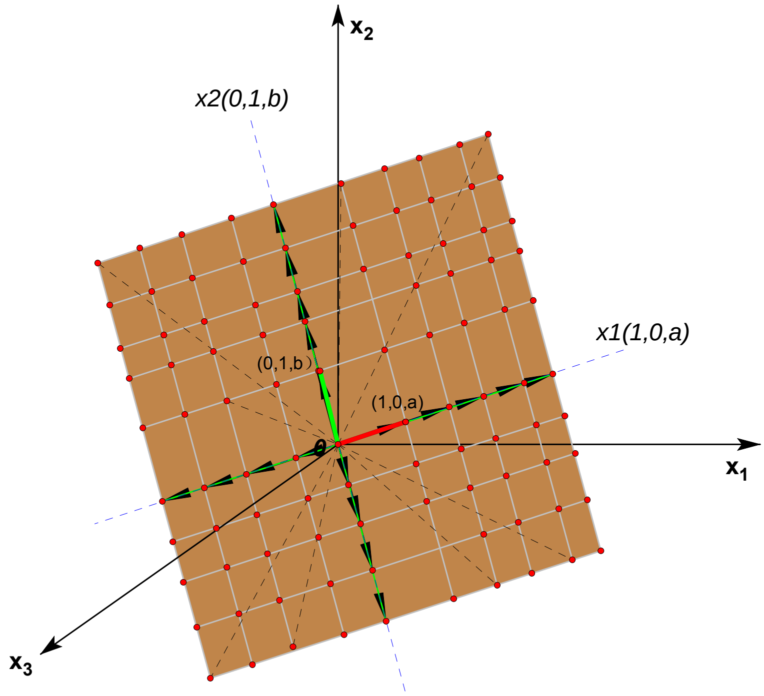
与二维变向量相类似，$(x_1,x_2,x_3)$ 有三个不定变元，表示三维向量空间中的任意一个向量，代表了整个三维空间。

变向量 $(a_1,x_2,x_3)$、$(x_1,a_2,x_3)$ 和 $(x_1,x_2,a_3)$ 各有两个不定变元，分别表示了平面 $x_1=a_1$，$x_2=a_2$ 和 $x_3=a_3$。每个变向量中任意一个向量的末端都在这个变向量所表示的平面上。这些平面分别垂直于一个坐标轴。

变向量 $(a_1,a_2,x_3)$、$(a_1,x_2,a_3)$ 和 $(x_1,a_2,a_3)$ 各有一个不定变元，分别表示了三根直线 $\begin{cases}x_1=a_1\\x_2=a_2\end{cases}$，$\begin{cases}x_1=a_1\\x_3=a_3\end{cases}$ 和 $\begin{cases}x_2=a_2\\x_3=a_3\end{cases}$，每个变向量中任意一个向量的末端都在这个变向量所表示的直线上。这些直线分别平行于一个坐标轴。

变向量 $(x_1,x_2,ax_1+bx_2)$ 有两个独立变元 $x_1,x_2$，第三个变元 $x_3$ 与 $x_1,x_2$ 有线性关系：$x_3=ax_1+bx_2$。为了方面看到这个变向量的几何图形，我们对它进行向量分解：

$$(x_1,x_2,ax_1+bx_2)=(x_1,0,ax_1)+(0,x_2,bx_2)=x_1(1,0,a)+x_2(0,1,b)$$




在 $x_1, x_2$ 独立地取遍所有不同的实数时，$x_1(1,0,a)+x_2(0,1,b)$ 所形成的无数个向量覆盖了一个平面。实际上，在后面的章节中，这个平面是一个向量空间，一个被常向量 $(1,0,a)$ 和 $(0,1,b)$ 所张成的向量平面空间，记为 $Span\{(1,0,a),(0,1,b)\}$。因此，我们可以有这样的一个等价式：

$$(x_1, x_2, ax_1+bx_2) \cong Span\{(1,0,a),(0,1,b)\}$$

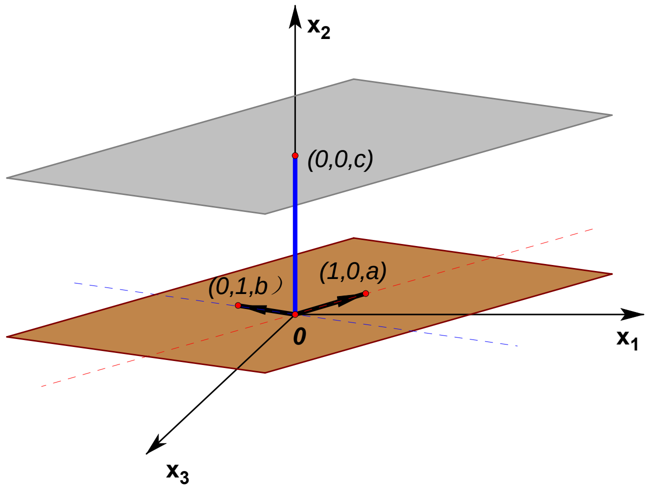
类似的，变向量 $(x_1, x_2, ax_1+bx_2+c)$ 也有两个独立变元 $x_1, x_2$，第三个变元 $x_3$ 与 $x_1, x_2$ 有线性关系：$x_3 = ax_1+bx_2+c$。同样，我们也对它进行向量分解：

$$(x_1, x_2, ax_1+bx_2+c) = (x_1, 0, ax_1)+(0, x_2, bx_2)+(0,0,c) = x_1(1,0,a)+x_2(0,1,b)+(0,0,c)$$

同变向量 $(x_1, x_2, ax_1+bx_2)$ 比较起来，变向量 $(x_1, x_2, ax_1+bx_2+c)$ 与它的关系可以表示为：

$$(x_1, x_2, ax_1+bx_2+c) = (x_1, x_2, ax_1+bx_2)+(0,0,c)$$

两个变向量差了一个常向量。因此变向量 $(x_1, x_2, ax_1+bx_2+c)$ 的几何图形是 $(x_1, x_2, ax_1+bx_2)$ 的图形加了一个常向量 $(0,0,c)$。得到的新图形仍然是个平面，只是沿着向量 $(0,0,c)$ 的方向平移了长度为 $c$ 的距离。




：一个变向量是和一个解析方程或方程组相对应的，因此变向量和方程一样能表示一个几何图形。实际上，变向量也叫向量函数。

## 变向量的应用

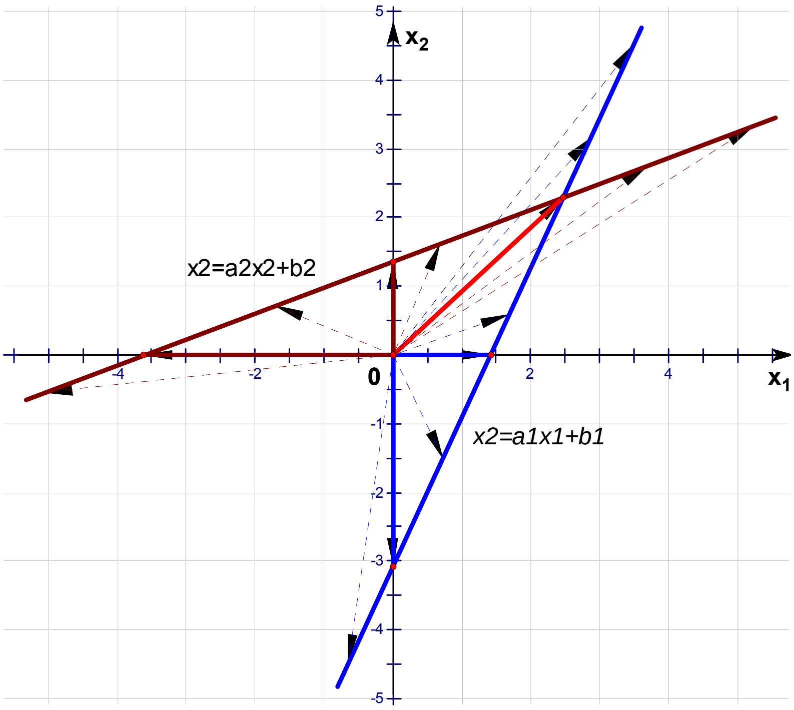
变向量简化了线性方程或方程组，但这种简化与线性方程组的向量方程或者矩阵方程对线性方程的简化又有不同。下面我们看看如何用变向量的概念去解线性方程组。

对于二阶线性方程组$\begin{cases} x_2 = a_1x_1 + b_1 \\ x_2 = a_2x_1 + b_2 \end{cases}$可以用变向量的交$(x_1, a_1x_1 + b_1) \cap (x_1, a_2x_1 + b_2)$来等价表示。可以想象，两个变向量相同的向量就是他们的交集。如果两个向量相同，那么向量的夹角就会为零，根据行列式的几何意义知道，两个夹角为零的向量构成的平行四边形为零，即行列式为零。因此得到：

$$\begin{vmatrix} x_1, a_1x_1 + b_1 \\ x_1, a_2x_1 + b_2 \end{vmatrix} = 0$$

化简整理得到 $x_1 = 0$ 或者 $x_1 = \dfrac{b_1 - b_2}{a_2 - a_1}$。进而得到向量对 $(0, b_1), (0, b_2)$ 和 $\left(\dfrac{b_1 - b_2}{a_2 - a_1}, \dfrac{a_2b_1 - a_1b_2}{a_2 - a_1}\right)$及其自身。

实际上，应用行列式可以求出所有的相关向量。如上式得到了两对线性相关的向量。其中一对向量重合，即得到了线性方程组的解$\left(\dfrac{b_1 - b_2}{a_2 - a_1}, \dfrac{a_2b_1 - a_1b_2}{a_2 - a_1}\right)$。




从图中我们看到，还有一对线性相关的向量没有求出来。但如果对线性方程组的变向量以 $x_2$ 为自变量时就可以得到这对相关量。

## 2.10 向量的“社会”关系

在这一节中，我们简要的探讨一下向量与主要几个数学概念或领域的联系，以期更好的理解向量。

### 向量和矢量

向量就是矢量。

Vector，物理界叫“矢量”，数学界叫“向量”。其实是一回事。在历史上，数学界曾把vector定名为“矢量”，而物理界曾把它定名为“向量”。后来，可能是为了尊重对方，却对换了一下，数学用向量，物理用矢量。矢量、向量的分歧，因为各自学科发展繁多，学科有分支，术语有派生，统一起来没那么容易，因此分歧一直维持到今。事实上，台湾物理界至今用的是还“向量”。

英文vector的中文译名是“向量”。顾名思义，就是指“既有方向又有大小的量”。可以说译得十分贴切。不过，中国在清末引进vector概念时，物理学家称之为“矢量”，至今两种译名并存。究其原因，早年的向量，只是物理学专门用来表示力和速度等物理量的工具，并不为数学家所重视。因为物理学用得多，矢量的译名自然流行。向量进入数学领域之后，渐渐有取代“矢量”之势。

### 向量与数

复数为“媒”，撑起“向量空间”一片天史载，古希腊的亚里士多德(公元前384—公元前322)已经知道两个力的合成，可以用平行四边形的法则得到。但是，集古希腊数学大成的《几何原本》，

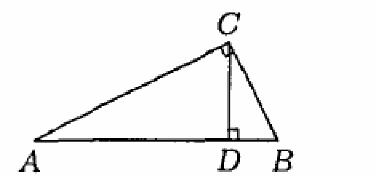
没有讨论向量。 以后的一千多年中，经过文艺复兴时期，牛顿创立微积分之后的17、18世纪，向量的知识没有什么变化。伽利略(1564—1642)清楚地叙述了“平行四边形法则”，仅此而已。这点向量知识，形不成多少有意义的问题，发展不成一个独立的学科，因而数学家没有把向量当作一回事。进入19世纪，事情开始发生变化。“复数”充当了催化剂。丹麦的魏塞尔(1745—1818)，瑞士的阿工(1768—1822)发现了复数的几何表示，德国高斯(1777-1855)建立了复平面的概念，从而使向量与复数建立起一一对应。这不但为虚数的现实化提供了可能，也为向量的发展开辟了道路。向量表示为-x˜有序的实数(a，6)，是一个重大的进步。当时的数学家想到，实数可看作一维向量，复数可看作二维向量，那么一定还有“三维数”、“四维数”，乃至“n维数”。令人失望的是，哈密发现，要形成有加减乘除四则运算的数系，只能是四元数，而且不得不放弃乘法的交换律。最后发现的八元数，连结合律也维持不了。除此而外，其他维数的向量，根本无法定义四则运算，谈不上构成数系。I1 J德国数学家格拉斯曼1844年引入了n维向量的概念。 令人深思的是，n维向量既然不能成为有四则运算的数系，那么它的结构是什么呢?这是19世纪抽象代数思想的发展的自然思考。研究表明，礼维向量全体，可以定义加法和减法，此外还有单个的“数”可以和向量相乘。这就是向量空间(线性空间)的来源。此外，两个向量可以有“内积”和“外积”，但是它们都没有逆运算，即没有除法。这是一个不同于“数系”的崭新的数学结构。果然，在向量空间的舞台上，产生了具有深远影响的数学成就。

尽管实平面$R^2$中的元素和复平面C是一一对应的（见图），并且加法也相同，但逻辑上两者是不同的，在$R^2$上，我们仅仅能对一个向量作标量乘法（向量的内积不能作为向量乘法，因为内积得到一个数，而不是向量，它不属于$R^2$），然而在C上，我们可以对任意两个复数相乘并得到第三个复数。

## 向量和几何（向量几何）

使“点”可以运算平面几何一向以综合法的演绎为主，以后引入坐标系，发展为坐标几何，即解析几何。数与形互相结合，使得几何学别开生面。如前所述，坐标表示的“点”，依然不能运算。有了向量之后，情况发生了改变。几何学家项武义认为“向量几何在本质上是坐标几何的返璞归真。”即向量几何揭示了坐标几何的本质，是坐标几何的向前发展。向量几何使用“向量的数量积”，使之成为超越坐标几何的有力工具。两条直线的夹角，当然也可以从两直线方程的系数求得。但是，在向量几何里，它可以用两直线的方向向量的数量积加以表示。本来很费事的夹角问题，通过一次运算就解决了。利用向量，许多几何命题迎刃而解。至于利用向量讨论直线与直线的垂直与平行，空间线面、面面之间的位置关系，比起综合方法需要“个别处理”的技巧，它是一个“一揽子”解决的手段，显得十分轻松。

以下是初中的教学事例：勾股定理的证明(事先准备知识：由一个角的两边的任何一点向另一边作投影，其压缩的比值相同)。



如图，已知CD是Rt$\Delta$ABC斜边AB上的高。在$\angle A$中，$AC=\alpha AB$，$AD=\alpha AC$，因此$AD=\alpha^2 AB$。

在$\angle B$中，$BC=\beta AB$，$BD=\beta BC$，因此$BD=\beta^2 AB$。

由于$AB=AD+BD=\alpha^2 AB+\beta^2 AB=(\alpha^2+\beta^2)AB$，因此$\alpha^2+\beta^2=1$。

所以，$AC^2+BC^2=\alpha^2 AB^2+\beta^2 AB^2=(\alpha^2+\beta^2)AB^2=AB^2$。

这里，将线段的投影，三角的余弦，以及未来的向量分解和数量积等知识都拧在一起，并用来证明勾股定理，在数学思想上更简约、更紧密了。总之，向量和几何的融合，已是不可阻挡的潮流。

## 向量和三角：

把三角看作向量的“投影”以来，三角学的处理引人注目。数学科普作家张景中院士认为数学课程中，三角至关重要。三角是联系几何与代数的一座桥梁，数形结合的一份经典，沟通初等数学和高等数学的一条通道。函数、向量、坐标、复数等许多重要的数学知识与三角有关，大量的实际问题的解决要用到三角知识。

在现行的课程中，三角函数是作为直角三角形的两边的比值引进的。这样的定义，依赖于有关相似三角形的知识，而且只能定义锐角的三角函数。其实，三角函数的定义，无非是给三角函数提供一个几何模型。张景中建议用面积方法建立三角学。首先采用新的角度认识正弦。定义：把边长为1，有一个角为 的菱形面积记作sinA。于是，容易得到平行四边形的面积公式

$$S_{YABCD}=AB\cdot AD\cdot\sin A\ ,$$

取其一半，得到三角形面积功式：

$$S_{\Delta ABC}=\frac{1}{2}bc\sin A=\frac{1}{2}ac\sin B=\frac{1}{2}ab\sin C$$

正弦可以看做将边长为1的正方形“压扁”为菱形时面积所打的折扣。折扣，是小学生就明白的数学常识。现在，如果我们把折扣观念用到向量的分解上，余弦和正弦就是向量OA在两个坐标轴上的投影在长度上所打的折扣。

将向量 a和b的数量积定为

$$ab=|a||b|\cos\alpha$$

这样也自然地把三角和向量联系起来了：单位圆上的半径向量OA，其在 轴和Y轴上的投影长度，就是余弦和正弦。于是，任意角的三角函数的符号变化就显得一清二楚，正弦、余弦函数的图象也很容易画出。

特别是，由向量的数量积(坐标定义)，导出余弦的差角公式简直是举手之劳，一步到位，简单极了。如图2所示，利用数量积可得

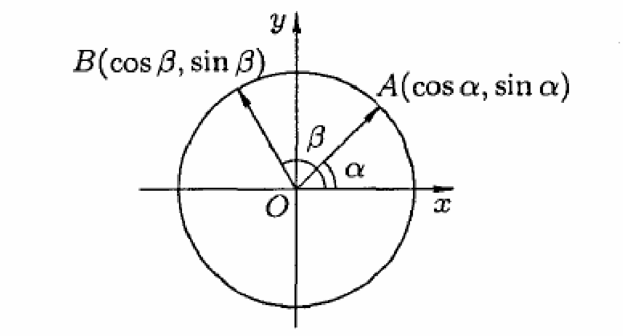



图2

$$\cos(\beta-\alpha)=\frac{\cos\alpha\cos\beta+\sin\alpha\sin\beta}{|\overrightarrow{OA}||\overrightarrow{OB}|}$$

$$因|\overrightarrow{OA}|=|\overrightarrow{OB}|=1$$

因此：

$$\cos(\beta-\alpha)=\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta$$

向量非常直观，几何意义很明显。即使是数量积，也非常直观。到商场购物，71，种物品的价格向量 $P(p_1, p_2, ...p_n)$ 和你购物的数量向量 $B=(b_1b_2,...b_n)$ 的数量积就是你的付款数：

$$A=P\cdot B=p_1b_1+p_2b_2+...+p_nb_n$$

所谓向量的数量积，其实就在我们身边。最后，采用数量积的坐标定义，能否得出数量积的运算律，特别是分配律?经过严格论证，完全可以非常简单地得到，而且也不会出现循环论证。

## 向量与解析几何

向量在今天越来越受人重视，为何呢?说来简单，无非是向量“能算”。

在中学，解析几何课程主要研究平面上一些点、线的几何位置和几何性质，所涉及的点、直线、曲线均共面，它们都落在坐标系所在的平面内。在大学，解析几何课程所讨论的主要是空间图形，所涉及的点、线、面的位置关系复杂。教科书从空间向量与坐标入手，主要以向量为工具，研究空间中的一些几何图形及性质，不论从知识上还是从思维上，都是对中学内容的加深和提高。由于平面上的几何图形比较直观，点、线之间的关系比较简单，主要以坐标为工具进行研究;而空间上的几何图形，大多比较复杂，构成图形的点、线、面及其之间的位置关系较复杂，仅用坐标工具远远不能达到研究问题的最佳境界，只有使用向量这一工具，才能比较好而深刻地探讨空间图形问题。实际上向量综合了图形（直线段或者说方向）和数字（向量终点用一组数表示）两种几何分析方法使然。

无论是直线、平面方程的讨论，还是空间曲面、曲线的参数方程以及通过方程去讨论图形，到处都需要向量的参与。它就象木匠的尺子、石匠的凿子，在整个学科的展开中处处需要，向量工具灵活、好用。它有时是三角形的边，三角形的边的关系，经过向量的参与就变成简单的和的关系。它有时又是一个方向，用它可以决定直线的方向、平面的倾斜度，而一族直线、一族平面之间的关系有时通过两个小小的向量之间的关系就可以解决了。在讨论空间曲线、空间曲面的方程时，复杂、多变的轨迹问题，又变成了有公共起点的变向量问题，寻找到变向量的变化规律后，问题就迎韧而解了。因此说，向量思想是解析几何学的灵魂，要学好《解析几何》这一课程就必须掌握向量的作用，并要牢牢把握好向量这一有效工具。

## 向量与线性代数

“向量”上台，“线性数学”大放异彩“线性”，是20世纪数学中使用十分广泛的词汇。无论是英文还是俄文，我们常说的“一次方程”和次函数”，原本都是“线性方程(Linear Equation)”和“线性函数(Linear Function)”。在大学里，则流行“线性”。 “线性代数”、“线性变换”、“线性常微分方程”、“线性偏微分方程”、“线性规划”、“线性算子”、“线性泛函”、“线性控制系统”、“拟线性”、“准线性”等等。为什么以向量为基本对象的“线性数学”会流行呢?

实际上，相对于“非线性数学”来说，线性数学比较简单。 微积分学的基本思想是“以直代曲”，局部地以切线代替曲线。 于是，在某种条件下，微分方程就可以近似地变成“线性代数方程组”。20世纪有了电子计算机，无论未知数的个数n有多大，都可以设法计算。于是，把以n维向量为对象的线性方程组搞清了，许多复杂的数学问题也就有解了(至少是近似的)。工程中十分有效的有限元方法，其基础就是求解线性方程组。总之，在n维向量空间基础上生成的线性代数，成了许多数学问题得以解决的必备工具。计算机技术的推动，使得各种各样的“线性数学”成熟起来。

矩阵是线性代数中的主要研究对象，矩阵是“函数”概念的推广，矩阵描述向量之间的变换。如果说，函数是“数”与“数”之间的对应关系，那么矩阵则是向量与向量之间的对应(线性对应关系)，可以看作“函数”概念的推广。矩阵作为一种新型的数学表示工具，已经或者正在成为现代公民的常识。

矩阵应用虽然十分广泛，但是在数学上最亲近的原型还是线性方程组的求解。从高斯消去法看矩阵的变化，自然是必不可少的。然而，给人的印象是“杀鸡用牛刀”，数学上缺乏非矩阵不可的那种探究动力。于是，将二阶线性方程组用向量作两种不同的考察，数学上的新意顿出，令人深思。种是常规的，即写成$\begin{pmatrix} a,b \\ c,d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} m \\ n \end{pmatrix}$形式。解方程就是找矩阵$\begin{pmatrix} a,b \\ c,d \end{pmatrix}$一个逆矩阵，恰将向量(m，n)变为(x ，y )。另一个角度就是写成x(a，c)+y(b，d)=(m，n)。解方程就是设法将(m，n)表示成(a，c)和(b，d)的线性组合。这样就很自然地引出了行列式。

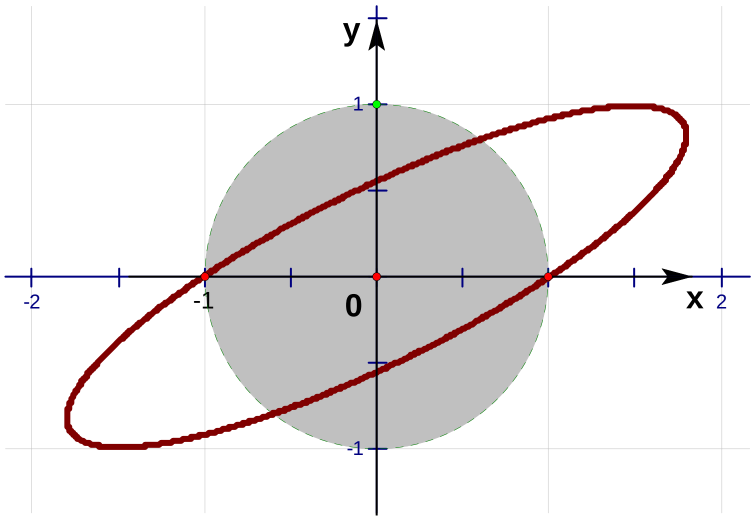
以下章节待续。。。

-----图解线性代数-----

# 线性代数的几何意义

# 之（3）

任广千　胡翠芳 编著




2010.06.01

# 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

# 第三章　行列式的几何意义

在中国古代，用筹算表示联立一次方程未知量的系数时，就有了行列式的萌芽-----排列的方式。日本吸收了这种思想，在 1683 年，日本学者关孝和（Seki Takakusu）对行列式的概念和它的展开已有了清楚的叙述。到 18 世纪，瑞士数学家克莱姆（G.Gramer）和法国数学家拉普拉斯（P.S.Laplace）建立了行列式理论。

行列式的几何意义具有深刻的含义。它是指行列式的行向量或列向量所构成的平行多面体的有向体积。这个有向体积是由许多块更小的有向面积或有向体积的累加。在我们逐步地讨论这个几何意义之前，先来回顾一下行列式的定义。

## 3.1.  行列式的定义

行列式是由一些数据排列成的方阵经过规定的计算方法而得到的一个数。当然，如果行列式中含有未知数，那么行列式就是一个多项式。它本质上代表一个数值，这点请与矩阵区别开来。矩阵只是一个数表，行列式还要对这个数表按照规则进一步计算,最终得到一个实数、复数或者多项式。

行列式分阶，比如二阶行列式、三阶行列式直至 n 阶行列式。下面我们罗列了各阶行列式的定义（以拉普拉斯展开定理的形式给出了定义，这样可以使各阶行列式看起来有规律），以方便后面的论述：

- 一阶行列式：



$$|a_1| = a_1$$

一阶行列式就等于元素 $a_1$ 自己。各位看官注意啊，$a_1$ 值可正可负，行列式两条竖线的记号不要当成绝对值的符号。

- 二阶行列式

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 \cdot |b_2| - a_2 \cdot |b_1| = a_1 b_2 - a_2 b_1$$

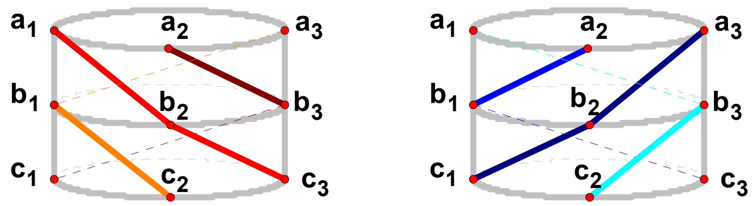
这个结果是不是很面熟？有点像三维向量叉积的第三个元素。你看，$\mathbf{a} \times \mathbf{b} = (a_1, a_2, a_3) \times (b_1, b_2, b_3) = (a_2 b_3 - a_3 b_2\text{，} a_3 b_1 - a_1 b_3\text{，} \underline{a_1 b_2 - a_2 b_1})$。注意叉积是和有向面积联系在一起的。

- 三阶行列式

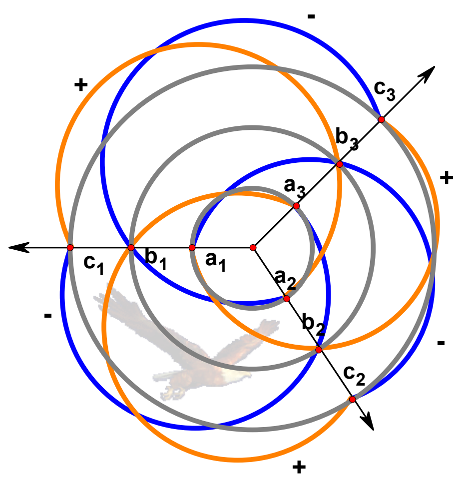
$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \cdot \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \cdot \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \cdot \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

$$= a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1$$

三阶行列式的计算比较常用，务必记住上式。上式帮助记忆之经典的展开运算法是对角线法，这个大家都知道，不再多讲。但为使对角线法看起来更有规律，这里稍稍改变了一下，图如下：




把行列式的元素依序排在圆柱体的外圆上，沿右下方向的乘积顺序得到正符号（左图），沿左下方向的乘积顺序得到负符号（右图）。如果把圆柱体拓扑变形为圆锥体，它的顶视图如下，运算顺序很有规律和美感，如下：




图中，行向量沿圆的逆时针方向排列元素，列向量沿原的半径方向排列。乘积项的元素分布在各个半圆弧上。

- 四阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12}\cdot\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13}\cdot\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14}\cdot\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}$$

$$\begin{aligned} = &+a_{11}a_{22}a_{33}a_{44} - a_{11}a_{22}a_{34}a_{43} + a_{11}a_{23}a_{34}a_{42} \\ &-a_{11}a_{23}a_{32}a_{44} + a_{11}a_{24}a_{32}a_{43} - a_{11}a_{24}a_{33}a_{42} \\ &-a_{12}a_{21}a_{33}a_{44} + a_{12}a_{21}a_{34}a_{43} - a_{12}a_{23}a_{34}a_{41} \\ &+a_{12}a_{23}a_{31}a_{44} - a_{12}a_{24}a_{31}a_{43} + a_{12}a_{24}a_{33}a_{41} \\ &+a_{13}a_{21}a_{32}a_{44} - a_{13}a_{21}a_{34}a_{42} + a_{13}a_{22}a_{34}a_{41} \\ &-a_{13}a_{22}a_{31}a_{44} + a_{13}a_{24}a_{31}a_{42} - a_{13}a_{24}a_{32}a_{41} \\ &-a_{14}a_{21}a_{32}a_{43} + a_{14}a_{21}a_{33}a_{42} - a_{14}a_{22}a_{33}a_{41} \\ &+a_{14}a_{22}a_{31}a_{43} - a_{14}a_{23}a_{31}a_{42} + a_{14}a_{23}a_{32}a_{41} \end{aligned}$$

可惜，在四阶以上的行列式中，没有了以上三阶行列式的特有美感的运算图解。这说明行列式的几何本质远没有如三阶行列式的图解一样简单。再者，手算四阶行列式很是恐怖，在 MATLAB 横行的年代也没此必要。把四阶行列式的展开算式罗列在此，只是给诸位一个感性认识。

- N 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13}\dots a_{1n} \\ a_{21} & a_{22} & a_{23}\dots a_{2n} \\ \dots\dots & \dots\dots & \dots\dots \\ a_{n1} & a_{n2} & a_{n3}\dots a_{nn} \end{vmatrix} = a_{11}\cdot\begin{vmatrix} a_{22} & a_{23}\dots a_{2n} \\ \dots\dots & \dots\dots \\ a_{n2} & a_{n3}\dots a_{nn} \end{vmatrix} - a_{12}\cdot\begin{vmatrix} a_{21} & a_{23}\dots a_{2n} \\ \dots\dots & \dots\dots \\ a_{n1} & a_{n3}\dots a_{nn} \end{vmatrix} + \dots + (-1)^{n+1}a_{1n}\cdot\begin{vmatrix} a_{21} & a_{22} & \dots a_{2n-1} \\ \dots\dots & \dots\dots & \dots\dots \\ a_{n1} & a_{n2} & \dots a_{nn-1} \end{vmatrix}$$

$$= \sum_{(j_1 j_2 \cdots j_n)} (-1)^t a_{1j_1} a_{2j_2} \dots a_{nj_n}.$$

式中，$j_1 j_2 \dots j_n$ 是 $1,2,\dots n$ 的一个排列，$\sum\limits_{(j_1 j_2 \cdots j_n)}$ 表示对 $j_1 j_2 \dots j_n$ 取遍 $1,2,\dots n$ 的一切 $(j_1 j_2 \dots j_n)$ 排列求和，$t$ 为排列 $j_1 j_2 \dots j_n$ 的逆序数。

N 阶的行列式的展开有 $n!$ 个乘积项，这些乘积项具有统一的表达式，因而具有统一的规律。通过观察以上各阶行列式的定义式，我们至少看到有两个小规律。一个规律是 n 阶行列式可以化为更低一阶的 n-1 阶行列式的和，这将会帮助我们简化行列式的计算。另一个规律是 n 阶行列式实际上是不同行不同列的 n 个元素的乘积项的代数和（和或差，每项符号由置换的逆序数的奇偶性决定）。在前面的向量一章中，我们介绍了向量的张量积，现在看来，行列式实际上是 n 个行向量的张量积中的部分结果。这个部分项的和具有独立的代数及几何意义，因而在线性代数中占有一席之地。

行列式的几何意义是什么呢？概括说来有两个解释：一个解释是行列式就是行列式中的行或列向量所构成的超平行多面体的**有向面积或有向体积**；另一个解释是矩阵 A 的行列式 detA 就是线性变换 A 下的图形面积或体积的**伸缩因子**。

这两个几何解释一个是静态的体积概念，一个是动态的变换比例概念。但具有相同的几何本质，因为矩阵 A 表示的(矩阵向量所构成的)几何图形相对于单位矩阵 E 的所表示的单位面积或体积(即正方形或正方体或超立方体的容积等于 1）的几何图形而言，伸缩因子本身就是矩阵矩阵 A 表示的几何图形的面积或体积，也就是矩阵 A 的行列式。

一个 2×2 矩阵 A 的行列式，是 A 的行向量（或列向量）决定的平行四边形的有向面积。

用几何观点来看，二阶行列式 $D=\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$ 是xoy平面上以行向量 $\mathbf{a}=(a_1\ a_2)$，$\mathbf{b}=(b_1\ b_2)$ 为邻边的平行四边形的有向面积：若这个平行四边形是由向量 $\mathbf{a}$ 沿逆时针方向转到 $\mathbf{b}$ 而得到的，面积取正值；若这个平行四边形是由向量 $\mathbf{a}$ 沿顺时针方向转到 $\mathbf{b}$ 而得到的，面积取负值；若 $\mathbf{a},\mathbf{b}$ 与 $\mathbf{a}',\mathbf{b}'$ 张成的平行四边形的有向面积符号相同，称他们有相同定向。所以平面上平行四边形有两种定向。

类似地，三阶行列式的值就是它的三个向量在Oxyz空间上张成的平行六面体的有向体积。这里空间平行六面体也有两种定向：当 $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成右手系时，体积正值；当 $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成左手系时，体积

取负值。这同时启发我们可以把n阶行列式定义为n个n维平行多面体的有向容积。

关于伸缩因子的几何解释，需要引入矩阵的概念。行列式被看作对矩阵的某种运算，并反映了矩阵的特定性质。实际上也是这样。

假设 A 是一个列向量（或行向量）为 $\mathbf{a},\mathbf{b}$ 的 2×2 矩阵。那么，这里的线性变换 A 是指将 $R^2$ 中的单位正方形变成 $R^2$ 中以 $\mathbf{a},\mathbf{b}$ 为邻边的平行四边形；如果原图形为一个园，则线性变换 A 将之变成一个椭圆。

同样，在 3×3 的情形下，A 将 $R^3$ 中的一个单位立方体映射成 $R^3$ 中由 A 的列向量确定的平行六面体；如果原图形为一个球，则线性变换 A 将之变成一个椭球。

一般地，一个 n×n 矩阵 A 将 $R^n$ 中的单位 n 立方体变成 $R^n$ 中由 A 列向量确定的 n 维平行体。对非单位正方形(立方体或超立方体)以同样的方式变换，即伸缩因子为

$$\frac{\text{像域的容积}}{\text{原域的容积}}$$

而 n×n 矩阵 A 的行列式 detA 就是这个伸缩因子。

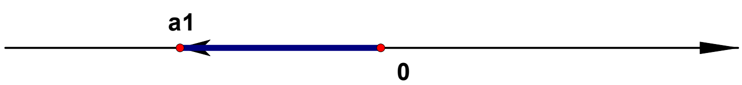
下面我们主要从向量及其张成的面积和体积的几何图像的角度分别就二阶和三阶行列式给出其几何意义，为了较深入的理解其几何含义，我们逐个地解释了行列式的主要性质。(从矩阵的线性变换的角度讲解行列式的伸缩因子的几何意义的内容放在了矩阵一章“矩阵的行列式”一节里面)。

# 3.2. 二阶行列式的几何意义

## 二阶行列式的几何意义

顺便先把一阶行列式的意义说一下。

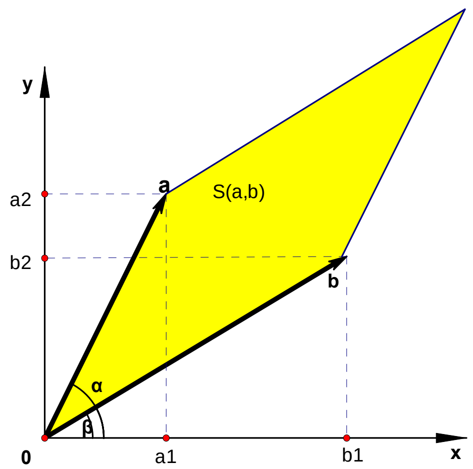
一阶行列式 $|a_1| = a_1$。意思就是 $a_1$ 的一阶行列式就是数 $a_1$ 或者讲是向量 $a_1$ 的本身，这个数 $a_1$ 的本身是一维坐标轴上的有向长度。这里我强调的是有向的，长度是有向的，是个向量，这一直是个很重要的概念。




再回来，我们继续讨论二阶行列式。

二阶行列式 $D=\begin{vmatrix}a_1,a_2\\b_1,b_2\end{vmatrix}$ 的几何意义是xoy平面上以行向量 $\mathbf{a}=(a_1\ a_2)$，$\mathbf{b}=(b_1\ b_2)$ 为邻边的平行四边形的有向面积。为什么？其实，我们可以推导出来这个几何意义的：

我们来考察这个平行四边形与构成它的两个向量之间的关系。




我们在二维几何空间 $R^2$ 中取定一个直角坐标系 $[0;\mathbf{e}_1;\mathbf{e}_2]$，设 $\mathbf{a}=a_1\mathbf{e}_1+a_2\mathbf{e}_2$，$\mathbf{b}=b_1\mathbf{e}_1+b_2\mathbf{e}_2$，则以 $\mathbf{a},\mathbf{b}$ 为边的平行四边形的面积为：

$$S(\mathbf{a},\mathbf{b})=abSin\langle\mathbf{a},\mathbf{b}\rangle$$

这里：$a=\sqrt{a_1^2+a_2^2}$，$b=\sqrt{b_1^2+b_2^2}$，$Sin\langle\mathbf{a},\mathbf{b}\rangle$ 为向量 $\mathbf{a},\mathbf{b}$ 之间的夹角正弦。

$$Sin\langle\mathbf{a},\mathbf{b}\rangle=Sin(\alpha-\beta)=Sin\alpha Cos\beta-Cos\alpha Sin\beta\text{，}$$

参照图中的关系把三角式用坐标值表示出来：

则

$$Sin\langle\mathbf{a},\mathbf{b}\rangle=\frac{b_2}{b}\cdot\frac{a_1}{a}-\frac{b_1}{b}\cdot\frac{a_2}{a}=\frac{a_1b_2-a_2b_1}{ab}$$

把上式整理得：$abSin\langle\mathbf{a},\mathbf{b}\rangle=a_1b_2-a_2b_1$。

又

$$\begin{vmatrix}a_1,a_2\\b_1,b_2\end{vmatrix}=a_1b_2-a_2b_1$$

因此

$$S\left(\mathbf{a},\mathbf{b}\right)=\begin{vmatrix}a_1,a_2\\b_1,b_2\end{vmatrix}$$

至此可以得到，二阶行列式的几何意义就是由行列式的向量所张成的平行四边形的面积。另外，两个向量的叉积也是这个公式。在向量一章中向量叉积的一个公式就是

$$\mathbf{a}\times\mathbf{b}=(ab\sin\theta)\mathbf{n}^0$$

这里，$\mathbf{n}^0$ 是垂直于 $\mathbf{a}$ 和 $\mathbf{b}$ 展成的平面的单位向量。

因此，二阶行列式的另一个意义就是是两个行向量或列向量的叉积的数值，这个数值是 z 轴上（在二维平面上，z 轴的正向想象为指向读者的方向）的叉积分量。如果数值是正值，则与 z 坐标同向；负值就与 Z 坐标反向。如果我们不强调叉积是第三维的向量，也就是忽略单位向量 $\mathbf{n}^0$，那么二阶行列式就与两个向量的叉积完全等价了。

## 二阶行列式性质的几何解释

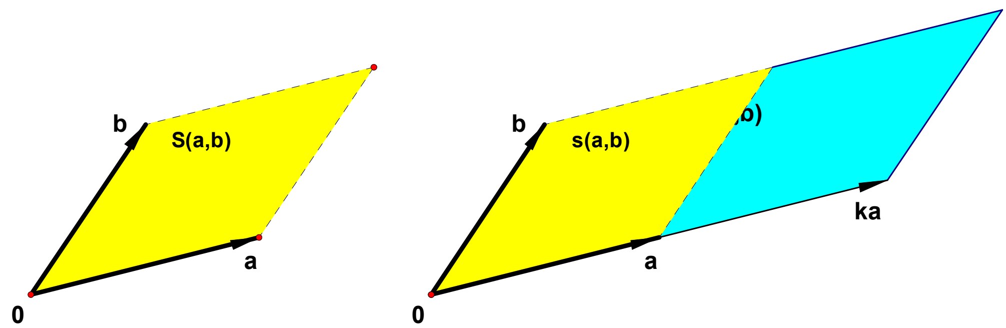
下面，我们仍然从向量的角度来解释或证明二阶行列式的几个主要性质。

**性质 1** $k\begin{vmatrix}a_{1,}a_2\\b_1,b_2\end{vmatrix}=\begin{vmatrix}ka_1,ka_2\\b_2,b_2\end{vmatrix}$，k 为实数；

这个性质是说，一个实数乘以行列式等于一个行向量乘以这个实数的行列式。几何解释就是

两个行向量 $\mathbf{a}=(a_1,a_2),\mathbf{b}=(b_1,b_2)$ 所张成的平行四边形的有向面积的 k 倍面积等于这样两个向量 $k\mathbf{a}=(ka_1,ka_2),\mathbf{b}=(b_1,b_2)$ 所张成的平行四边形的面积，也就是 $S(k\mathbf{a},\mathbf{a})=kS(\mathbf{a},\mathbf{b})$ 。

通过下面的图例容易得到几何解释。



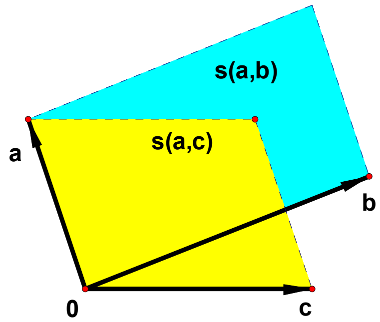

图中，$S(\mathbf{a},\mathbf{b})$ 可以看作以 $\mathbf{a}$ 为底的平行四边形的面积，$S(k\mathbf{a},\mathbf{b})$ 是以 $k\mathbf{a}$ 为底的平行四边形的

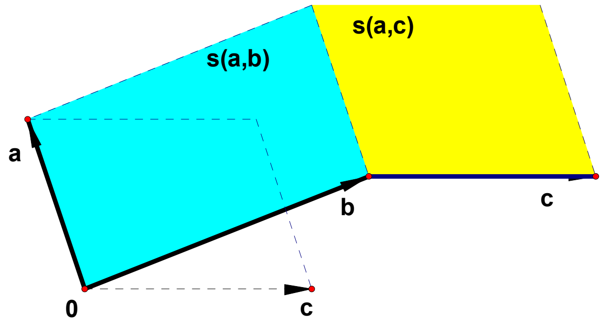
面积，高相同。因此，向量**a** 变化了 k 倍，面积也变化了 k 倍。

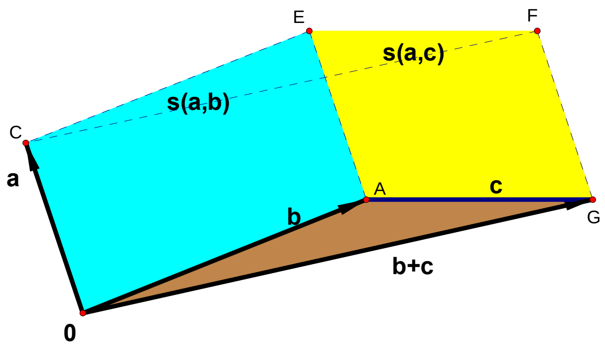
**性质 2** $\begin{vmatrix} a_1, a_2 \\ b_1+c_1, b_2+c_2 \end{vmatrix} = \begin{vmatrix} a_{1,} a_2 \\ b_1, b_2 \end{vmatrix} + \begin{vmatrix} a_{1,} a_2 \\ c_1, c_2 \end{vmatrix}$

对于三个向量 $\mathbf{a}=(a_1,a_2),\mathbf{b}=(b_1,b_2),\mathbf{c}=(c_1,c_2)$，那么向量 **a** 和 **b** 张成的平行四边形有向面积与向量 **a** 和 **c** 张成的平行四边形有向面积之和等于向量 **a** 和 **b+c** 张成的平行四边形有向面积，即 $S(\mathbf{a},\mathbf{b}+\mathbf{c})=S(\mathbf{a},\mathbf{b})+S(\mathbf{a},\mathbf{c})$。

下面的三个图形对此进行了图解说明。




看图之温馨提示 我们在“向量叉积的分配律的几何解释 2”中应用有向面积的概念解释了叉积的分配率。因为二阶行列式等同于叉积运算，因此具有类同的几何意义，它们都是一个意义：有向面积满足三角形法则。

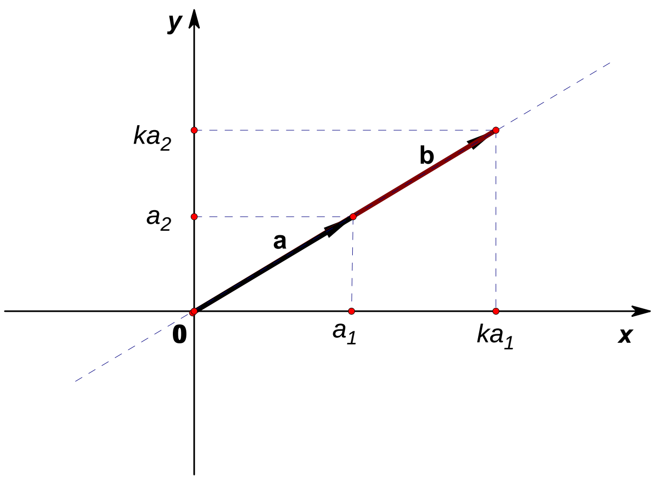
**性质 3** $\begin{vmatrix} a_{1,} a_2 \\ ka_1, ka_2 \end{vmatrix} = 0$

此行列式是说两行对应元素成比例，则行列式为零。

对于两个向量 $\mathbf{a}=(a_1,a_2),\mathbf{b}=(ka_1,ka_2)=k(a_1,a_2)$，显然成比例，比例系数为 k。如果把成比

列的两个向量的始端都移动到原点，则两向量会在同一条直线上，显然围成的四边形的面积为零，即 $S(\mathbf{a},\mathbf{b})=0$，因此行列式为零。

当然如果两个向量相同，同样道理，行列式的值也为零。



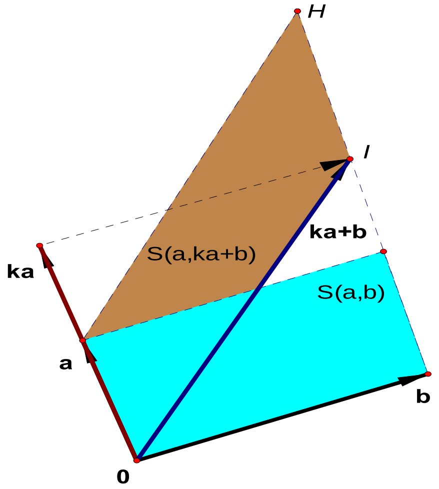

**性质 4** $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix} = -\begin{vmatrix} b_1, b_2 \\ a_1, a_2 \end{vmatrix}$

交换行列式的两行则行列式换号。

这个性质由行列式的叉积特性得到，交换行列式的两行，就是改变了向量 **a** 和向量 **b** 的叉积顺序，根据 $\mathbf{a}\times\mathbf{b}=-\mathbf{b}\times\mathbf{a}$，因此行列式换号。由此我们得到一个印象：就是一个给定的行列式，它的行向量顺序也给定了，不能随意改变其顺序。

**性质 5** $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix} = \begin{vmatrix} a_1 & , & a_2 \\ b_1+\lambda a_1, & & b_2+\lambda a_2 \end{vmatrix}$

把行列式的一行的 k 倍加到另一行，则行列式值不变，即 $S(\mathbf{a},\mathbf{b})=S(\mathbf{a},k\mathbf{a}+\mathbf{b})$。




由图，两个平行四边形阴影的面积相等，因为这两个平行四边形同底（都等于 **a** ）同高。显然，把行列式的一行的 k 倍加到另一行的操作，相当于把原平行四边形在保持同底同高的情况下发生了切变。

**性质 6** $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix} = \begin{vmatrix} a_1, b_1 \\ a_2, b_2 \end{vmatrix}$

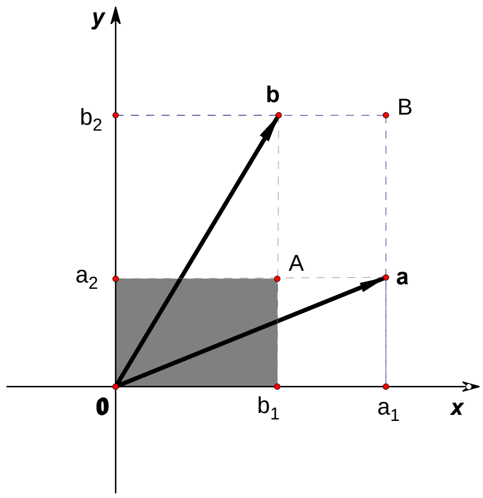
矩阵的行列式等于其转置矩阵的行列式。

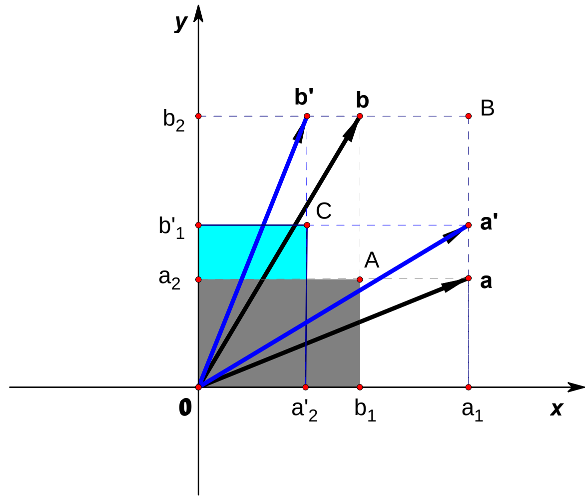
如果要讲清楚转置行列式的几何意义必须再一次使用行列式叉积的定义。另外，我们要回顾向量一章中两个向量的叉积的解析定义及意义。从前面的分析知道，两个向量的叉积等于这两个向量的各个分量（各坐标轴方向的分量）分别进行叉积的求和。

各个分量互相垂直，因而进行叉积运算张成的四边形是方形的面积。

对于二阶矩阵的行列式$\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$， $a_1\mathbf{i}$ 和 $b_2\mathbf{j}$ 叉乘得到的四方形 $Oa_1Bb_2$ 的有向面积是 $a_1b_2$，$a_2\mathbf{i}$ 和 $b_1\mathbf{j}$ 叉乘得到的 $Oa_1Bb_2$ 的有向面积是- $a_2b_1$（注意是负值！）。

所以矩阵的行列式$\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$ = “ $Oa_1Bb_2$ 的有向面积” + “ $Oa_1Bb_2$ 的有向面积” $= a_1b_2 - a_2b_1$。从几何图形上看，行列式等于大四方形的面积减去小四方形的面积（因为小四方形是负向面积）。




行列式$\begin{vmatrix}\mathbf{a}\\ \mathbf{b}\end{vmatrix}=\begin{vmatrix}a_1,a_2\\ b_1,\ b_2\end{vmatrix}$转置后得到$\begin{vmatrix}\mathbf{a}'\\ \mathbf{b}'\end{vmatrix}=\begin{vmatrix}a_1,b_1\\ a_2,b_2\end{vmatrix}$。重新对这个新的行列式进行分量叉积运算。

我们会发现，大四方形没有变化，而小四方形进行了基于 xy 角平分线的镜像变化。有向面积的绝对值和方向都没有变化。因而

$$\begin{vmatrix}\mathbf{a}\\ \mathbf{b}\end{vmatrix}=\begin{vmatrix}\mathbf{a}'\\ \mathbf{b}'\end{vmatrix}=a_1b_2-a_2b_1$$

性质得证。

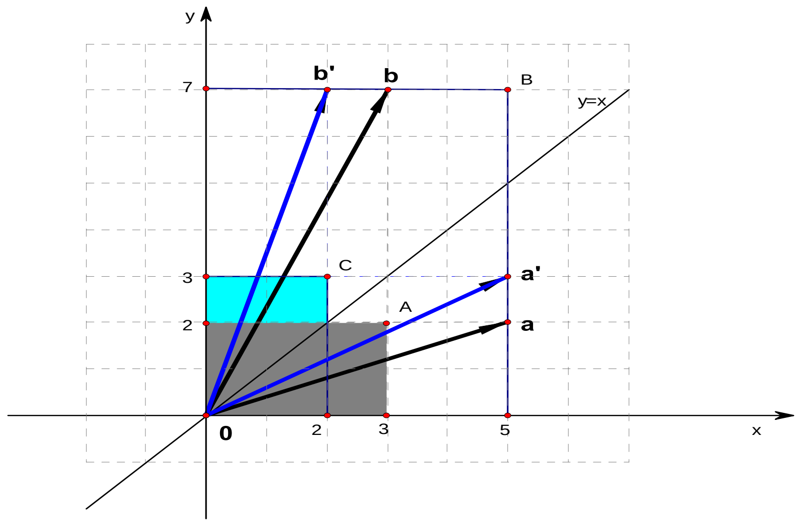
下面是上述转置行列式的具体数据，用于帮助读者得到具体感知。数据如下：

$$\begin{vmatrix}a_1,a_2\\ b_1,b_2\end{vmatrix}=\begin{vmatrix}5,2\\ 3,7\end{vmatrix}=5\times7-2\times3=29\text{。}$$

大长方形的面积是 35，小长方形的面积是 6，差为行列式的值。行列式转置后，

$$\begin{vmatrix}a_1,b_1\\ a_2,b_2\end{vmatrix}=\begin{vmatrix}5,3\\ 2,7\end{vmatrix}=5\times7-3\times2=29\text{。}$$

大长方形的面积是仍然是 35，小长方形经过基于 y=x 直线的镜像或者对折，面积仍然是 6，差为行列式的值不变。




矩阵的行列式等于其转置矩阵的行列式这条性质同时揭示了一个认识就是：按矩阵行向量构成的平行四边形的有向面积等于列向量构成的平行四边形的有向面积；换句话讲，对于矩阵的行列式几何意义，处理成行向量的图形与处理成列向量的图形是等效的。

总结一下，前面的行列式的性质在变换时总可以归结到下面的的三个变换性质：

（1）用一个数k乘以向量$\mathbf{a},\mathbf{b}$中之一的$\mathbf{a}$，则平行四边形的面积就相应地增大了k倍；

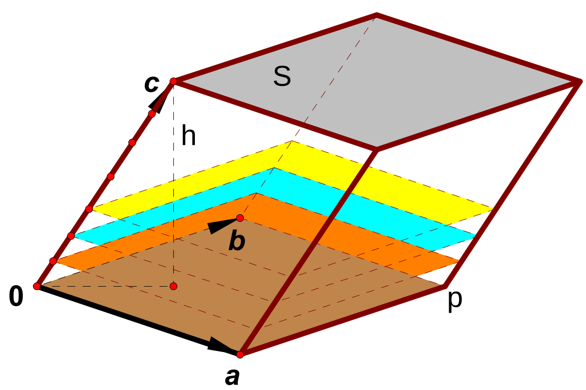
（2）把向量$\mathbf{a},\mathbf{b}$中的一个乘以数k之后加到另一个上，则平行四边形的面积不变；

（3）以单位向量(1，0)，(0，1)构成的平行四边形(即单位正方形)的面积为1。

这三条性质，便是我们用公理化的方法定义行列式的几何背景。

## 3.3. 三阶行列式的几何意义

一个 3×3 阶的行列式是其行向量或列向量所张成的平行六面体的有向体积。这个结论可以从两个向量所张成的平行四边形推知。如下图所示。




由两个向量$\mathbf{a},\mathbf{b}$张成的平行四边形为$0\mathbf{a}P\mathbf{b}$，面积 S 为$\mathbf{a},\mathbf{b}$构成的行列式。那么沿着第三个向

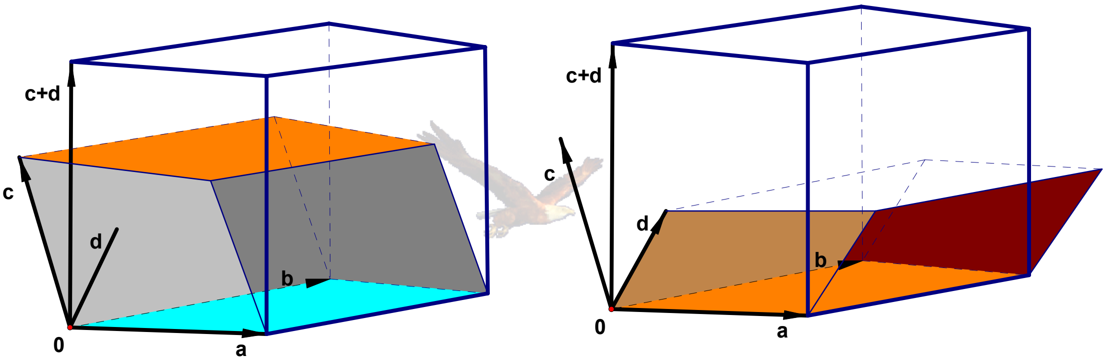
量 $\mathbf{c}$ 方向生长出无数个平行于原四边形的新的平行四边形来，直至到向量 $\mathbf{c}$ 的末端为止。显然，所有的这些平行四边形构成一个以向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体，这些四边形的面积叠加起来正是平行六面体的体积。

下面我们对 3×3 阵的行列式基本性质的几何意义进行解释。为了书写及描述方便，我们以 det（列向量）的方式表述三阶行列式。列向量用大写的 $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ 来表示。那么，$\det(\mathbf{a},\mathbf{b},\mathbf{c})=\begin{vmatrix} a_1,b_1,c_1 \\ a_2,b_2,c_2 \\ a_3,b_3,c_3 \end{vmatrix}$。

性质 1 $\det(\mathbf{a},\mathbf{b},\mathbf{c}+\mathbf{d})=\det(\mathbf{a},\mathbf{b},\mathbf{c})+\det(\mathbf{a},\mathbf{b},\mathbf{d})$

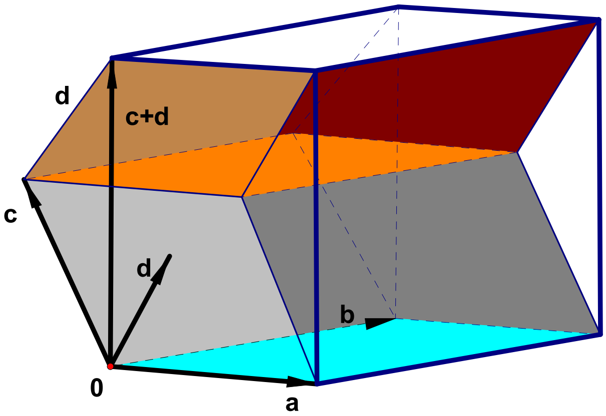
一个行列式可以通过拆分某一个列向量得到两个行列式的和。

看看这个性质的数学表达式，把 det 看作算子，有点像分配律的公式什么的。几何解释如下图示。

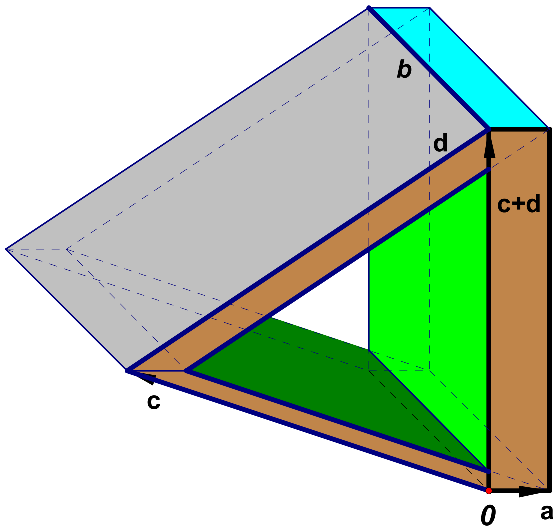



这里，把行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{c}+\mathbf{d})$ 的第三列拆分，变成两个行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$ 和 $\det(\mathbf{a},\mathbf{b},\mathbf{d})$。三阶行列式可以看作是平行六面体的有向体积。上图左表示由向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 张成的平行六面体代表行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$，上图右表示由向量 $\mathbf{a},\mathbf{b},\mathbf{d}$ 张成的平行六面体代表行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{d})$。这两个平行六面体共有一个底面积 $\det(\mathbf{a},\mathbf{b})$。这两个平行六面体的和就是由向量 $\mathbf{a},\mathbf{b},\mathbf{c}+\mathbf{d}$ 张成的粗实线平行六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c}+\mathbf{d})$。

实际上，我们把灰色的六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{d})$ 上移，摞在六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$ 上，刚好得到六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c}+\mathbf{d})$。显然，他们的棱向量 $\mathbf{c}$，$\mathbf{d}$，$\mathbf{c}+\mathbf{d}$ 满足向量和的三角形法则。

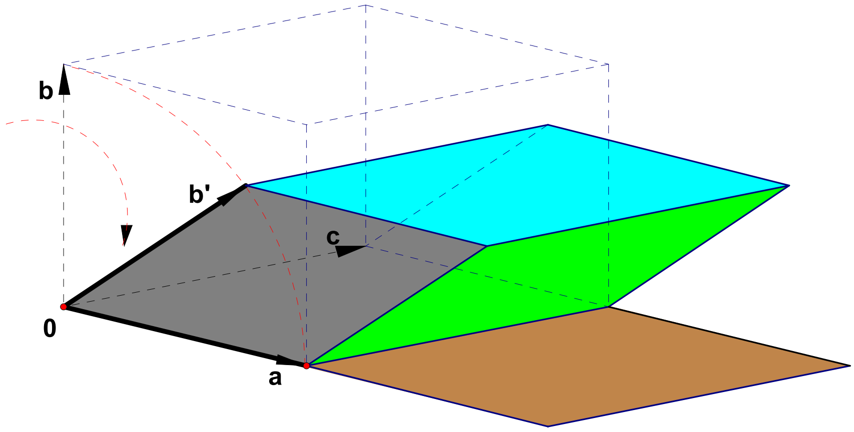



看图之温馨提示 和前面的有向面积满足三角形法则类似，这里的有向体积同样满足三角形法则。为了更清楚地表达有向面积的三角形法则，把上图变形，得到下面的明显的立方三角形图形。




**性质 2** $\det(\mathbf{a}, \mathbf{a}, \mathbf{c}) = 0$

行列式的有两行或者两列元素相同，它对应的空间平行六面体的两条邻边重合，相当于三维空间中六面体被压成了高度为零的二维平面，显然， 这个平面的三维体积 $\det(\mathbf{a}, \mathbf{a}, \mathbf{c})$ 为零。如下图所示。

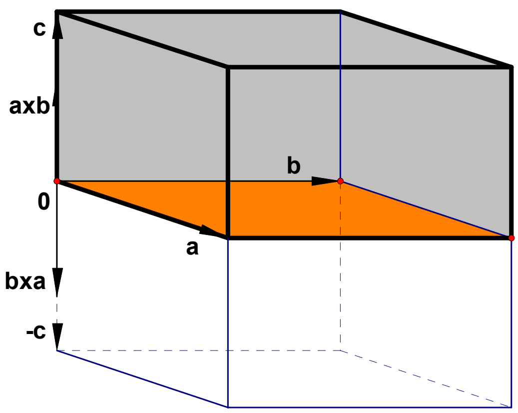



上图中，我们把平行六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$ 中的棱向量 $\mathbf{b}$ 以坐标原点为轴沿一弧线下压（六面保持对应面平行）切变为 $\det(\mathbf{a},\mathbf{b}',\mathbf{c})$，显然六面体的高度变小了，行列式值变小了。继续下压，直到 $\mathbf{b}'$ 重合与 $\mathbf{a}$，即变成了 $\det(\mathbf{a},\mathbf{a},\mathbf{c})$，平行六面体变成了一个平行四边形平面。

**性质 3** $\det(\mathbf{a},\mathbf{b},\mathbf{c}) = -\det(\mathbf{b},\mathbf{a},\mathbf{c})$

一个行列式对应着一个数值，这个数值是对行列式中的元素经过运算得到的。这个运算是与元素的位置有关系的，因此你改变了行列式中列向量或行向量的位置当然会改变行列式的结果。幸而只改变结果的符号。一般地，一个行列式的值对应矩阵 A 的列向量的一个固定顺序。当 detA 为负值时，它确定原象的一个反射。所以，这种变换改变了原象的定向。

实际上，$\det(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} = -(\mathbf{b}\times\mathbf{a})\cdot\mathbf{c} = -\det(\mathbf{b},\mathbf{a},\mathbf{c}) = \det(\mathbf{b},\mathbf{a},-\mathbf{c})$，因此如果交换了向量 $\mathbf{a},\mathbf{b}$ 的位置，也就是改变了叉乘的顺序，因此叉乘结果改变了方向，也即改变符号。

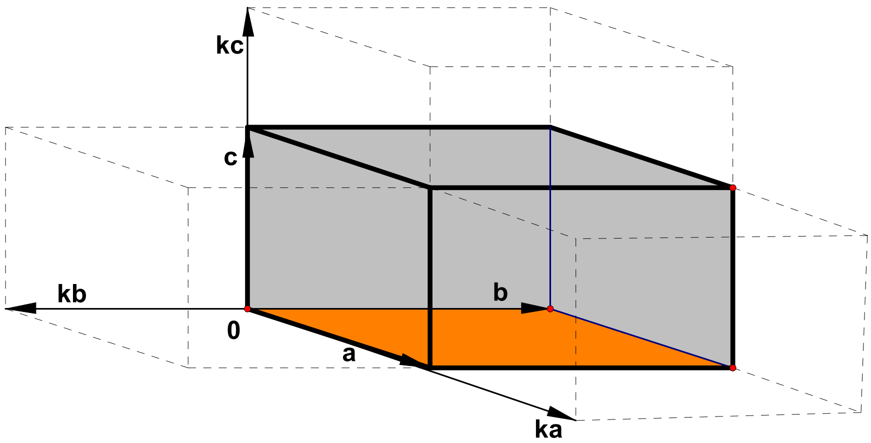



如图，阴影的平行六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$，$\mathbf{a}\times\mathbf{b}$ 的方向与向量 $\mathbf{c}$ 同向（据右手法则）；当向量 $\mathbf{a},\mathbf{b}$ 的位置交换后，$\mathbf{b}\times\mathbf{a}$ 的方向与 $\mathbf{a}\times\mathbf{b}$ 相反，因而与向量 $\mathbf{c}$ 点乘后得到向下的平行六面体。所以平行六

面体 $\det(\mathbf{b},\mathbf{a},\mathbf{c})$ 和 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$ 以 $\mathbf{a},\mathbf{b}$ 张成的平面为镜面互为反射。

**性质 4**　$k\det(\mathbf{a},\mathbf{b},\mathbf{c})=\det(k\mathbf{a},\mathbf{b},\mathbf{c})=\det(\mathbf{a},k\mathbf{b},\mathbf{c})=\det(\mathbf{a},\mathbf{b},k\mathbf{c})$

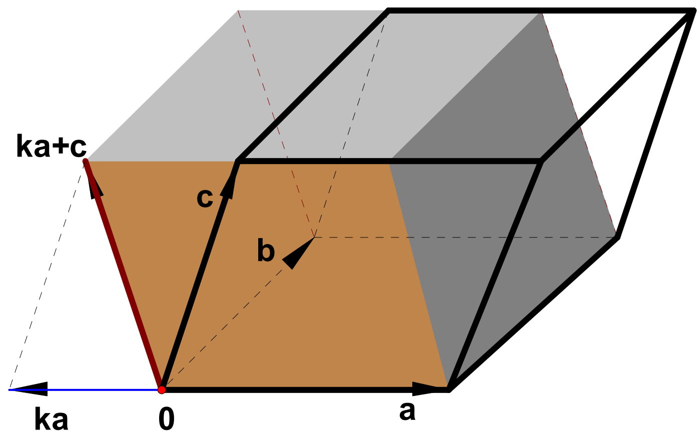
这就是说，平行六面体的体积的 k 倍等于六面体的三条棱中一条棱长的 k 倍。这是显然的。因为立方体的体积增大可以沿着立方体某一棱方向增大相同的倍数。如下图：



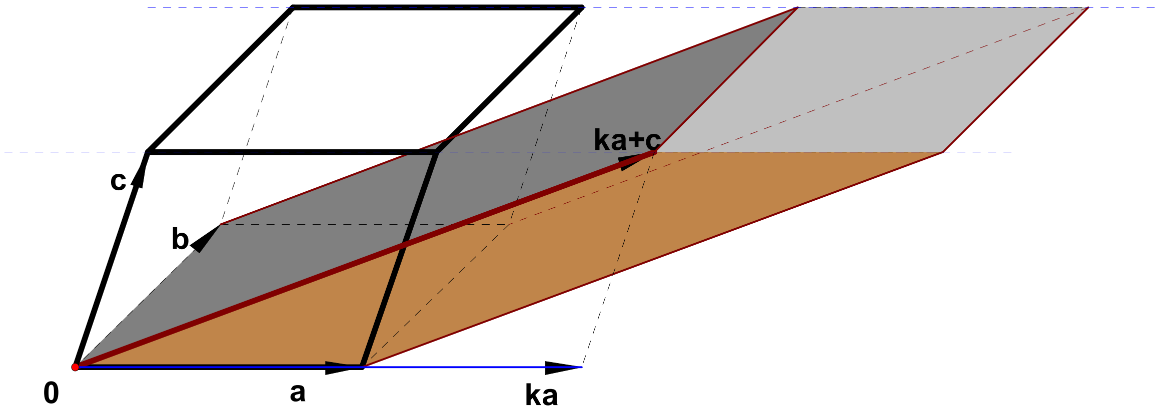

**性质 5**　$\det(\mathbf{a},\mathbf{b},\mathbf{c})=\det(\mathbf{a},\mathbf{b},k\mathbf{a}+\mathbf{c})$

此性质表述了以 $\mathbf{a}\times\mathbf{b}$ 为底面积的平行六面体在 $\mathbf{a}$ 方向上进行了切向变换，变换的后的六面体因为底面积不变，高也不变，因此体积不变。

下图中，原有六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$，是由向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 张成。向量 $\mathbf{a}$ 乘以一个负值的 k 值后与向量 $\mathbf{c}$ 相加后得到新的向量 $k\mathbf{a}+\mathbf{c}$，三个向量 $\mathbf{a},\mathbf{b},k\mathbf{a}+\mathbf{c}$ 构成了一个新的平行六面体 $\det(\mathbf{a},\mathbf{b},k\mathbf{a}+\mathbf{c})$，这个六面体与原六面体 $\det(\mathbf{a},\mathbf{b},\mathbf{c})$ 同底等高，因而体积相同。



右图向量 $\mathbf{a}$ 乘以一个正值的 k 值后与向量 $\mathbf{c}$ 相加的结果，与作图类似。




通过观察，我们发现，切变后的平行六面体与 k 值无关。K 值不同，向量 $k\mathbf{a}+\mathbf{c}$ 终端在始终在一条与向量 $\mathbf{a}$ 平行的直线上滑动，因而保持了六面体的等高。

**性质 6** $\det\mathbf{A}=\det\mathbf{A}'$

矩阵 A 的行列式等于矩阵 A 转置的行列式。

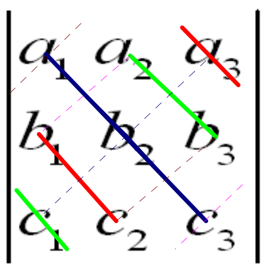
为了便于讨论，我们把三阶行列式及其转置行列式的按照第一行元素展开式列举在这里：

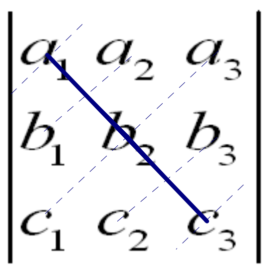
$$\det(A)=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=a_1\cdot\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-a_2\cdot\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}+a_3\cdot\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}=a_1b_2c_3+a_2b_3c_1+a_3b_1c_2-a_1b_3c_2-a_2b_1c_3-a_3b_2c_1$$

$$\det(A')=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}=a_1\cdot\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}-b_1\cdot\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}+c_1\cdot\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}=a_1b_2c_3+b_1c_2a_3+c_1a_2b_3-a_1c_2b_3-b_1a_2c_3-c_1b_2a_3$$

显然，上述两式是相等的。这个相等有里外两层几何意义可以解释。先说外层的几何解释：

对照一下这两个图片：



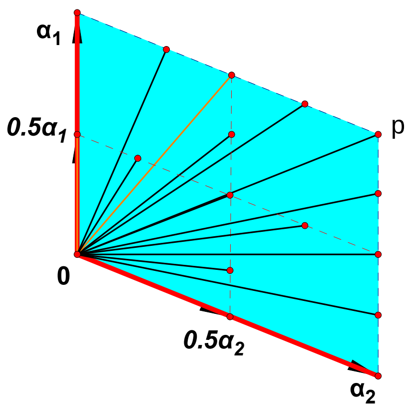



左图是行列式展开时元素相乘的计算图，右图是行列式转置时以 $a_1b_2c_3$ 为反射轴的元素交换图。行列式转置顺序与计算顺序多么一致，怪不得不改变结果呢。

行列式和它的转置相等还有一个深层次的几何意义，可以在下面逆序数的几何意义中得到体现。行列式的乘积项及其逆序数的几何意义实际上是行列式最根本的几何意义，因而可以解释所有的行列式的定义及其性质。后面的章节我们会逐步探讨到这个重要的课题。

：什么是向量张成的平行四边形及平行六面体？

以两个从原点出发的向量为邻边可以画出一个平行四边形；以三个从原点出发的不同方向向量为相邻的棱可以画出一个平行六面体，这个平行六面体与立方体仿射等价。




上图中是以两个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 张成的平行四边形 $0\boldsymbol{\alpha}_1 P\boldsymbol{\alpha}_2$，在这个平行四边形中有无数的向量，这些向量统统满足线性组合 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 \quad (0 \le k_1, k_2 \le 1)$ 的定义。比如，对角线向量 0P 是最大的向量，满足 $k_1 = k_2 = 1$ 的条件；$\boldsymbol{\alpha}_1 P$ 上的向量满足 $k_1 = 1, \ 0 \le k_2 \le 1$ 的条件等等。

我们可以推广到 n 维欧几里德空间中的 m 个向量张成的 m 维超平行多面体。如果这 m 个向量表示为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \ldots \boldsymbol{\alpha}_m$，那么这个 m 维超平行多面体是由以下无穷个向量

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 \ldots + k_m\boldsymbol{\alpha}_m \qquad (0 \le k_1, k_2 \ldots k_m \le 1)$$

组成的。

在后面的章节中，我们还要讨论由向量张成的空间的概念。这时候，实数 $k_i$ 的范围不是在区间 $[0,1]$ 之上了，而是整个实数域 $R$，因此空间的范围也将变得无穷大的了。

## 3.4. 行列式化为对角形的几何解释

一个行列式的第 i 行加上 j 行的 K 倍，可以使第 i 行的某一个元素变为 0，而这个行列式的值不变。这个性质在化简行列式时非常有用。在前面的二阶和三阶行列式的性质中我们也已看到了它的

几何意义。不过前面对它的几何解释一般都是使用向量的“箭头”的图形形式给出的，没有考虑向量的元素的变化的几何意义。在本阶中我们就此性质的应用过程中来探究行列式中每一个元素的几何意义。

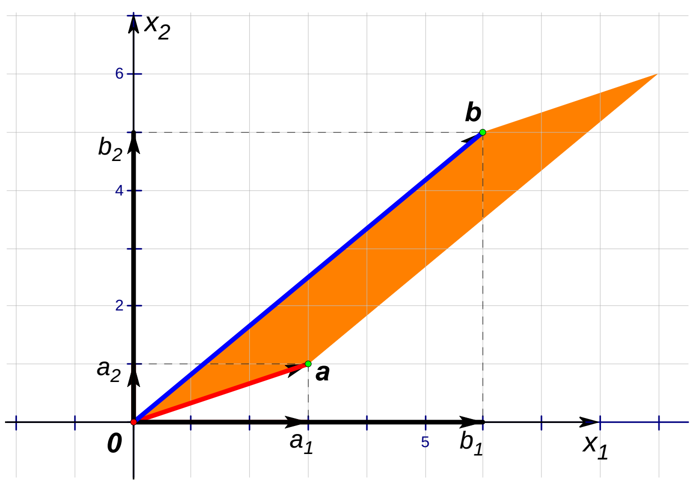
把一般二阶行列式化为上对角形式的行列式，只要把行列式第二行第一列化为 0 就可以了。因此，第一行向量元素$(a_1\ \ a_2)$乘以-$\frac{b_1}{a_1}$(设$a_1 \neq 0$)后加到第二行上，即可得到如下的三角式：

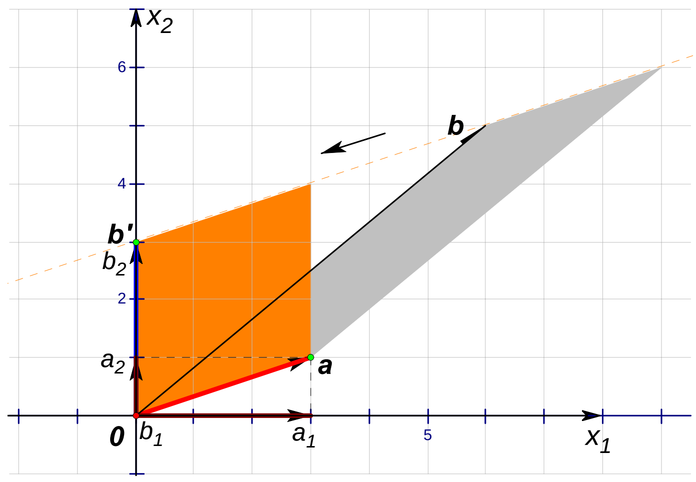
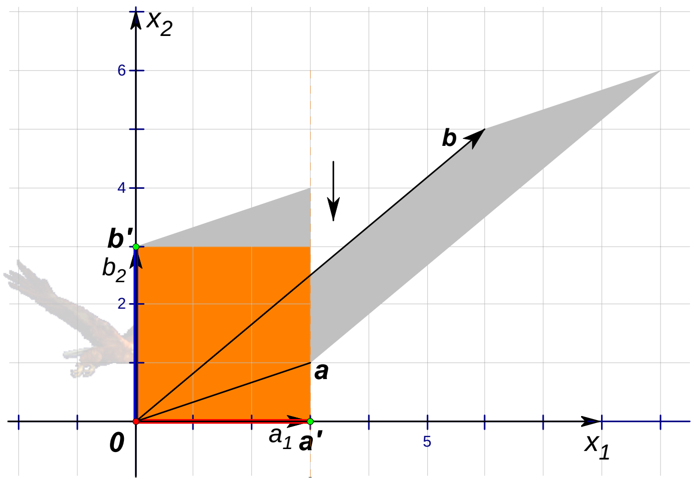
$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ 0 & \dfrac{a_1b_2 - a_2b_1}{a_1} \end{vmatrix}$$

继续化简，把第二行的元素$(0\quad \frac{a_1b_2 - a_2b_1}{a_1})$乘以$-\frac{a_1a_2}{a_1b_2 - a_2b_1}$后加到第一行上，得到如下的对角形行列式：

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & 0 \\ 0 & \dfrac{a_1b_2 - a_2b_1}{a_1} \end{vmatrix}$$

这个化简的过程我们可以有一个很形象的几何变化过程。

| 顺序 | 描述 | 对应的几何图形 |
|---|---|---|
| 1. 1 | 一个二阶行列式$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$。<br>其值等于右图阴影平行四边形的有向面积。 |  |

| | | |
|---|---|---|
| 1.2 | 化简到如下的三角式：<br>$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ 0 & \dfrac{a_1b_2-a_2b_1}{a_1} \end{vmatrix}$。<br>其几何变化过程是向量 **b** 沿着 0**a** 的平行线 **b'b**，滑行到 $x_2$ 坐标轴上，此时 $b_1$ 化为 0。显然两个阴影平行四边形面积不变。 |  |
| 1.3 | 继续化简到对角形：<br>$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & 0 \\ 0 & \dfrac{a_1b_2-a_2b_1}{a_1} \end{vmatrix}$。<br>其几何变化过程是向量 **a** 沿着 0**b** 的平行线 **a'a**，滑行到 $x_1$ 坐标轴上，此时 $a_2$ 化为 0。显然三个阴影平行四边形面积相同。 |  |
| 1.4 | 至此，一个二阶行列式所表示的平行四边形被变成了一个对角行列式所表示的正（长）方形。 | |

三阶行列式有类似的变换情形，对角化的过程会把一个平行六面体变化为一个等体积的立方体或长方体。具体的图形不再绘出。

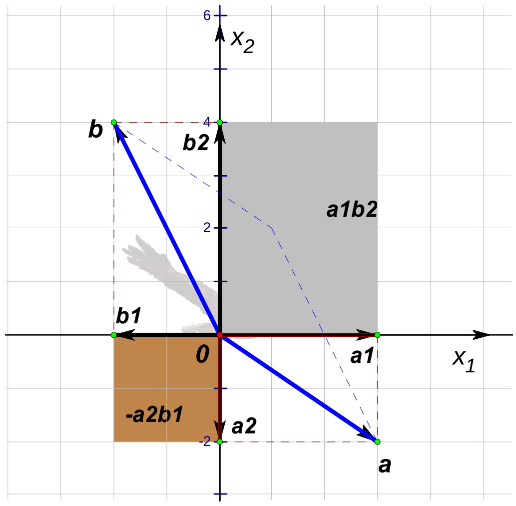
那么 n 阶行列式我们亦不怀疑的认为也可以被表示成一个 n 维的长方体的几何图形。这个结论对于我们理解行列式的展开式很有帮助。

## 3.5. 行列式乘积项的几何意义

n 阶行列式的展开式是乘积项 $(-1)^t a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ 的和，这些乘积项其实也可以有几何解释的。下面我们详细的了解一下各阶行列式的情况。

## 二阶行列式乘积项的几何意义

对于二阶行列式而言，既然二阶行列式的几何图形是一个有方向的面积，那么从二阶行列式公理化定义 $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix} = a_1b_2 - a_2b_1$ 看，又是如何构成这个面积的呢？显然，式中 $a_1b_2$ 项和 $-a_2b_1$ 项的和构成了这个面积。$a_1b_2$ 项和 $-a_2b_1$ 项又是什么呢？易知，$a_1$ 是向量 $\mathbf{a}$ 在 $x_1$ 轴上投影，$b_2$ 是向量 $\mathbf{b}$ 在 $x_2$ 轴上投影，$a_1b_2$ 项是 $\mathbf{a}$ 在 $x_1$ 轴上投影和 $\mathbf{b}$ 在 $x_2$ 轴上投影所围成的长方形面积。同样知道，$-a_2b_1$ 项是 $\mathbf{a}$ 在 $x_2$ 轴上投影和 $\mathbf{b}$ 在 $x_1$ 轴上投影所围成的长方形面积，不过这个面积表达式表现为负数的表达式。如下如图中长方形阴影的面积：




这里有两个问题要注意：

1、　因为二阶行列式表达的是向量 $\mathbf{a}$ 和 $\mathbf{b}$ 所围成的面积，因此向量 $\mathbf{a}$ 自己的分向量 $a_1$ 和 $a_2$ 围成的面积 $a_1a_2$ 不在统计范围之内，向量 $\mathbf{b}$ 自己的分向量围成的面积 $b_1b_2$ 也不在统计范围之内；

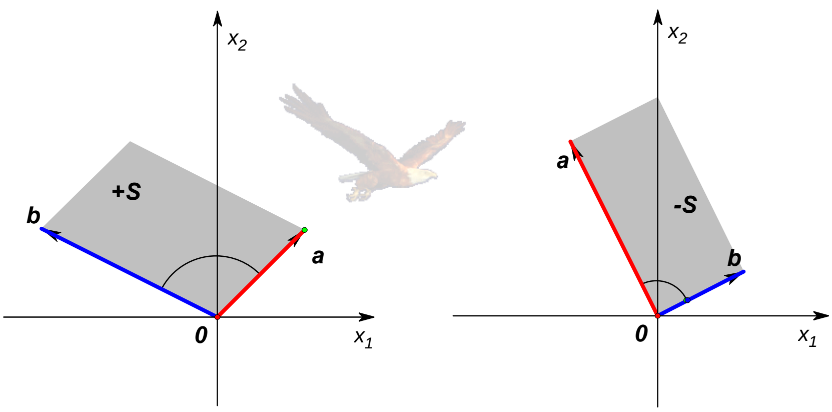
2、　向量 $\mathbf{a}$ 的分向量 $a_1$ 和向量 $\mathbf{b}$ 的分向量 $b_1$ 围成的面积 $a_1b_1$ 要统计，但可惜的是 $a_1b_1 = 0$，因为两个分向量投影到同一根坐标轴 $x_1$ 上。同样的理由，$a_2b_2 = 0$。

前面我们在解释二阶行列式的转置时讲过，两个向量的叉积等于这两个向量的各个分量（各坐标轴方向的分量）分别进行叉积的求和。各个分量互相垂直，因而进行叉积运算张成的四边形是方形的面积。向量 $\mathbf{a}, \mathbf{b}$ 的叉积（有向面积）等于每个向量的分向量所有可能的叉积之和。因为在直角坐标系下，向量 $\mathbf{a}, \mathbf{b}$ 落在同一坐标轴上的分量叉积显然为零。因此进行叉积的分量必然是不在的坐

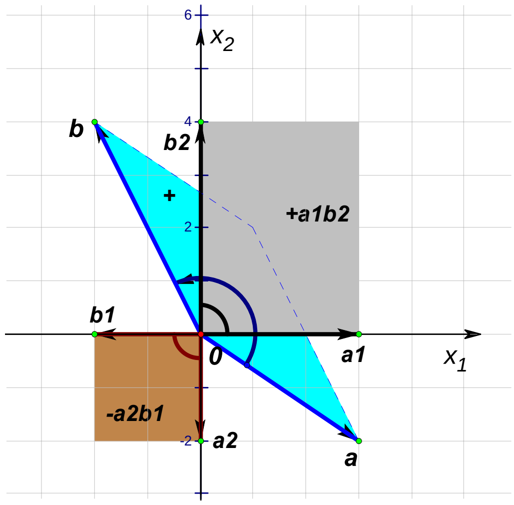
标轴上。

$a_1b_2$ 项和 $-a_2b_1$ 项都是有向面积，$a_1b_2$ 项的面积和此二阶行列式的面积方向相同，因此符号为正，$-a_2b_1$ 项的面积和此二阶行列式的面积方向相反，因此符号为负。

那么二阶行列式的面积及其乘积项的面积的方向如何确定呢？对二阶行列式来说，有不同的方法可以确定，我们这里使用叉积的右手法则来判断。这里，我们把行列式 $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$ 看作由两个行向量 $\mathbf{a}, \mathbf{b}$ 所组成，其中排在第一行的元素组成第一个向量 $\mathbf{a} = (a_1, a_2)$，排在第二行的元素组成第二个向量 $\mathbf{b} = (b_2, b_2)$，（看成两个列向量也有类似的结果，因为转置行列式值不变）。这是一个顺序，不要错了。那么行列式 $\begin{vmatrix} a_1, a_2 \\ b_1, b_2 \end{vmatrix}$ 的面积方向就是由第一向量转向第二向量时右手大拇指所确定的方向。




上图左，行列式的方向为指向读者；图右，行列式的方向指向页面内，背向读者。下图中，行列式的方向是指向读者，那么，分向量 $\mathbf{a_1}$ 和 $\mathbf{b_2}$ 张成的长方形（右上角图块）面积的方向也是指向读者的，因此面积为正，故记为 $a_1b_2$；分向量 $\mathbf{a_2}$ 和 $\mathbf{b_1}$ 张成的长方形（左下角图块）面积的方向指向页面内，背向读者，与行列式方向相反，因此面积为负，故记为 $-a_2b_1$。




## 三阶行列式乘积项的几何意义

依据三阶行列式的公式 $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1$，与二阶行列式的乘积项的几何解释类似，三阶行列式的乘积项如 $a_1b_2c_3$，$a_2b_3c_1$ 等，可以看成具有有方向的小长方体的体积。也就是说，在三阶方阵张成的三维平行六面体可以分解为一个个由各座标分量混合积构成的小长方体。这些小长方体共有六块，其体积具有方向。

既然平行六面体和小长方体都是具有方向的三维几何图形，那么这些方向是如何确定呢？与上一节的论述类似的，也是由张成这些三维几何图形的有序的三元向量组所决定的。三个无关向量 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的有序组根据右手法则可以确定一个方向。这个确定的方法类同于叉积定义的方法：

从向量 $\mathbf{a}$ 的方向转过一个 $0\sim\pi$ 之间的角度到达向量 $\mathbf{b}$ 的方向，右手拇指方向是向量 $\mathbf{a},\mathbf{b}$ 所确定平面的正向，如果向量 $\mathbf{c}$ 在此平面正向的一面，$\mathbf{a},\mathbf{b},\mathbf{c}$ 的有序组就确定了一个正的方向；如果向量 $\mathbf{c}$ 在正向平面的另一面，$\mathbf{a},\mathbf{b},\mathbf{c}$ 的有序组就确定了一个负的方向。和向量的叉积联系起来讲，第三个向量 $\mathbf{c}$ 与向量 $\mathbf{a}\times\mathbf{b}$ 在 $\mathbf{a},\mathbf{b}$ 所确定平面的同一侧时确定一个正的方向。

换一种说法就是，从向量 $\mathbf{c}$ 的的逆方向看过去，向量 $\mathbf{a}$ 的旋转过一个 $0\sim\pi$ 之间的角度到达向量 $\mathbf{b}$ 的方向符合逆时针方向。

因此，我们有一些结论：

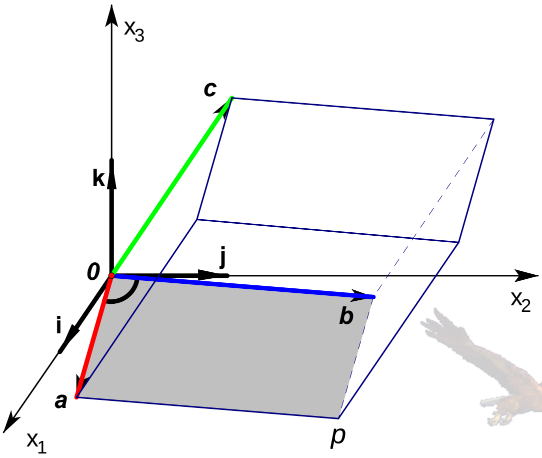
- $\mathbf{b},\mathbf{a},\mathbf{c}$ 有序组与 $\mathbf{a},\mathbf{b},\mathbf{c}$ 有序组具有相反的方向；

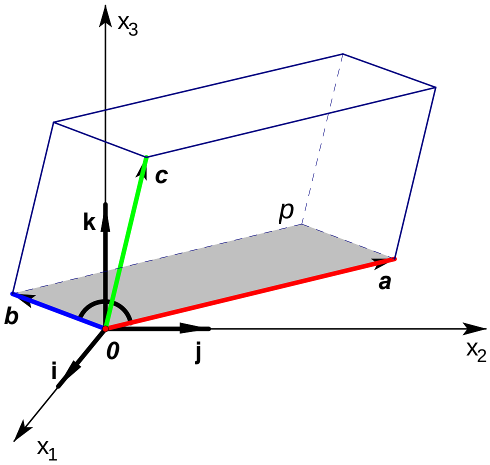
- $\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}$ 有序组与三维坐标向量的有序组 $\mathbf{i},\mathbf{j},\mathbf{k}$ 的方向永远相同；

- 如果 $\mathbf{a},\mathbf{b},\mathbf{c}$ 有序组与坐标向量有序组 $\mathbf{i},\mathbf{j},\mathbf{k}$ 有同样的定向，我们就称 $\mathbf{a},\mathbf{b},\mathbf{c}$ 有序组相对于坐标系 $(x_1,x_2,x_3)$ 有正向；如果有相反的方向，就称为负向；

- $\mathbf{a},\mathbf{b},\mathbf{c}$ 有序组有正向的充分必要条件是行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{c})>0$；

- 行列式 $\det(\mathbf{a},\mathbf{b},\mathbf{c})>0$ 意味着 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}>0$。

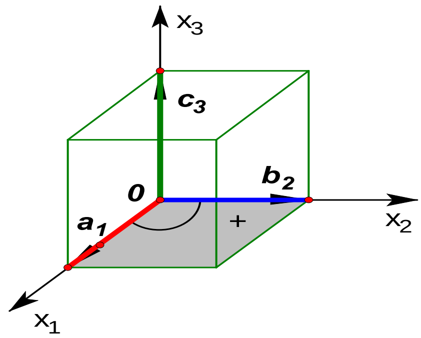
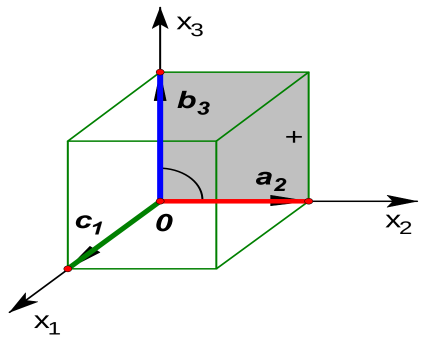
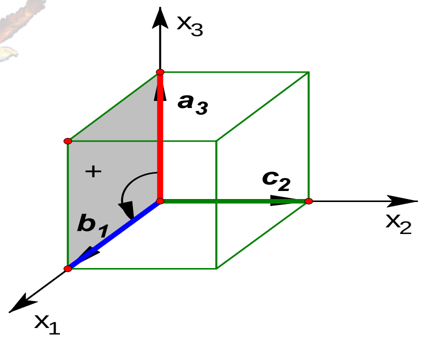
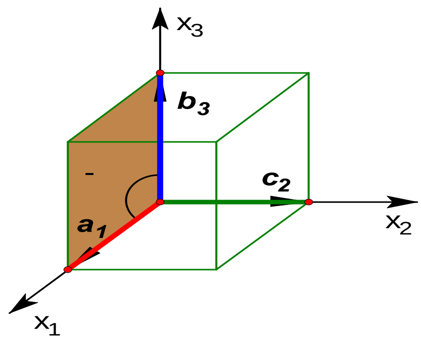


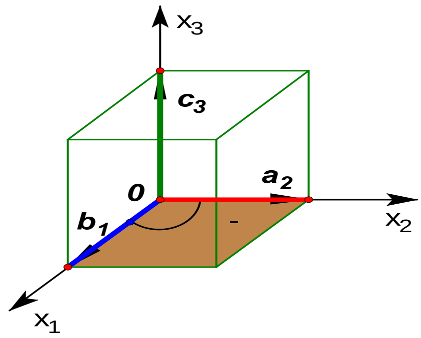
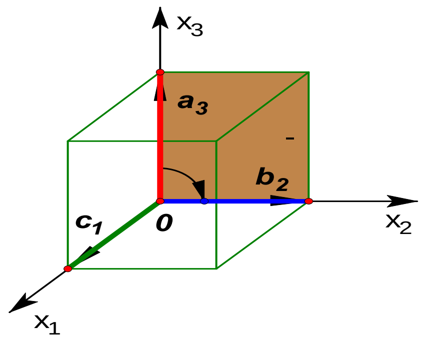




$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}>0$，也就是说，向量 $\mathbf{a}\times\mathbf{b}$ 的方向与向量 $\mathbf{c}$ 的方向作成一个锐角，也就是向量 $\mathbf{c}$ 与向量 $\mathbf{a}\times\mathbf{b}$ 在 $\mathbf{a},\mathbf{b}$ 所确定平面的同一侧,这和前面的解释是一致的。

同样，这些小长方体的方向也是由三个坐标轴上的向量投影所确定：与坐标向量有序组 $\mathbf{i},\mathbf{j},\mathbf{k}$ 有同样的定向，我们就称 $\mathbf{a},\mathbf{b},\mathbf{c}$ 有序组相对于坐标系 $(x_1,x_2,x_3)$ 有正向；如果有相反的方向，就称为负向。因为小长方体是由三个坐标轴上的三个不同向量的投影向量所张成，三个投影分量互相垂直，因此我们可以方便地使用右手法则确定其方向。如果是正向，则长方体体积的乘积项大于 0，否则小于 0。下面我们画出 6 个小长方体的方向确定的几何图形。

| 顺序 | 乘积项及其解释 | 对应的几何图形 |
|---|---|---|

| | | |
|---|---|---|
| 1．1 | $+a_1b_2c_3$ 项：<br>以右手法则，向量有序组<br>$(\mathbf{a}_1,\mathbf{b}_2,\mathbf{c}_3)$ 与坐标系方向一致，<br>乘积项为正。 |  |
| 1．2 | $+a_2b_3c_1$ 项：<br>以右手法则，向量有序组<br>$(\mathbf{a}_2,\mathbf{b}_3,\mathbf{c}_1)$ 与坐标系方向一致，<br>乘积项为正。 |  |
| 1．3 | $+a_3b_1c_2$ 项：<br>以右手法则，向量有序组<br>$(\mathbf{a}_3,\mathbf{b}_1,\mathbf{c}_2)$ 与坐标系方向一致，<br>乘积项为正。 |  |
| | $-a_1b_3c_2$ 项：<br>以右手法则，向量有序组<br>$(\mathbf{a}_1,\mathbf{b}_3,\mathbf{c}_2)$ 与坐标系方向相反，<br>乘积项为负。 |  |

| | | |
|---|---|---|
| | $-a_2b_1c_3$ 项:<br>以右手法则，向量有序组<br>$(\mathbf{a}_2,\mathbf{b}_1,\mathbf{c}_3)$ 与坐标系方向相反，<br>乘积项为负。 |  |
| | $-a_3b_2c_1$ 项:<br>以右手法则，向量有序组<br>$(\mathbf{a}_3,\mathbf{b}_2,\mathbf{c}_1)$ 与坐标系方向相反，<br>乘积项为负。 |  |

## n 阶行列式乘积项的几何意义

N 阶行列式的展开式是乘积项 $(-1)^t a_{1j_1}a_{2j_2}...a_{nj_n}$ 的和，$t$ 为排列 $j_1j_2...j_n$ 的逆序数。行列式的行向量在坐标系中向各个坐标轴的投影就是行列式中的元素 $a_{1j_1},a_{2j_2},...a_{nj_n}$，这些元素当然也是行向量在各个坐标轴上分向量。这些坐标轴上分向量也可以张成一个个小小的 n 维超长方体（共 $n!$ 个）。我们刚好知道，N 阶行列式的超平行多面体的几何图形是由行(或列)向量张成的，而且这个 n 维超平行多面体与一个 n 维超长方体等体积。那么大家就会容易得到：

- $n!$ 个小的 n 维长方体叠加成了大的表示 n 阶行列式的 n 维超长方体；
- $n!$ 个小的 n 维长方体就是 $n!$ 个乘积项 $(-1)^t a_{1j_1}a_{2j_2}...a_{nj_n}$ 的几何图形。

实际上，一个 n 阶行列式可以分拆成 $n!$ 个只有坐标轴分向量组成的 n 阶对角行列式，而 $n!$ 个 n 阶对角行列式就是 $n!$ 个乘积项 $(-1)^t a_{1j_1}a_{2j_2}...a_{nj_n}$。

比如一个二阶行列式可以分拆成两个这样的二阶对角行列式：

$$\begin{aligned}\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix} &= \begin{vmatrix}a_1 & 0\\ b_1 & b_2\end{vmatrix}+\begin{vmatrix}0 & a_2\\ b_1 & b_2\end{vmatrix}\\ &= \begin{vmatrix}a_1 & 0\\ b_1 & 0\end{vmatrix}+\begin{vmatrix}a_1 & 0\\ 0 & b_2\end{vmatrix}+\begin{vmatrix}0 & a_2\\ b_1 & 0\end{vmatrix}+\begin{vmatrix}0 & a_2\\ 0 & b_2\end{vmatrix}\\ &= \begin{vmatrix}a_1 & 0\\ 0 & b_2\end{vmatrix}+\begin{vmatrix}0 & a_2\\ b_1 & 0\end{vmatrix}\\ &= a_1b_2-a_2b_1\end{aligned}$$

一个三阶行列式可以拆分成六个（其余的行列式值等于零）三阶对角行列式：

$$\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix} = \begin{vmatrix}a_1 & 0 & 0\\ 0_1 & b_2 & 0\\ 0 & 0 & c_3\end{vmatrix}+\begin{vmatrix}0 & a_2 & 0\\ 0 & 0 & b_3\\ c_1 & 0 & 0\end{vmatrix}+\begin{vmatrix}0 & 0 & a_3\\ b_1 & 0 & 0\\ 0 & c_2 & 0\end{vmatrix}+\begin{vmatrix}a_1 & 0 & 0\\ 0 & 0 & b_3\\ 0 & c_2 & 0\end{vmatrix}+\begin{vmatrix}0 & a_2 & 0\\ b_1 & 0 & 0\\ 0 & 0 & c_3\end{vmatrix}+\begin{vmatrix}0 & 0 & a_3\\ 0 & b_2 & 0\\ c_1 & 0 & 0\end{vmatrix}$$

$$=a_1b_2c_3+a_2b_3c_1+a_3b_1c_2-a_1b_3c_2-a_2b_1c_3-a_3b_2c_1$$

因此，二阶行列式的有向面积是由两块小长方形有向面积累加的代数和。三阶行列式的有向体积是由六块小长方体有向体积累加的代数和。

决定每一个面积块和体积块的符号或方向的因素是乘积项中元素的排列顺序，这个排列顺序决定了逆序数。通过上面的行列式拆解式，我们看到，每一个乘积项对应一个同样阶数的行列式，乘积项中每一个元素对应者一个向量（在坐标轴上）。

一个行列式的整体几何意义是有向线段（一阶行列式）或有向面积（二阶行列式）或有向体积（三阶行列式及以上）。因此，行列式最基本的几何意义是由各个坐标轴上的有向线段所围起来的所有有向面积或有向体积的累加和。这个累加要注意每个面积或体积的方向或符号，方向相同的要加，方向相反的要减，因而，这个累加的和是代数和。

## 3.6. 克莱姆法则的几何意义

1750 年，瑞士的克莱姆发现了用行列式求解现行方程组的克莱姆（Cramer）法则。这个法则在表述上简洁自然，思想深刻，包含了对多重行列式的计算，是对行列式与线性方程组之间关系的深刻理解。如果我们不能从几何上解释这个法则，就不可能领会向量、行列式和线性方程组之间的真正关系。

### 二阶克莱姆法则的几何解释

二阶线性方程组：

$$\begin{cases}a_1x+b_1y=c_1\\ a_2x+b_2y=c_2\end{cases}$$

其克莱姆法则的解：

$$x=\frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}},\ y=\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_1 & b_2 \end{vmatrix}}$$

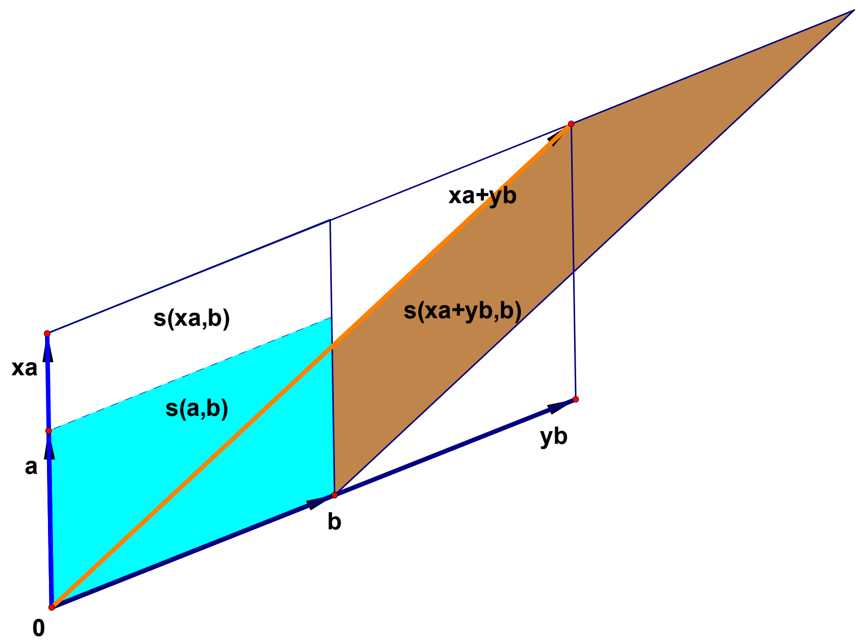
在这里，为了解释其几何意义，同样，从行列式的列或行向量的角度入手。2 阶列向量 $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{x}$ 分别表示为：

$$\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix},\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix},\mathbf{c}=\begin{pmatrix} c_1 \\ c_2 \end{pmatrix},\mathbf{x}=\begin{pmatrix} x \\ y \end{pmatrix}$$

上述二阶线性方程组使用向量的形式表示为 $(\mathbf{a}\ ,\mathbf{b})\mathbf{x}=\mathbf{c}$。下面我们推导出二阶克莱姆法则：对于二阶线性方程组的系数行列式为 $|\mathbf{a},\mathbf{b}\,|$，我们构造出：

$$x|\mathbf{a},\ \mathbf{b}\,|=|x\mathbf{a},\ \mathbf{b}\,|=|x\mathbf{a}+y\mathbf{b},\ \mathbf{b}\,|=|\mathbf{c},\ \mathbf{b}\,|。$$

因此：$x=\dfrac{|\mathbf{c},\ \mathbf{b}\,|}{|\mathbf{a},\ \mathbf{b}\,|}$，据此推理过程，绘出几何图形如下：




在上图中，向量 $\mathbf{a},\ \mathbf{b}$ 的面积是 $s(\mathbf{a},\ \mathbf{b})=|\mathbf{a},\ \mathbf{b}\,|$，这个面积乘以一个数值 $x$（这里 $x$ 大于 1），

则 $s(\mathbf{a},\mathbf{b})$ 伸展为 $s(x\mathbf{a},\mathbf{b})$；此时进行右向切变换，面积不变，表达式则为 $s(x\mathbf{a}+y\mathbf{b},\mathbf{b})$ 即 $s(\mathbf{c},\mathbf{b})$。

由克莱姆法则的表达式 $x=\dfrac{|\mathbf{c},\mathbf{b}|}{|\mathbf{a},\mathbf{b}|}=\dfrac{s(\mathbf{c},\mathbf{b})}{s(\mathbf{a},\mathbf{b})}$，我们可以归纳出：

**$x$ 为由面积 $s(\mathbf{a},\mathbf{b})$ 伸缩和切变到 $s(\mathbf{c},\mathbf{b})$ 的面积之比例，是变化前后的面积之比。**

同理，我们构造出 $y|\mathbf{a},\mathbf{b}|=|\mathbf{a},y\mathbf{b}|=|\mathbf{a},x\mathbf{a}+y\mathbf{b}|=|\mathbf{b},\mathbf{c}|$，得到 $y=\dfrac{|\mathbf{b},\mathbf{c}|}{|\mathbf{a},\mathbf{b}|}$。几何图形类同 $x=\dfrac{|\mathbf{c},\mathbf{b}|}{|\mathbf{a},\mathbf{b}|}$，不再给出了。

## 三阶克莱姆法则的几何解释

类似的，三阶线性方程组如下：

$$\begin{cases} a_1x+b_1y+c_1z=d_1 \\ a_2x+b_2y+c_2z=d_2 \\ a_3x+b_3y+c_3z=d_3 \end{cases}$$

其克莱姆法则的解：

$$x=\frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}},\ y=\frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}},\ z=\frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}$$

相类似的，三阶列向量 $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{x}$ 分别表示为：

$$\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix},\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},\mathbf{c}=\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},\mathbf{d}=\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}\mathbf{x}=\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

三阶线性方程组表示为：$(\mathbf{a}\ ,\mathbf{b},\mathbf{c})\mathbf{x}=\mathbf{d}$，与二阶克莱姆法则的推导类似，下面我们推导出三阶克莱姆法则：

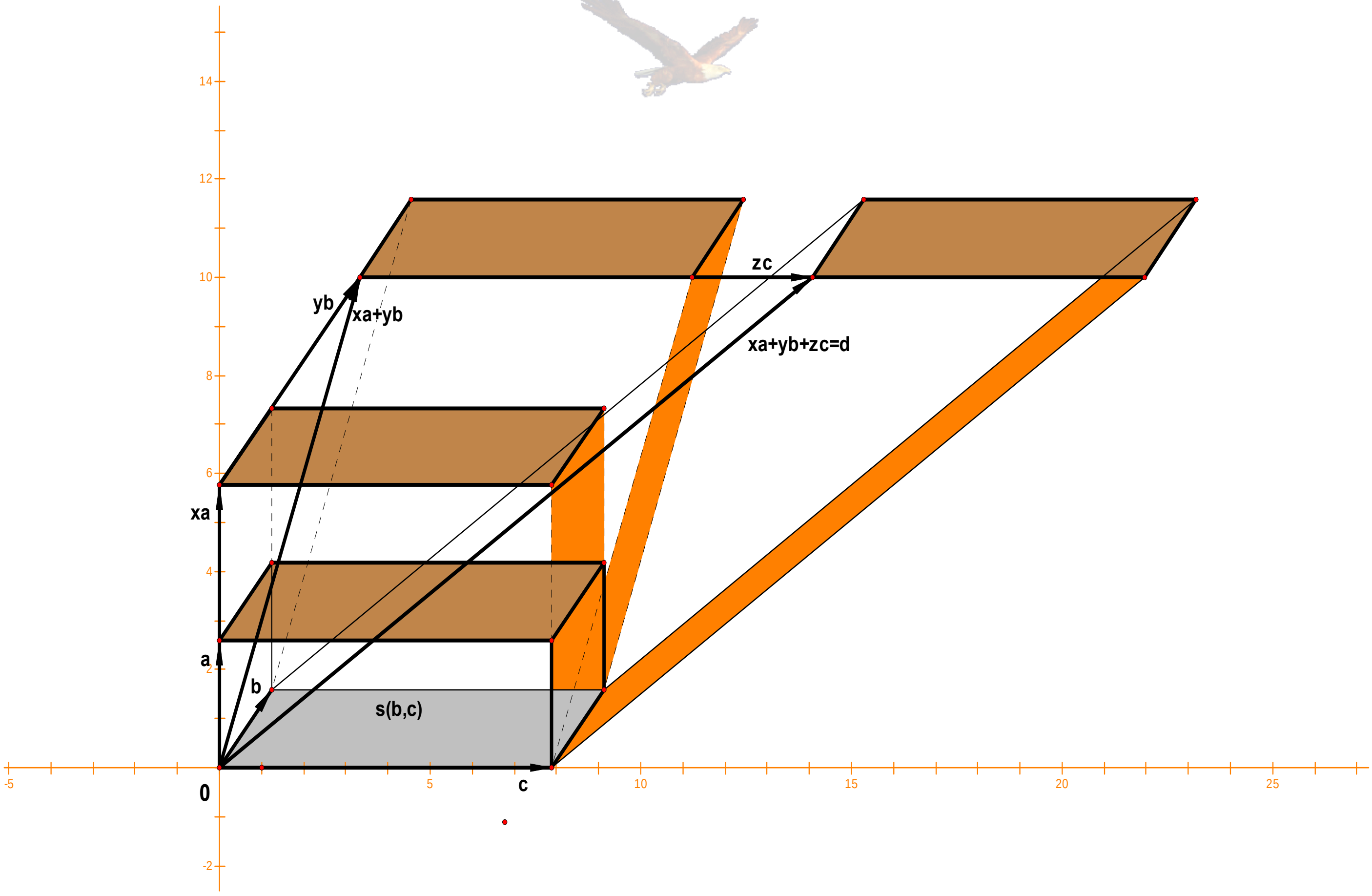
对于三阶线性方程组的系数行列式为$|\mathbf{a},\mathbf{b},\mathbf{c}|$，我们构造出：

$$x|\mathbf{a},\mathbf{b},\mathbf{c}|=|x\mathbf{a},\ \mathbf{b},\mathbf{c}\ |=|x\mathbf{a}+y\mathbf{b},\ \mathbf{b},\mathbf{c}\ |=|x\mathbf{a}+y\mathbf{b}+z\mathbf{c},\ \mathbf{b},\mathbf{c}\ |=|(\mathbf{a}\ ,\mathbf{b},\mathbf{c})\,\mathbf{x},\ \mathbf{b},\mathbf{c}\ |=|\mathbf{d},\mathbf{b},\mathbf{c}|。$$

因此：$x=\dfrac{|\mathbf{d},\mathbf{b},\mathbf{c}|}{|\mathbf{a},\mathbf{b},\mathbf{c}|}$。

类似的，我们可以得到 $y=\dfrac{|\mathbf{a},\mathbf{d},\mathbf{c}|}{|\mathbf{a},\mathbf{b},\mathbf{c}|},z=\dfrac{|\mathbf{a},\mathbf{b},\mathbf{d}|}{|\mathbf{a},\mathbf{b},\mathbf{c}|}$。

据上述推理过程，绘出 $x=\dfrac{|\mathbf{d},\mathbf{b},\mathbf{c}|}{|\mathbf{a},\mathbf{b},\mathbf{c}|}$ 几何图形如下：




由上图看到，由向量$\mathbf{d},\mathbf{b},\mathbf{c}$张成的平行六面体的体积$|\mathbf{d},\mathbf{b},\mathbf{c}|$除以以向量$\mathbf{a},\mathbf{b},\mathbf{c}$ 张成的平行六面

体的体积$|\mathbf{a},\mathbf{b},\mathbf{c}|$就是向量$\mathbf{a}$ 的伸缩量 x 值。因为我们看到，在上图所示的变换中，以$\mathbf{b},\mathbf{c}$两向量张成的平行四边形面积$\mathbf{s}(\mathbf{b},\mathbf{c})$为底的平行六面体的变化过程中，只有向量$\mathbf{a}$ 进行了伸缩变化，其余的平行六面体的变化只是先后沿着向量$\mathbf{b}$ 和$\mathbf{c}$ 的方向进行切向变化，因此向量$\mathbf{a}$ 方向的高度没有变化。

克莱姆法则的意义是可以用方程组的系数和常数项的行列式把方程组的解简洁的表达出来。但在实际工程应用中由于计算量较大，常常采用高斯消元法来解大型的线性方程组。在后面的线性方程组一章，我们将探讨高斯消元法的意义。

**在以上的几何解释中，除了伸缩、旋转、镜像就是切变，向量的大小和方向在变化，但没有对向量进行弯曲，扭曲变化，全部是直线段的变化。所有的变化保持直线性。这就是线性变换的本质含义！**

## 3.7. 行列式的一些应用

结合实际生活中的例子解释行列式的性质，如：一句玩笑话“把人都挤成照片了”，引申到维数的变化。人是三维的物体，体积不为零。挤成二维的照片，体积就变成了零。行列式也是这样：三阶行列式表示平行六面体的有向体积，如果其中有某两列相等，就是说平行六面体的三条相邻的棱中有两条重合，平行六面体退化成平面图形，也就是被“挤成照片”了，体积变成零。类似地，二阶行列式表示平行四边形的有向面积，如果两列相等，“平行四边形”的相邻两边重合，平行四边形退化为一条线段，面积为零。一般地，n 阶行列式可以想象成一个 n 维立体的 n 维体积，如果它有某两列相等，“n 维立体”退化为 n-1 维或者更低维数的图形。“n 维体积”当然就等于零。

线性代数中的行列式可以方便的应用到几何以及其它课程中，简洁地表示出一些结论，以帮助学生记忆和理解这些结论。

过平面两点$(x_1,y_1)$，$(x_2,y_2)$的直线方程为$\begin{vmatrix} x,y,1 \\ x_1,y_1,1 \\ x_2,y_2,1 \end{vmatrix}=0$，

再推广到空间有不在同一条直线上的三点$(x_1,y_1,z_1)$，$(x_2,y_2,z_2)$，$(x_3,y_3,z_3)$的平面方程为

$$\begin{vmatrix} x,y,z,1 \\ x_1,y_1,z_1,1 \\ x_2,y_2,z_2,1 \\ x_3,y_3,z_4,1 \end{vmatrix}=0$$

例。在空间中，有三点 A(a1，a2，a3)，B(b1，b2， b3)，C(c1， c2， c3)，令过点 A，B，C 的平面是π，M(x，y，z)是π上的任意一点，令$\overrightarrow{AB}=(b_1-a_1,\ b_2-a_2,\ b_3-a_3)^T,\ \overrightarrow{AC}=(c_1-a_1,\ c_2-a_2,\ c_3-a_3)^T$，

则由 $\overrightarrow{AM}\cdot(\overrightarrow{AB}\times\overrightarrow{AC})=0$，推出π的方程可用行列式表示为

$$\begin{vmatrix} x-a_1 & y-a_2 & z-a_3 \\ b_1-a_1 & b_2-a_2 & b_3-a_3 \\ c_1-a_1 & c_2-a_2 & c_3-a_3 \end{vmatrix}=0$$

进一步，根据行列式的性质，上式左端的行列式变形如下

$$\begin{vmatrix} x & y & z \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}+\begin{vmatrix} -a_1 & y & z \\ -a_1 & b_2 & b_3 \\ -a_1 & c_2 & c_3 \end{vmatrix}+\begin{vmatrix} x & -a_2 & z \\ b_1 & -a_2 & b_3 \\ c_1 & -a_2 & c_3 \end{vmatrix}+\begin{vmatrix} x & y & -a_3 \\ b_1 & b_2 & -a_3 \\ c_1 & c_2 & -a_3 \end{vmatrix}$$

$$=\begin{vmatrix} x & y & z \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}-a_1\begin{vmatrix} 1 & y & z \\ 1 & b_2 & b_3 \\ 1 & c_2 & c_3 \end{vmatrix}+a_2\begin{vmatrix} 1 & x & z \\ 1 & b_1 & b_3 \\ 1 & c_1 & c_3 \end{vmatrix}-a_3\begin{vmatrix} 1 & x & y \\ 1 & b_1 & b_2 \\ 1 & c_1 & c_2 \end{vmatrix}$$

$$=\begin{vmatrix} 1 & a_1 & a_2 & a_3 \\ 1 & x & y & z \\ 1 & b_1 & b_2 & b_3 \\ 1 & c_1 & c_2 & c_3 \end{vmatrix}=-\begin{vmatrix} 1 & x & y & z \\ 1 & a_1 & a_2 & a_3 \\ 1 & b_1 & b_2 & b_3 \\ 1 & c_1 & c_2 & c_3 \end{vmatrix}=-D$$

故，π的方程还可用行列式表示为 D=0。

在平面上过点 A(a1，a2)，B(b1，b2)的直线方程可用行列式为

$$\begin{vmatrix} x & y & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix}=0$$

在线性代数教材[4]的例题中，推出了过平面上三点 Ai(xi，yi)， (i=1，2，3)且对称轴平行于 y 轴的抛物线方程为：

$$\begin{vmatrix} x^2 & x & 1 & y \\ x_1^2 & x_1 & 1 & y_1 \\ x_2^2 & x_2 & 1 & y_2 \\ x_3^2 & x_3 & 1 & y_3 \end{vmatrix}=0$$

还推出了以点 A(a1，a2，a3)，B(b1，b2，b3)，C(c1， c2， c3)，D(d1，d2，d3)为顶点的四面体体积公式为：

$$V = \pm\frac{1}{6}\begin{vmatrix} 1 & a_1 & a_2 & a_3 \\ 1 & b_1 & b_2 & b_3 \\ 1 & c_1 & c_2 & c_3 \\ 1 & d_1 & d_2 & d_3 \end{vmatrix}$$

等等。可以说，这是行列式作为一个基本工具的重要应用，这充分体现了线性代数与几何的联系。

请关注第四章：

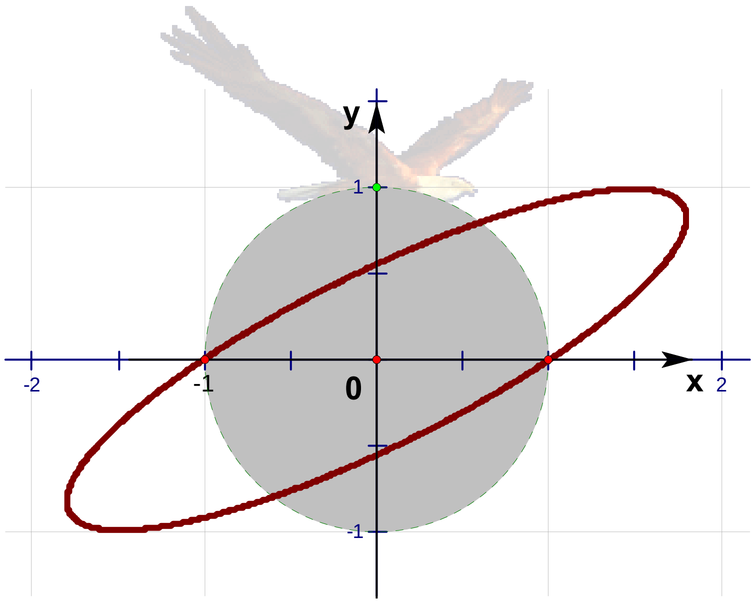
# 第四章　向量组及向量空间的几何意义

待续

---图解线性代数-----

# 线性代数的几何意义之（4）

任广千　胡翠芳 编著




2010.07.01

# 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

# 第三章　向量组及向量空间的几何意义

向量组的关键概念是线性相关性及其秩的概念，在向量组张成的向量空间里，基、维数、坐标及基变换等也是些有点让人头疼的东东。本章就是要从几何图形上弄清这些概念，让抽象的概念回归形象的几何解释。

## 4.1.　向量组的几何意义

向量组是对有限个向量集合的研究，对向量组的性质了解清楚后，就会对矩阵和线性方程组的性质认识更加深入。因为矩阵实际上就是一个有序向量组。线性方程组实际上就是向量组的线性表示。因此，在开展矩阵及方程组的研究之前，有必要先研究清楚向量组的问题。

向量组里的向量们有哪些特性呢？有什么共性？有没有不变量？

有，这个不变的特性就是一个叫“秩”的东东。下面我们来慢慢探究一下。

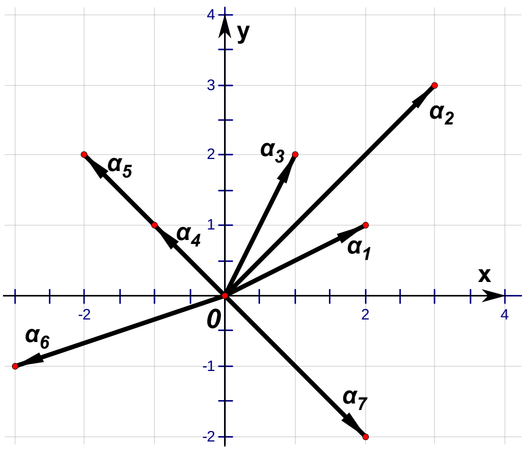
### 4.1.1. 线性表示、组合及相关性的意义

**向量的线性表示和组合的几何意义**

先看看下面的图中平面向量集合。这个集合有七个二维向量。用坐标表示出来就是：

$\boldsymbol{\alpha}_1=(2,1)$，$\boldsymbol{\alpha}_2=(3,3)$，$\boldsymbol{\alpha}_3=(1,2)$，$\boldsymbol{\alpha}_4=(-1,1)$，$\boldsymbol{\alpha}_5=(-2,2)$，$\boldsymbol{\alpha}_6=(-3,-1)$，$\boldsymbol{\alpha}_7=(2,-2)$。

仔细观察后发现：




向量 $\boldsymbol{\alpha}_2$ 可以两个向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_3$ 向加得到，即 $\boldsymbol{\alpha}_2=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3$；

向量 $\boldsymbol{\alpha}_5$ 可以由 $\boldsymbol{\alpha}_4$ 乘以 2 得到，也可以由 $\boldsymbol{\alpha}_7$ 乘以-1 得到，即 $\boldsymbol{\alpha}_5=2\boldsymbol{\alpha}_4$ 或 $\boldsymbol{\alpha}_5=-\boldsymbol{\alpha}_7$；

甚至，向量$\boldsymbol{\alpha}_6$也可以由$\boldsymbol{\alpha}_2$的数乘和$\boldsymbol{\alpha}_7$数乘之和得到，即$\boldsymbol{\alpha}_6=-\frac{2}{3}\boldsymbol{\alpha}_2-\frac{1}{2}\boldsymbol{\alpha}_7$；

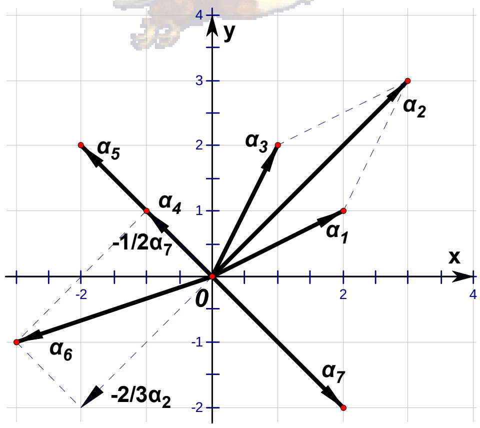
…

上面的例子是说，一个向量可以由另外一个或几个向量（向量组）用数乘之和的形式表示出来，一般表达式就是：$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\ldots+x_s\boldsymbol{\alpha}_s$， $x_1,x_2,\ldots x_s$是常数；这里称之为向量$\boldsymbol{\beta}$可以由向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots\boldsymbol{\alpha}_s\}$线性表示。或者讲，向量$\boldsymbol{\beta}$是向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots\boldsymbol{\alpha}_s$的线性组合。

我们研究一下线性表示的表达式$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\ldots+x_s\boldsymbol{\alpha}_s$，明显的：向量$\boldsymbol{\alpha}_1$数乘$x_1$，向量$\boldsymbol{\alpha}_2$数乘$x_2$，…，然后把数乘后的向量相加起来就得到了一个新的向量$\boldsymbol{\beta}$。反过来讲，一个向量$\boldsymbol{\beta}$被分解为几个向量的倍数。所以我们得到一个结论：

**线性组合或表示式实质上是向量的数乘和加法的综合。**

在第二章的向量介绍中，我们知道，数乘的几何解释就是在原向量的直线上向量长度的伸长或缩短；两向量相加的几何解释就是进行依照平行四边形法则对向量合并；因此向量的线性组合的几何意义就是对向量组内的向量长度进行缩放后依照平行四边形法则进行合并加；线性表示的几何意义就是可以把一个向量依照平行四边形法则分解（或投影）为向量组上的和。如下图所示。




**向量的线性相关和线性无关的几何意义**

如果一个向量可以由一个向量组线性表示，我们就称这个向量和向量组线性相关。另外的说法就是，一个向量组里，只要有一个向量可以由其它向量线性表示，我们就称这个向量组线性相关。反之，如果向量组里的任意一个向量都不能由其它向量线性表示，我们就称向量组线性无关。

上图中线性相关的向量组例举如下：

{ $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$}线性相关，因为$\boldsymbol{\alpha}_2=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3$；

$\{\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5\}$，$\{\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_7\}$，$\{\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7\}$都线性相关，因为$\boldsymbol{\alpha}_5=2\boldsymbol{\alpha}_4=-2\boldsymbol{\alpha}_7$；

$\{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_6,\boldsymbol{\alpha}_7\}$线性相关，因为$\boldsymbol{\alpha}_6=-\frac{2}{3}\boldsymbol{\alpha}_2-\frac{1}{2}\boldsymbol{\alpha}_7$；

…

仔细研究，我们就会发现一些规律，就是在二维平面上：

- 如果两个向量线性相关，那么这两个向量必然在一条直线上（这也是“线性相关”术语的由来），两个向量的方向或者相同或者相反；反之也成立，例如向量$\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7$在一条直线上，因此两两相关；

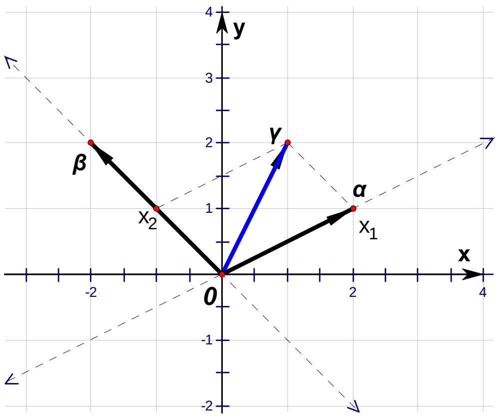
解释：因为在一条直线上的所有向量中的两个向量都具有$x$倍数的关系：$\boldsymbol{\beta}=x\boldsymbol{\alpha}$，所有向量都线性相关。

- 不在一条直线上的任意两个向量一定线性无关；如向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4(or\,\boldsymbol{\alpha}_5\,or\,\boldsymbol{\alpha}_7),\boldsymbol{\alpha}_6$两两线性无关。

解释：因为一个向量必然不能被另外一个向量线性表示。如果一个向量能够被另外一个向量线性表示的话即$\boldsymbol{\beta}=x\boldsymbol{\alpha}$，那么这两个向量就是$x$倍数的关系，就会在一条直线上。这与命题的前提矛盾。

- 在二维平面空间上，任意三个向量必然相关，如{ $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$}，{$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_6,\boldsymbol{\alpha}_7$}等。当然，三个以上的向量也必然线性相关。

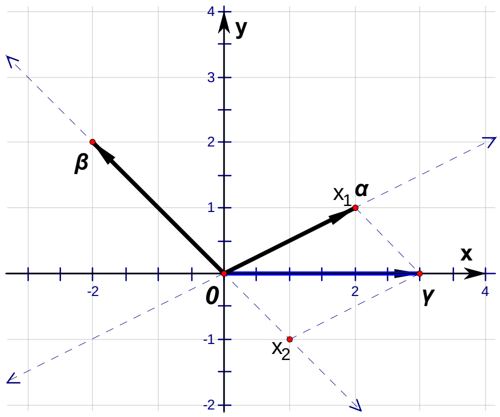
解释：我们看看任意一个三向量{$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$}，如果其中任意两个如$\boldsymbol{\alpha},\boldsymbol{\beta}$线性相关，不用说这个三向量的向量组也线性相关；下一步我们设$\boldsymbol{\alpha},\boldsymbol{\beta}$线性无关，看看为什么这个三向量的向量组就现行相关了呢？如下图：

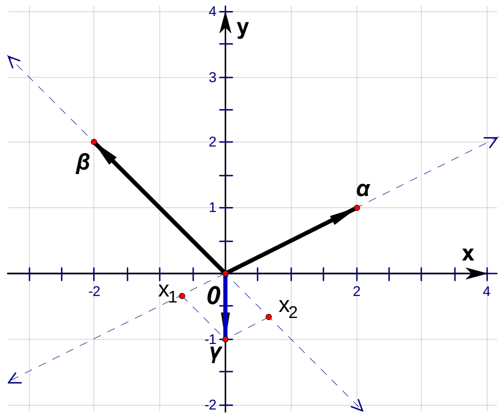



上图中，我们随意画出不在一条直线上的两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$，分别作出过此两向量的直线（所示虚线）；然后过第三个向量$\gamma$的头部点分别作$\boldsymbol{\alpha},\boldsymbol{\beta}$的平行线，分别交$\boldsymbol{\alpha},\boldsymbol{\beta}$直线的交点为$x_1,x_2$。至此，

我们构造出了向量 $\boldsymbol{\gamma}$ 的分解式 $\boldsymbol{\gamma}=x_1\boldsymbol{\alpha}+x_2\boldsymbol{\beta}$ 。

改变向量 $\boldsymbol{\gamma}$ 的位置如下图，我们仍然同样得到同样的表达式，只是 $x_1$ , $x_2$ 的值不同罢了。



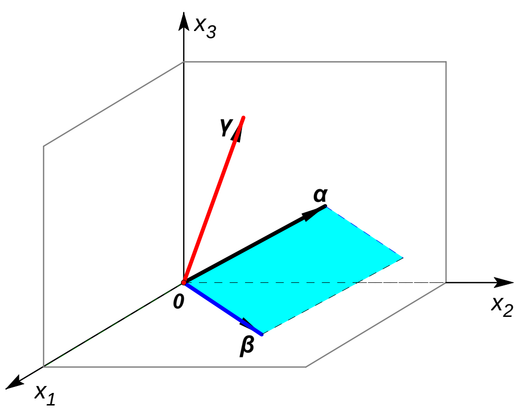

这个分解式就是线性表示式。因此{ $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ }线性相关（再强调一下，此命题在二维向量空间中成立）。

对于通常的三维空间的向量，线性相关和线性无关也有直观的几何意义。

先考察三维空间中两个向量的线性关系。由定义，它们线性相关是说其中一个是另一个的线性组合，即只与另一个向量相差一个常数因子，这就是说，这两个向量落在同一条直线上，方向相同或者方向相反。反之，如果两个向量落在同一条直线上，那么它们就线性相关。两个向量线性无关则说这两个向量不平行，不能平移地搬到同一条直线上去。

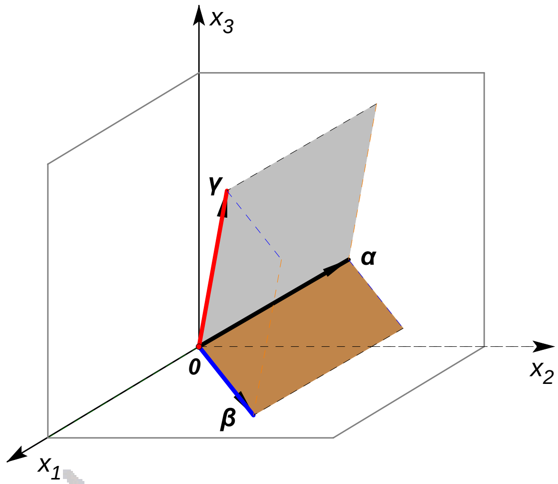
现在来考察三维空间中三个向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 的线性关系。

设向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性无关。那么就有 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 两两线性无关，任意假设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性无关，则说明 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 不在同一条直线上，向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 所在的两根相交直线构成了一个平面（或者讲 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 张成的平面），这个平面上的所有的向量是 $x_1\boldsymbol{\alpha}+x_2\boldsymbol{\beta}$ （ $x_1,x_2$ 为任意实数）。



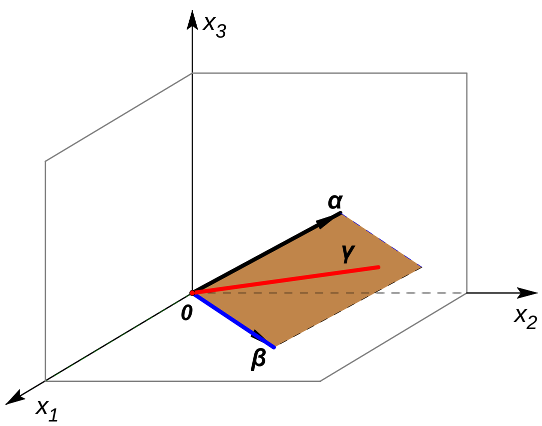
既然 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性无关，那么就没有 $x_1\boldsymbol{\alpha}+x_2\boldsymbol{\beta}=\boldsymbol{\gamma}$ 的可能，也就是说向量 $\boldsymbol{\gamma}$ 必然不在这个平面上，而是在这个平面之外。

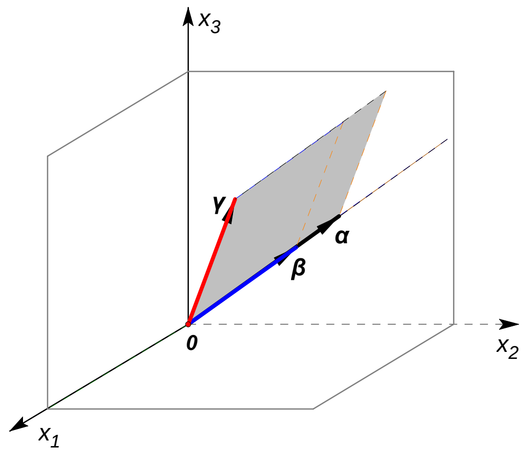
实际上，这三个向量中任意一个向量，都不在其余两个向量所张成的平面内，如果用这三个三维向量构成一个三阶行列式，那么必然张成一个平行六面体，同时行列式的值不等于零。




如果设向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性相关。这里有几种情况。

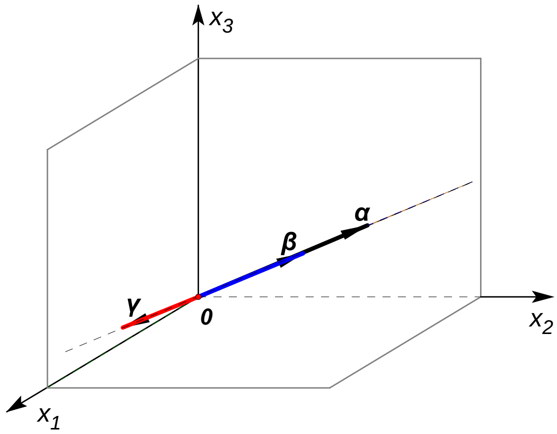
一种情况是三个向量在一个平面上（共面）。正如前面讨论三个向量线性无关的情况相反。任意的，如果 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 不相关，那么 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 会张成一个平面 $\{x_1\boldsymbol{\alpha}+x_2\boldsymbol{\beta}\}$；又因为 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性相关，则必会有 $\boldsymbol{\gamma}$ 落在这个平面上即 $x_1\boldsymbol{\alpha}+x_2\boldsymbol{\beta}=\boldsymbol{\gamma}$ （$x_1,x_2$ 因 $\boldsymbol{\gamma}$ 的不同而不同），如下图左。如果 $\boldsymbol{\gamma}$ 不落在这个平面上，会得到 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性无关的矛盾。




任意的，如果 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性相关，但第三个向量 $\boldsymbol{\gamma}$ 与 $\boldsymbol{\alpha}$（或 $\boldsymbol{\beta}$）线性无关，则也会有一个平面出来，仍然会得到三个向量共面的情况，如上图右。

一种情况是三个向量在一条直线（共线）。继续前面的讨论。也就是说，如果$\boldsymbol{\alpha},\boldsymbol{\beta}$线性相关，第三个向量$\boldsymbol{\gamma}$也与$\boldsymbol{\alpha}$（或$\boldsymbol{\beta}$）线性相关。那么三个向量会无意外的落在同一条直线上，如下图。




### 4.1.2. 向量组等价及秩的几何意义

#### 向量组的等价的几何解释

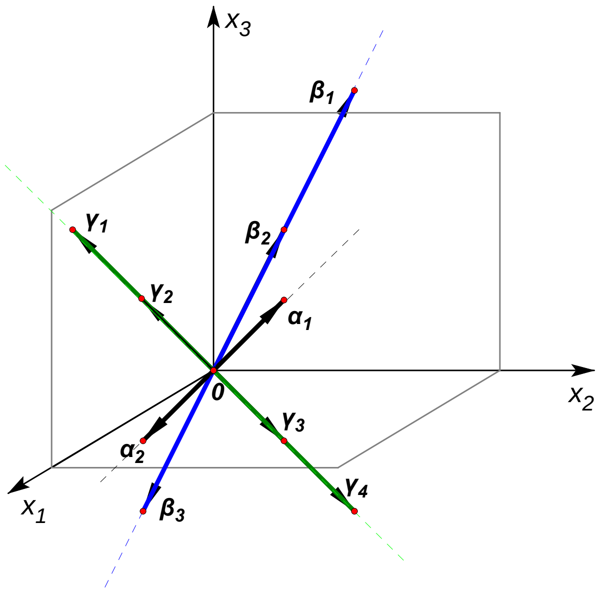
两个向量组 A 和 B 的等价就是这两个向量组能够互相被线性表示。详细地说，向量组 A 中的每一个向量都可以被向量组 B 线性表示；同样，向量组 B 中的每一个向量也可以被向量组 A 线性表示。或者说，如果把一个向量组中的任意一个向量拿出来放到另外一个向量组中，那么另外这个扩大的向量组就会线性相关，而且不论原向量组是否线性相关。采用向量空间（后面会探讨向量空间）的概念，就会清晰的明了向量组等价的几何意义：

向量组 A 中每一个向量都在向量组 B 张成的向量空间中；同样，向量组 B 中的每一个向量也在向量组 A 张成的向量空间中。或者说，两个向量组如果张成的向量空间相同或重合，就讲这两个向量组等价。

先看看 3 维的空间中的向量组的等价关系

**直线上的等价向量组：**

下图三维空间中，共有三条分离的不共面直线，每条直线上分别由两个、三个和四个向量。两向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$在一条直线上；三向量$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$在另外一条直线上；四向量$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4$在第三条直线上。




由此，我们可以验证以下的命题：

- $\{\boldsymbol{\alpha}_1\},\{\boldsymbol{\alpha}_2\},\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 是等价向量组；

- $\{\boldsymbol{\beta}_1\},\{\boldsymbol{\beta}_2\},\{\boldsymbol{\beta}_3\},$<br>$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\},\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_3\}\{\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\},$ 是等价向量组；<br>$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$

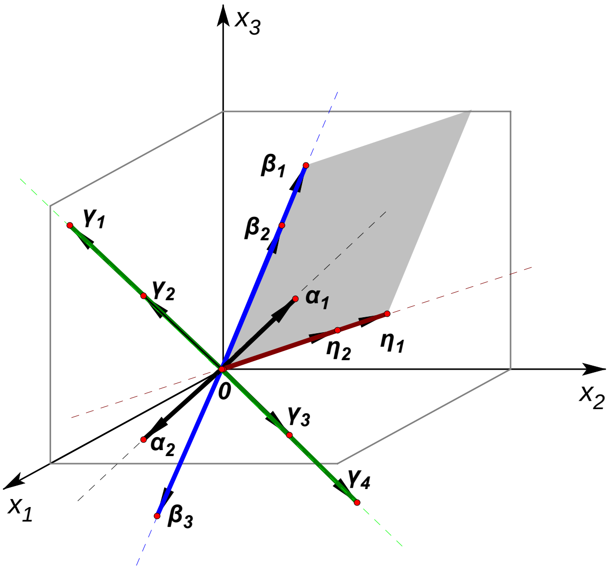
- $\{\boldsymbol{\gamma}_1\},\{\boldsymbol{\gamma}_2\},\{\boldsymbol{\gamma}_3\},\{\boldsymbol{\gamma}_4\},$<br>$\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2\},\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_3\},\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_4\},\{\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3\},\{\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_4\},\{\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\}$ 是等价向量组；<br>$\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3\},\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_4\},\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\},\{\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\}$<br>$\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\}$

其实上述的命题不用验证也可以知道，因为我们罗列的等价向量组是在一条直线上，而在一条直线上的向量是可以互相线性表示的。

另外，更多的向量组，如果他们所属的直线集合是相等的，那么这些向量组也是等价的，比如三个向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\gamma}_2\},\{\boldsymbol{\alpha}_2,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\},\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4\}$ 是等价的，因为每一组的向量都在 $\boldsymbol{\alpha}$ , $\boldsymbol{\gamma}$ 两条直线上。

**平面上的等价向量组：**

类似的，三维空间中，我们可以看看直线上和平面上的向量组之间的等价关系。如下图中，向量 $\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\eta}_i$ 分别所在的三条直线共面（阴影平行四边形），因此向量 $\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\eta}_i$ 中的任何一类可以被其它两类线性表示例如有关系 $\boldsymbol{\alpha}_i = x_1\boldsymbol{\beta}_i + x_2\boldsymbol{\eta}_i$ 。




那么，除了前面介绍过的直线上的等价向量组外，再考虑平面的话就有一下的等价关系：

- $\{\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i\},\{\boldsymbol{\alpha}_i,\boldsymbol{\eta}_i\},\{\boldsymbol{\beta}_i,\boldsymbol{\eta}_i\},\{\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\eta}_i\}$ 是等价向量组；

比如$\{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2\},\{\boldsymbol{\alpha}_2,\boldsymbol{\eta}_1\},\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\eta}_2\},\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\eta}_1,\boldsymbol{\eta}_2\}$是等价向量组。

如果一个平面在加一个平面外的一条直线$\boldsymbol{\gamma}$，有以下的等价关系：

- $\{\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\gamma_i\},\{\boldsymbol{\alpha}_i,\boldsymbol{\eta}_i,\gamma_i\},\{\boldsymbol{\beta}_i,\boldsymbol{\eta}_i,\gamma_i\},\{\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\eta}_i,\gamma_i\}$ 是等价向量组；

比如$\{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2,\gamma_1\},\{\boldsymbol{\alpha}_2,\boldsymbol{\eta}_1,\gamma_2\},\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\eta}_2,\gamma_1,\gamma_3\},\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\gamma_1,\gamma_2,\gamma_3,\gamma_4\}$是等价向量组。

## 向量组的秩及极大无关向量组

向量空间中的向量无穷多，因此，可以有无数个向量组等价。而且等价的向量组中的向量个数也不尽相同。

从前面的罗列中，还可以看出最短的等价向量组是只有一个向量元素的向量组；长的等价向量组的元素可以无穷多。这里，最短的向量组实际上就是极大无关向量组，最大无关向量组的元素的个数就是等价向量组的秩。如果等价向量组最小只有一个向量，则等价向量组的秩等于 1。

我们可以这样理解极大无关向量组：从原来的较长的向量组中挑出一部分向量组成了一个新的向量组，这个新的向量组在某种意义下可以代表原来的向量组（因为两者等价，可以互相表出）；同时这个新的向量组中很纯净，没有躲在别人后面滥竽充数的向量，多余的向量被剔出了，向量之互相独立，个顶个，既不代表谁也不被代表（任一个向量都不能被其它向量线性表示）。这些个顶个的向量个数就是这些互相等价的向量组的秩。

从几何意义上讲，在一个向量组里，如果有多个向量在一条直线上，哪么这些向量只要一个向量就可以了，其他的同直线的向量可以被代表了。这个向量代表可以是任意一个非零向量；进而，

如果向量组里还有多个向量构成且存在于一个平面上，那么只要有两个非零非共线的向量就可以代表其他的共面向量了；继续，如果向量组里还有多个向量构成且存在于一个立体空间里，那么只要有三个非零非共线非共面的向量就可以代表其它的同立体向量了…。

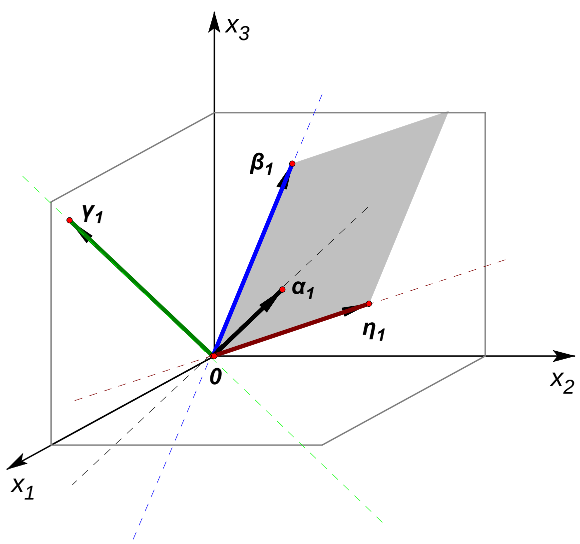
所以，一个 n 维向量组就可以通过几何意义上筛选得到一个极大无关向量组，筛选的过程可以这样：

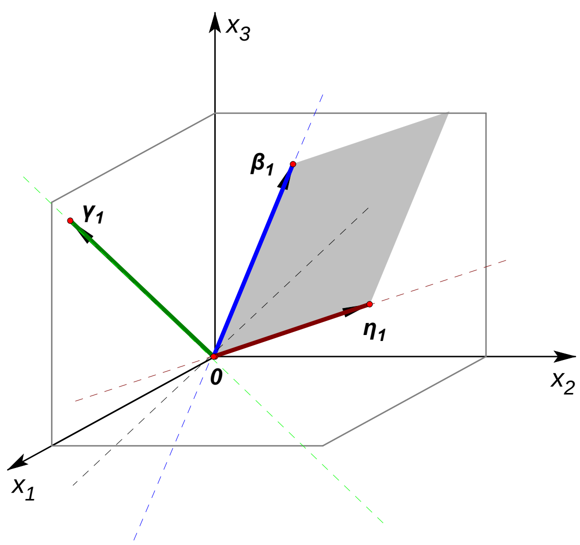
1、 首先把共线的向量全部找出来，然后一个直线留一个向量代表，其余的向量删除；
2、 把上述精简后的向量组里所有共面的向量全部找出来，然后一个平面留两个向量，其余的向量删除；
3、 把上述精简后的向量组所有共立方体的向量全部找出来，然后一个立方体六三个向量，其余的向量删除；
4、 ……；
5、 把上述精简后的向量组所有共超立方体（n-1 维）的向量全部找出来，然后一个超立方体六 n-1 个向量，其余的向量删除；
6、 最后必然留下的向量数小于等于 n，筛选结束，剩下的向量则为极大无关向量组。

注意：

最后精简后的向量的个数就是原 n 维向量组的秩。在整个精简过程组，每一步留下的向量组虽然个数逐步减少，但每一步向量组的秩却一直没有变。秩是一个不变量。

我们把前面的例子中的向量组进行筛选操作，首先一根直线留一个向量（可任意的），如左图，得到了一个向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1,\boldsymbol{\eta}_1,\boldsymbol{\gamma}_1\}$；然后把共面的向量留两个向量（可任意的）得到了包含三个向量的向量组$\{\boldsymbol{\beta}_1,\boldsymbol{\eta}_1,\boldsymbol{\gamma}_1\}$，如右图，筛选过程结束。




使用向量空间的概念，我们可以有一个更全面的关于线性相关的几何意义的结论：

- **一个向量空间中，一个向量组线性相关的话，那么这个向量组中全部向量会属于一个子向**

量空间中，且子空间的维数要小于向量组元素的个数；

- 一个向量空间中，一个向量组线性无关的话，那么这个向量组中全部向量会属于一个子向量空间中，且子空间的维数要大于或等于向量组元素的个数；

（上面两句话意思就是说，一个向量组应该可以张成一个和向量组元素个数相同的子空间，一个向量张成一维的子空间，两个向量应张成二维的字空间…；如果一个向量组 n 个元素，张成一个小于 n 的子空间，那么这个向量组就线性相关；如果总是张成一个 n 维的子空间，那么这个向量组就线性无关）。

- 线性相关或无关的向量组的秩就是可以张成的最大子空间的维数；
- 两个向量组等价，就是两个向量组张成的向量子空间相同或重合；

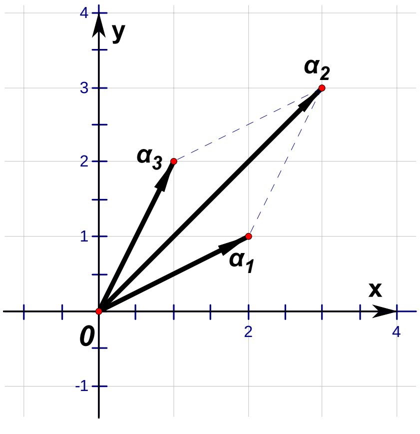
### 4.1.3. 向量组例题的图解

下面介绍两个容易搞错的命题，以加深印象：

***错误命题 1：若向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$中向量两两线性无关，则 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 线性无关。***

图证如下：

向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$的向量定义如下：$\boldsymbol{\alpha}_1=(2,1)$，$\boldsymbol{\alpha}_2=(3,3)$，$\boldsymbol{\alpha}_3=(1,2)$，中显然$\boldsymbol{\alpha}_1$和$\boldsymbol{\alpha}_2$，$\boldsymbol{\alpha}_1$和$\boldsymbol{\alpha}_3$，$\boldsymbol{\alpha}_2$和$\boldsymbol{\alpha}_3$线性无关（不在一条直线上），但$\boldsymbol{\alpha}_3=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，所以向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$线性相关。

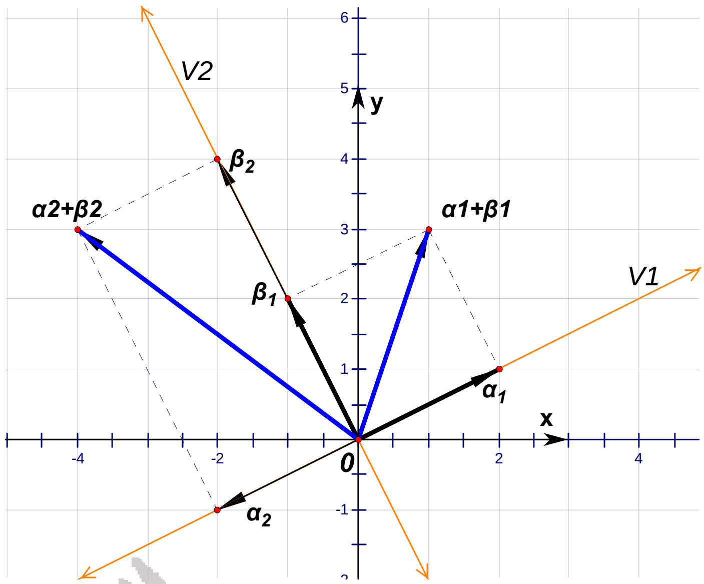



这个例题说明，向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$都是二维向量，属于二维向量空间，而向量组的元素个数是 3，超过了向量的维数，因而线性相关。一般的结论是，n+1 个以上的 n 维向量线性相关。

***错误命题 2：若$\boldsymbol{\alpha}_1$和$\boldsymbol{\alpha}_2$线性相关，$\boldsymbol{\beta}_1$和$\boldsymbol{\beta}_2$也线性相关，那么，$\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1$和 $\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2$也线性相关。***

图证如下：

向量定义如下：$\boldsymbol{\alpha}_1=(2,1)$，$\boldsymbol{\alpha}_2=(-2,-1)$，$\boldsymbol{\beta}_1=(-1,2)$，$\boldsymbol{\beta}_2=(-2,4)$。则有 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1=(1,3)$，$\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2=(-4,3)$；$\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1$ 和 $\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2$ 不在一条直线上，因而 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1$ 和 $\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2$ 线性无关。



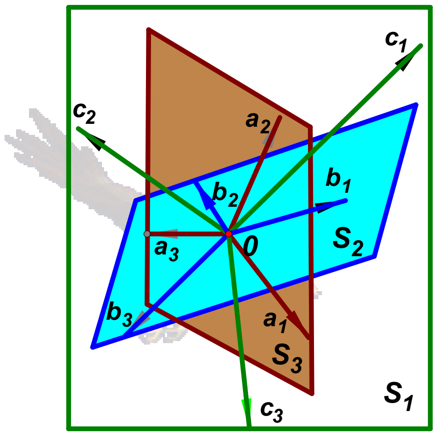

这个例题说明，$\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 在空间 $V_1$（直线）上，$\boldsymbol{\beta}_1$ 和 $\boldsymbol{\beta}_2$ 在空间 $V_2$（直线）上，两个不同空间上的向量（零向量除外）相加，必然会进入第三个空间（直线）、第四个空间（直线）…，以致布满整个二维空间（平面），显然 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1$ 和 $\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2$ 绝大部分线性无关。

## 4.2. 向量空间的几何意义

向量种类繁多，五花八门。形形色色的向量方向、长短各异，应该给他们分类，划分成向量集合。由于向量的概念具有几何的特质，因此向量的集合通常叫做向量空间（空间也是几何的概念）。这个向量空间里的规矩很多，有人给出八条铁律，还有的是十条。其实只有两项基本原则：一是任意两个向量叠罗汉（相加）不能出空间；另一个是任意一个向量伸头缩脑（数乘）也不能超出空间。

我们常常发现，在向量空间这所大房子里又有好多居室，每个居室里的向量们也严格坚守着自己居室的同样的两项基本原则：叠罗汉和伸缩头脑不能出室（呵，有些象我们的小学生，端坐笔直象一个个向量，下课了活动活动还不能出教室），这些大大小小的居室就是子空间。

需要注意的是，这些居室有个特点，就是共有一个原点，或者讲都要包括零向量，所以一个大概的空间和子空间的关系图形可以这样描述：




上图使用平面方框表述空间及其子空间之间的关系。向量空间 $S_3$，包含向量 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{b}_2$；向量空间 $S_2$ 中，向量 $\mathbf{a}_3,\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3$ 都属于 $S_2$ 中；同样，向量空间 $S_3$ 中，$S_1$ 和 $S_2$ 都是 $S_3$ 的子空间，向量 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3$ 都属于 $S_3$ 空间。

在数学教科书中，向量空间的标准定义一般是这样：

设 $V$ 是非空的 n 维向量的集合（n=1，2，3...），如果 $V$ 中的向量对加法和数乘两种运算封闭，也即：

- 若 $\mathbf{a},\mathbf{b}\in V$，则 $\mathbf{a}+\mathbf{b}\in V$；
- $\mathbf{a}\in V$，则 $k\mathbf{a}\in V$，$k$ 为任意实数。

则称 $V$ 为向量空间。

向量空间主要有两种：一种是由 $V$ 中的一个向量组张成的空间（比如由特征向量张成的子空间等）。另外一种齐次线性方程组的解集组成的解空间。实际上，线性方程组的解空间也是解向量所张成。我们后面将会看到，这两种空间里都包含有无穷多的向量。下节我们首先看看由向量所张成

的空间的意义。

### 4.2.1. 向量张成的空间

实际上，向量空间的概念就是对向量们的一个分类。那么一个向量空间如何用数学式子表达呢？换句话就是说，一个空间里面的所有向量(无穷多)如何用有限的数学算式表达呢？

前面已说过，一个向量空间满足两个基本原则：对加法和数乘的运算封闭。把加法和数乘综合到一块，就是线性组合式 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2...+x_n\boldsymbol{\alpha}_n$。所以，我们可以使用一个向量组的线性组合式来表达一个空间里的全部的无穷向量。这个向量组常常是最大向量无关组，也可以是向量的相关组。

一个向量组可以线性表示出一个空间里的所有向量，反过来讲，空间里的所有向量都可以分解为这个向量组的线性表示。那么这个空间我们就叫向量组张成的空间。

下面我们看看它的数学定义式。

设一个向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$，这个向量组的所有的线性组合生成一个向量集合：

$$\{x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2...+x_n\boldsymbol{\alpha}_n \mid x_1,x_2...x_n\in R\}$$

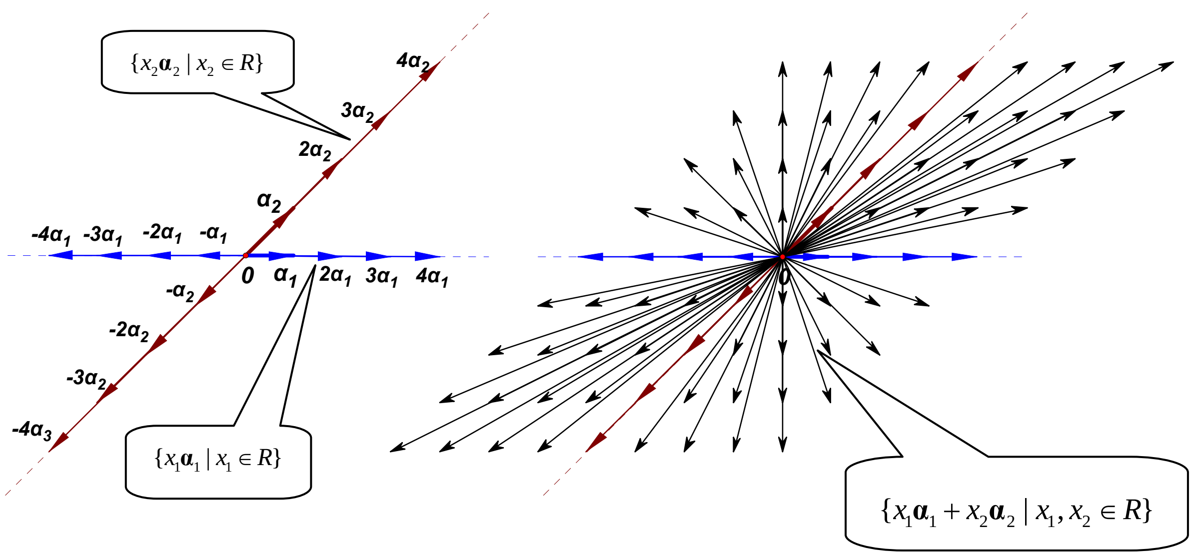
此集合常称为 $Span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$，称为由向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$ 张成的向量空间。

请注意：这个向量空间的数学定义和前面的加法和数乘的定义是等价的。

下图中给出了由向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 张成的向量空间平面 $S$ 的例子：

$$S=Span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}=\{x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2 \mid x_1,x_2\in R\}$$

图中显示，由两个不相关的向量使用平行四边形法则可以生成平面上所有的向量。



我们在讲行列式的几何意义时说，行列式的 m 维超平行多面体像是一个枝繁叶茂的大树所构成的一个物理空间，主枝干就是行列式的 m 个行向量或列向量。虽然向量数量可以是无穷多，但这个物理空间是有限的，空间的体积就是行列式的值。

由向量所张成的线性空间是无穷大的，空间里的向量也是无穷多的。 因为在向量空间的数学定义式$\{x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2...+x_n\boldsymbol{\alpha}_n \mid x_1,x_2...x_n \in R\}$中，因为系数的可以无穷大，因此可以张成无穷大的空间。

## 4.2.2. 子空间的几何意义

子空间的一般定义是这样的：

如果$V$和$H$都是向量空间，而且$H \subset V$，则称$H$是$V$的子空间。

具体说来，由向量空间中的一些向量张成的子空间，其定义如下：

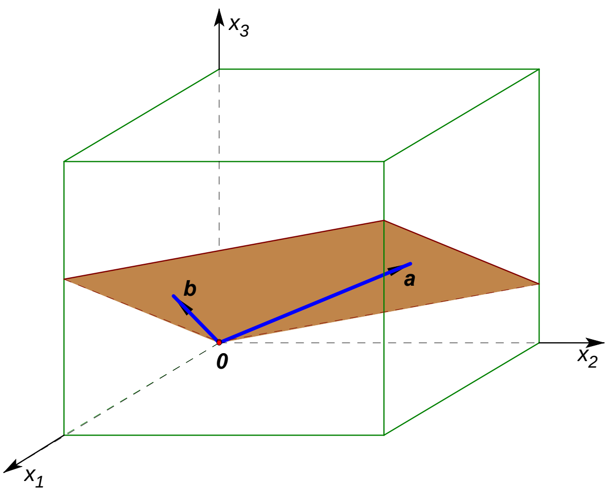
设$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_m\}$是 n 维向量空间$V$的一个向量组，$m \le n$；这个向量组的所有的线性组合生成一个向量空间：

$$Span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_m\}=\{x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2...+x_m\boldsymbol{\alpha}_m \mid x_1,x_2...x_m \in R\}$$

向量空间$Span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_m\}$称为由向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_m$张成的子空间。

这里要提醒一下，$\mathbf{0}$向量是唯一的，既属于$V$空间也属于$H$空间。任意一个子空间$H$都要包含$\mathbf{0}$向量，否则就不能满足加法和数乘的封闭运算。

下面来一个三维空间中由两个三维向量$\mathbf{a}=(a_1,a_2,a_3)$和$\mathbf{b}=(b_1,b_2,b_3)$张成的一个平面二维空间的例子。




这里注意一个小细节：三维的向量张成了一个二维的空间。平面是三维空间的子空间。

**n 维实线性空间 $\mathbf{R}^n$ 的子空间：**

$\mathbf{R}^n$ 表示所有 n 维实向量所构成的集合。每个向量中的元素是实数，元素个数是 n 个。如 $\mathbf{R}^2$ 表示平面实向量集合，$\mathbf{R}^3$ 表示三维空间实向量集合。

三维向量空间 $\mathbf{R}^3$ 的所有子空间包括：

三维子空间：本身 $\mathbf{R}^3 = Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$（$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关），作为自身的子空间表现为一个立体空间，同自身一样，也包含原点；

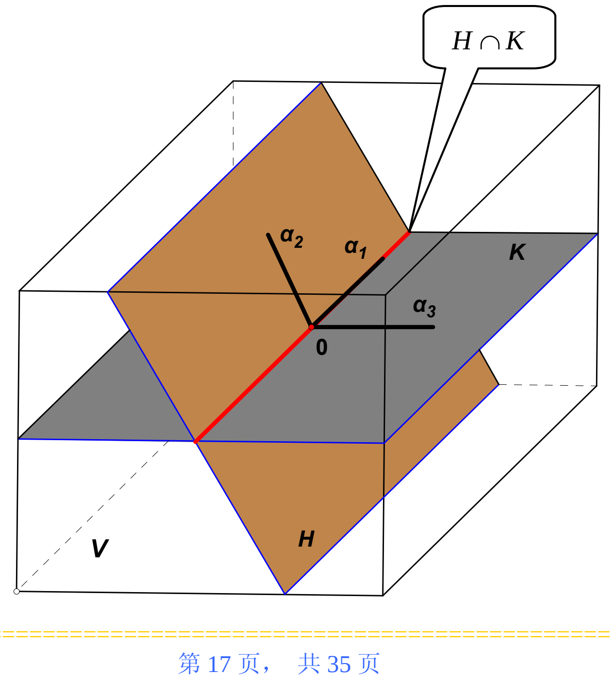
二维子空间：如 $Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ （$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 线性无关），表现为通过原点的任意一个平面（注意：二维空间 $R^2$ 是不是 $\mathbf{R}^3$ 的子空间）；

一维子空间：如 $Span\{\boldsymbol{\alpha}_1\}$ $(\boldsymbol{\alpha}_1 \neq 0)$，表现为通过原点的任意一条直线；

零维子空间：只包含原点 **0** 向量，只有零空间；

下面的图形给出了 $\mathbf{R}^3$ 的所有子空间的图形。

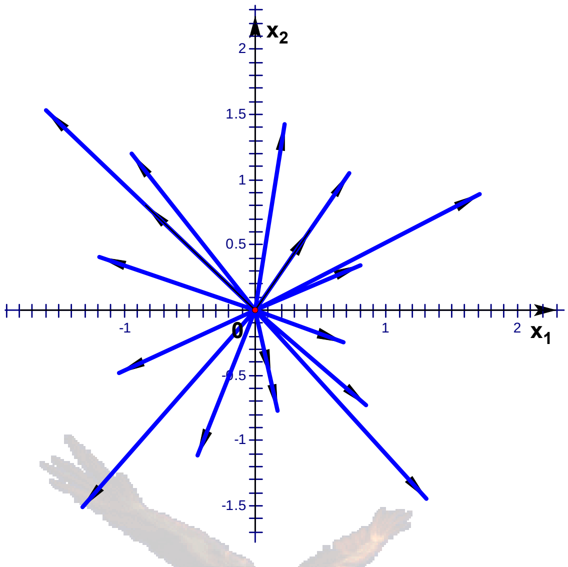
图中，$V$ 三维线性空间即 $\mathbf{R}^3$，它可以由 $Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$（$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关）表示；$H = Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ （$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 线性无关）表示一个二维子空间；$K = Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3\}$（$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3$ 线性无关），表示另一个二维子空间的例子；$H$ 和 $K$ 的公共集合交集 $H \cap K = Span\{\boldsymbol{\alpha}_1\}$ $(\boldsymbol{\alpha}_1 \neq 0)$，是一维向量子空间的例子；上述所有的子空间皆包含零向量 **0**=（0，0，0）；当然零向量自身可以组成一个零空间。



**子空间一定要过原点的几何意义：**

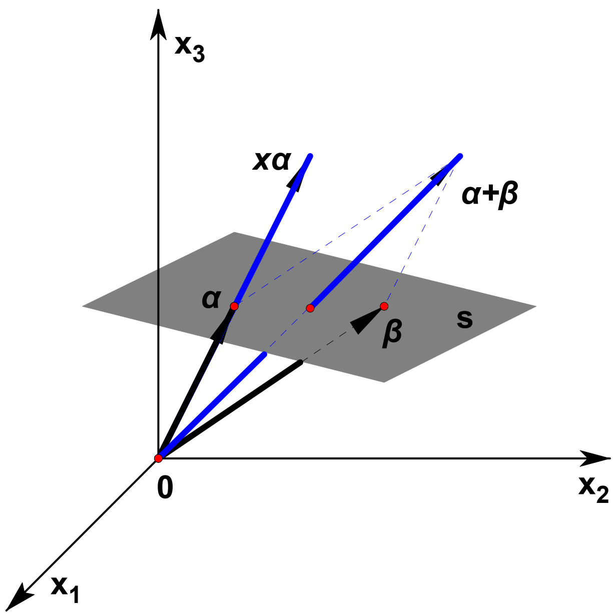
前面在介绍子空间的概念时，总是在强调过原点，或者所有的子空间一定要包含零空间在内。为什么？这是硬性规定吗？

实际上，我们现在讨论的向量，不能称之为自由向量，因为所有的向量的尾巴都被拉到了原点上，或者说，所有响亮空间里的向量都是从原点出发的，大家都有一个共同的零空间，这就是为什么所有的子空间一定要包含零空间的原因了。




为什么要把向量的尾部都拉到原点呢？在前面向量的基础几何意义一章讲过，那就是为了研究向量的方便，因为这样就可以把向量和空间中的点一一对应起来。空间中一旦建立起了坐标系，点有坐标值，那么我们就用点的坐标表示与点对应的向量，这样向量就有了解析式，就有了向量的坐标表达式，我们就可以方面的使用代数中的矩阵技术进行分析及计算了。

如果一个子空间没有通过原点，那么从原点出发的向量必然首尾不顾，造成了向量头在子空间中尾在空间外（因为原点在空间外）。当然，向量的加法和数乘也都跑到子空间外面去了。如下图。




上图中，严格讲来向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 并不在平面 $S$ 中，因为只有向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的头部在平面上。但如果

$\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 向量是采用坐标解析式表达的话，我们就会认为这两个向量在平面上。即使这样，相量 $\boldsymbol{\alpha}$ 的数乘 $x\boldsymbol{\alpha}$ 超出了平面 $S$ 之上，$\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 相加的和向量 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 也超出了平面之上。因此，他们对向量的加法和数乘运算不封闭。所以平面 $S$ 不是二维向量子空间。

1、 **实际上，在三维几何向量空间中，凡是过原点的平面或直线上的全体向量组成的集合都是 $\mathbf{R}^3$ 的子空间，而不过原点的平面或直线上的全体向量组成的集合都不是 $\mathbf{R}^3$ 的子空间。**

## 4.2.3. 基、维数及其坐标的几何意义

### 基、维数、坐标的定义

对于向量空间 $V$ 中的一个有序向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$，若满足：

- $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n$ 线性无关；
- $V$ 中任意一个向量 $\boldsymbol{\alpha}$ 都可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n$ 线性表示：$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2...+x_n\boldsymbol{\alpha}_n$。

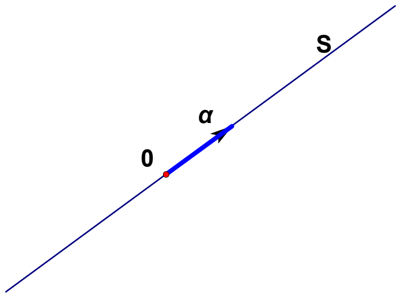
那么称向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$ 为向量空间 $V$ 的一个**基**；称向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$ 的元素个数 n 为向量空间 $V$ 的**维数**；称有序数组 $(x_1,x_2...x_n)$ 为向量 $\boldsymbol{\alpha}$ 在基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2...\boldsymbol{\alpha}_n\}$ 上的**坐标**。

### 基的几何意义

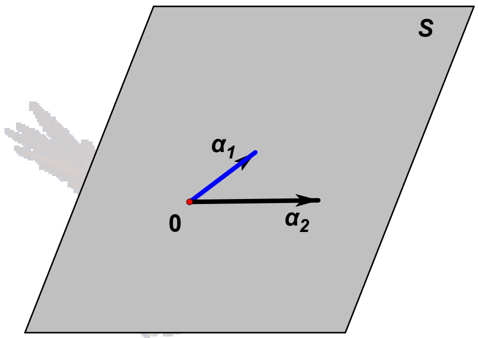
基是向量空间的一组很“结实”的向量集合，每一个基向量可以象房屋的地基的每一块石块一样支撑衍生出空间中的全部向量，因此，首先一个基能代表或衍生出空间里的所有的向量，缺一不可。其次，作为基的每一个向量都是个顶个，谁也不能代表谁，他们必须线性无关，它是一个最大无关向量组。

我们给一个向量空间找一个基，目的是为了给这个空间定一个坐标系，以方面我们定位和计算向量。一个基实际上就是选取的一个坐标系，另外一个基就是选取的一个新的坐标系。基是坐标系在线性空间中的推广。基向量对应坐标系的坐标轴，有几个基向量就有几个坐标轴 ，n 维空间的一个基就需要有 n 个基向量。下面我们看看 $\mathbf{R}^n$ 空间中的几个基的例子：

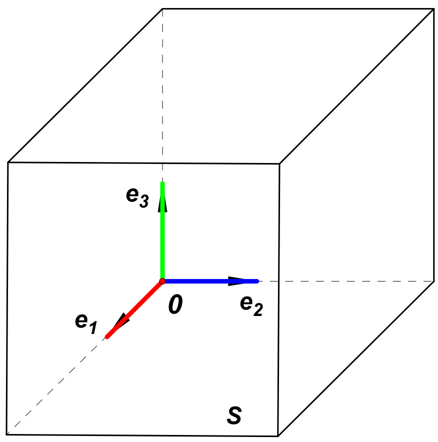
图 1 中，一维向量空间 $S$ 是一条过 0 点的直线，向量 $\boldsymbol{\alpha}\neq\mathbf{0}$ 并属于直线 $S$，因而可以讲 $S$ 是向量 $\boldsymbol{\alpha}\neq\mathbf{0}$ 张成的向量空间，$S$ =Span($\boldsymbol{\alpha}$)，所以向量组 $\{\boldsymbol{\alpha}\}$ 是向量空间 $S$ 的一个基。




在图 2 中，如果二维向量空间 $S$ 是 $R^3$ 中的一个平面，且 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 是平面 $S$ 上的任意的两个向量，其中任意一个都不是另外一个向量的倍数，因此向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ 线性无关 。因此平面 $S$ 可以看作是向量 $\boldsymbol{\alpha}_1$、 $\boldsymbol{\alpha}_2$ 张成的向量空间 $S = Span\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$，所以向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ 是向量空间的一个基。




在图 3 中，三维向量空间 S 是 $R^3$，三个标准单位向量 $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{(1,0,0),(0,1,0),(0,0,1)\}$，因为 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 彼此线性无关，可以生成 $R^3$，因此向量组 $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ 是 $R^3$ 的一个基。这个基的基向量是由标准单位向量组成，因此又称为标准基。



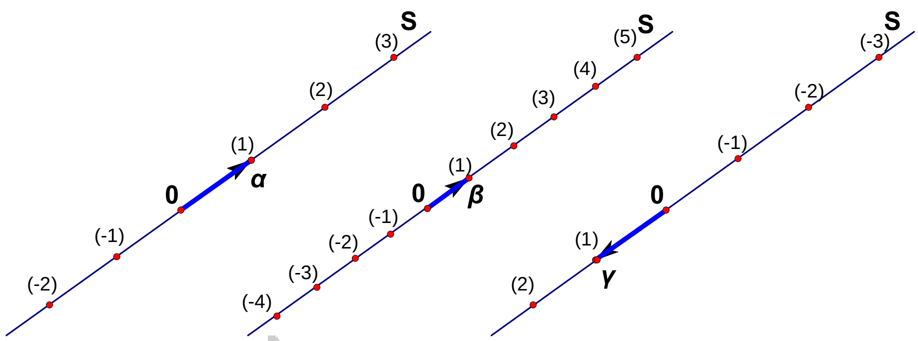

坐标与维数的几何意义

一个基包含的向量个数就是坐标轴的个数，也就是向量空间的维数。维数是空间的一个本质特征，它不依赖于基的选取。无论你怎么选取不同的基，但基向量的个数不会改变，维持支撑空间的维数不会改变。这就是为何称之为“维数”的原因。

一个向量空间的基选定后，其坐标是什么？如何求取？下面我们接着看看几个图示的例子。

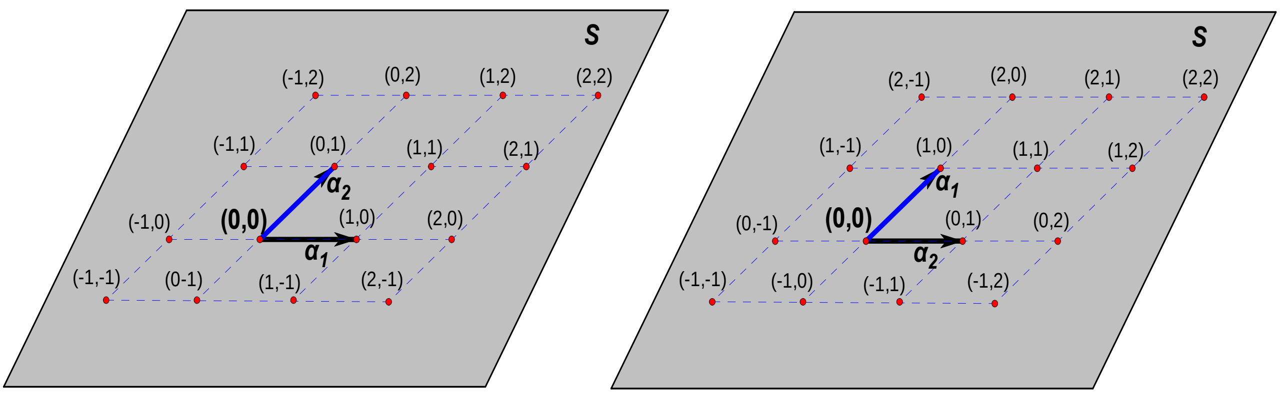
**一维基及其坐标：**

一维空间 $S$，当选取基为$\{\boldsymbol{\alpha}\}$时，坐标选取如图左；当选取的新基向量$\boldsymbol{\beta}$为$\boldsymbol{\alpha}/2$时，坐标刻度的密度加大一倍，如图中；当选取的新基向量$\boldsymbol{\gamma}$为$-\boldsymbol{\alpha}$时，坐标轴方向也随之反转，如图右。




**二维基及其坐标：**

二维空间 $S$，当选取基为$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$时，坐标选取如图左。两个坐标轴是与向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$重合的，刻度的划分是遵循向量加法的平行四边形规则（注意，在这里我们不要沿袭笛卡尔坐标系的习惯，试图把空间中的一点（对应一个向量）向坐标轴作垂直投影。当然，你如果坚持这样做的话，你就会发现基的施密特正交化方法的秘密）。



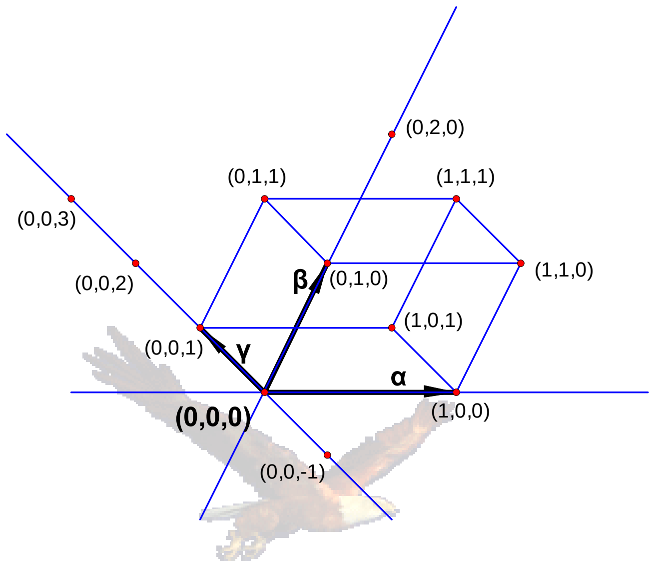

另外，在二维空间中，确定基向量的顺序是必要的。在向量组的讨论中我们不强调向量组中向量们的顺序，但作为一个基的向量组就要有顺序，显然，如果基向量顺序进行了调整，坐标值也相应进行调整。在图右中，我们把空间 $S$ 的基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$调换了顺序成为一个新的基$\{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1\}$，当然空间中的坐标也变了。

另外，我们在上述的例子中也看到了基与我们的直角坐标系的不同，两个基向量不一定垂直；

在刻画坐标网络时不是直角坐标系的垂直投影，而是平行四边形坐标网络，分割一个坐标轴的坐标线是与另外一个坐标轴平行的关系。一个基向量的方向是对应坐标轴的正方向，坐标单位是基向量的长度。

三维基及其坐标：

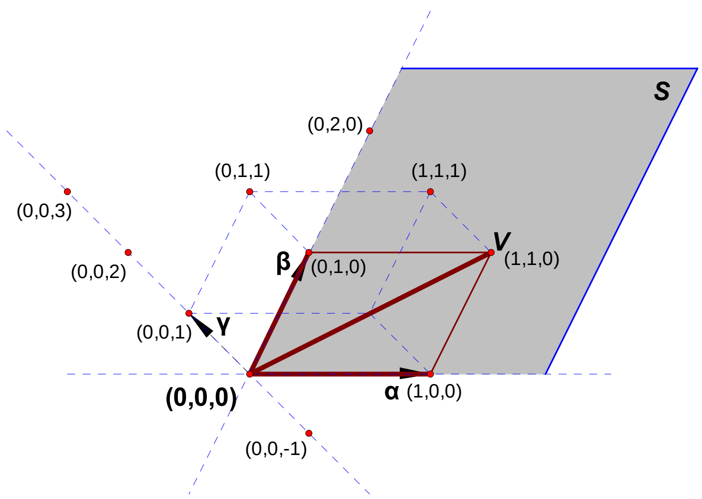
三维空间的一个基包含了三个线性无关的向量$\{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\}$，空间以$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$为基的坐标刻画满足平行六面体法则，如向量$(1,1,1)$是与原点相对应的平行六面体的对角点。




下面对于一个三维空间中的二维子空间，我们看看他的基及其坐标是如何刻画的。

三维空间的子空间的基及其坐标：

在上面的三维空间的例子中，向量$\mathbf{V}$的坐标是$(1,1,0)$。如果我们要研究由向量组$\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$张成的子空间$S=Span\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$，这个二维的子空间现以向量组$\boldsymbol{\alpha},\boldsymbol{\beta}$为基，那么向量$\mathbf{V}$在$S$中的坐标是什么？从图中可以看出，向量$\mathbf{V}$坐标是$(1,1)$。



总结一下：向量 **V** 在三维空间 $Span\{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\}$ 中的 **B** 坐标是 $(1,1,0)$，其中 **B**=$\{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\}$；而向量 **V** 在二维空间 $Span\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ 中的 **B** 坐标变成了 $(1,1)$，其中 **B**=$\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$。

如果向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 是属于三维欧几里德空间的任意一组线性无关组，向量形式是由笛卡尔坐标给出。这个三维欧几里德空间如果使用{$\{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\}$}作为基，那么一个向量如何求其 **B** 坐标？如果这个向量也在 $Span\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ 二维空间中，又如何求取以 $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ 为基的 **B** 坐标呢？这个问题就是我们下面要讨论的基变换。

**注：** 空间坐标系

前面说过，建立坐标系的目的就是把空间向量的线性变换转化为坐标的运算。在我们讨论向量和矩阵以及向量方程中仿射坐标系和直角坐标系最为有用。

在空间中任取一点 0，以点 0 为起点任意作三个不共面的向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$，这就建立了一个仿射坐标系，记为 $\{0;\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3\}$，有了这个仿射坐标系，我们就可以建立空间向量和三元有序数组（即坐标值）之间的一一对应的关系了。就是说，在坐标系中的任意向量 **a** 可用唯一的有序数组 $(a_1,a_2,a_3)$ 来表示关系：$\mathbf{a}=a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2+a_3\boldsymbol{\varepsilon}_3$。

在这里，点 0 称为坐标原点，$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 称为坐标向量或基，过原点且与坐标向量同向的直线称为坐标轴，有序数组 $(a_1,a_2,a_3)$ 称为向量 **a** 的坐标。

仿射坐标系按手征性分为左手坐标系和右手坐标系。

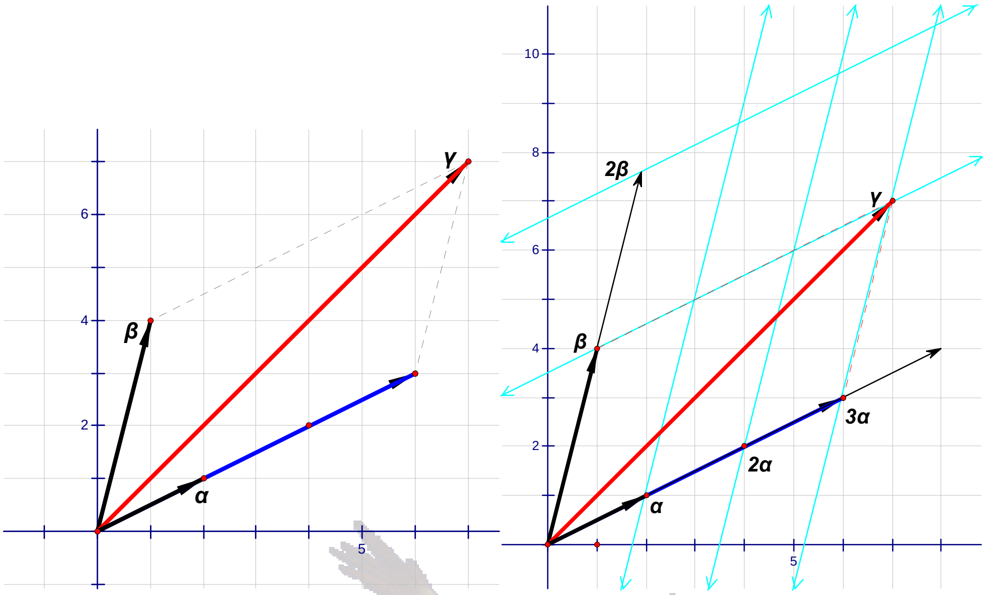
特别的，如果仿射坐标系 $\{\mathbf{o},\mathbf{i},\mathbf{j},\mathbf{k}\}$ 的坐标向量 $\mathbf{i},\mathbf{j},\mathbf{k}$ 是两两垂直的单位向量，则称之为直角坐标系。$\mathbf{i},\mathbf{j},\mathbf{k}$ 所在的坐标系分别称为 x 轴、y 轴和 z 轴。我们常用的直角坐标系为右手直角坐标系。

**基及其坐标的例子：**

在直角平面坐标系（实际上为单位正交基 $\mathbf{i},\mathbf{j}$）的空间中，有两个向量 $\boldsymbol{\alpha}=(2,1)$ 和 $\boldsymbol{\beta}=(1,4)$。这两个向量线性无关（不成倍数关系），我们可以让他们构成平面空间一对新基。那么有另外一个向量 $\boldsymbol{\gamma}=(7,7)$ 可以写为这两个新基的线性组合：

$$\boldsymbol{\gamma} = 3\boldsymbol{\alpha} + \boldsymbol{\beta}$$

因此，向量 $\boldsymbol{\gamma}$ 相对于基 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的坐标向量为 $(3,1)$。图解如下：




## 4.2.4. 基变换的几何意义

前面的讨论告诉我们，在一个 n 维的线性空间中，可以取不同的向量组（每个向量组含有 n 个线性无关的向量）作为基。那么这个线性空间的任意两个基之间必有关联，这个关联是什么？还有，一个向量 **a** 在一个确定的基下有一个确定的坐标，这个向量在不同的基下有不同的坐标。第二个问题是，任意两个基上的坐标之间有什么关联？

一个 n 维线性空间 $\mathbf{V}$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots\boldsymbol{\beta}_n$ 是 $\mathbf{V}$ 的两个基。先将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots\boldsymbol{\alpha}_n$ 作为空间的基，那么向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots\boldsymbol{\beta}_n$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots\boldsymbol{\alpha}_n$ 线性表示：

$$\boldsymbol{\beta}_1 = a_{11}\boldsymbol{\alpha}_1 + a_{12}\boldsymbol{\alpha}_2 + \ldots a_{1n}\boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n)\begin{pmatrix} a_{11} \\ a_{12} \\ \cdots \\ a_{1n} \end{pmatrix} = (a_{11}, a_{12}, \ldots a_{1n})\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix}$$

$$\boldsymbol{\beta}_2 = a_{21}\boldsymbol{\alpha}_1 + a_{22}\boldsymbol{\alpha}_2 + \ldots a_{2n}\boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n)\begin{pmatrix} a_{21} \\ a_{22} \\ \cdots \\ a_{2n} \end{pmatrix} = (a_{21}, a_{22}, \ldots a_{2n})\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix}$$

........................

$$\boldsymbol{\beta}_n = a_{n1}\boldsymbol{\alpha}_1 + a_{n2}\boldsymbol{\alpha}_2 + \ldots a_{nn}\boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n)\begin{pmatrix} a_{n1} \\ a_{n2} \\ \cdots \\ a_{nn} \end{pmatrix} = (a_{n1}, a_{n2}, \ldots a_{nn})\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix}$$

将以上 n 个表示式合并为：

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n)\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

或者

$$\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \cdots \\ \boldsymbol{\beta}_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix}$$

上式 1 或 2 称为由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots \boldsymbol{\beta}_n$ 的基变换公式，其中，矩阵 $\mathbf{P} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$ 或 $\mathbf{P}' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ 称为由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots \boldsymbol{\beta}_n$ 的过渡矩阵。

注意两点：

1、$\mathbf{P}$ 和 $\mathbf{P}'$ 互为转置矩阵，这取决于你对基向量 $\boldsymbol{\alpha}_i$ 和 $\boldsymbol{\beta}_i$ 看作是行向量还是列向量；

2、过渡矩阵 $\mathbf{P}$ 的列向量和 $\mathbf{P}'$ 的行向量分别是 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots \boldsymbol{\beta}_n$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n$ 上的坐标（按序排列成矩阵），换句话说，把 $\boldsymbol{\beta}_i$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_n$ 上的坐标一列列的排列而成过渡矩阵 $\mathbf{P}$。

以上得到的矩阵 $\mathbf{P}$ 和 $\mathbf{P'}$ 给出了第一个问题的答案，就是线性空间的两个基之间是可以互相转换或变换的，变换的矩阵称为过渡矩阵。下面我们看看一个向量的坐标的转换是什么？

设向量 $\mathbf{a}$ 在两个基下 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...\boldsymbol{\beta}_n$ 的坐标分别是 $(x_1,x_2,...,x_n)$ 和 $(x_1^{'},x_2^{'},...,x_n^{'})$，则它们之间的关系可有：

向量 $\mathbf{a}$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...\boldsymbol{\alpha}_n$ 上表示为 $\mathbf{a}=(x_1,x_2,...,x_n)\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \cdots\\ \boldsymbol{\alpha}_n\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...\boldsymbol{\alpha}_n)\begin{pmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{pmatrix}$ ----------1;

同时，向量 $\mathbf{a}$ 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...\boldsymbol{\beta}_n$ 上表示为 $\mathbf{a}=(x_1^{'},x_2^{'},...,x_n^{'})\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \cdots\\ \boldsymbol{\beta}_n\end{pmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...\boldsymbol{\beta}_n)\begin{pmatrix}x_1^{'}\\ x_2^{'}\\ \cdots\\ x_n^{'}\end{pmatrix}$ -----2;

把前面的基变换公式代入 2 得到

$$\mathbf{a}=(x_1^{'},x_2^{'},...,x_n^{'})\mathbf{P}'\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \cdots\\ \boldsymbol{\alpha}_n\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...\boldsymbol{\alpha}_n)\ \mathbf{P}\begin{pmatrix}x_1^{'}\\ x_2^{'}\\ \cdots\\ x_n^{'}\end{pmatrix} \text{-----------------3}$$

因为一个向量在同一个基下的坐标是唯一的，因此，对比上式 3 和 1 式中的坐标部分，得到：

$$(x_1,x_2,...,x_n)=(x_1^{'},x_2^{'},...,x_n^{'})\mathbf{p}'$$

或

$$\begin{pmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{pmatrix}=\mathbf{P}\begin{pmatrix}x_1^{'}\\ x_2^{'}\\ \cdots\\ x_n^{'}\end{pmatrix}$$

至此，我们得到了坐标转换的公式。

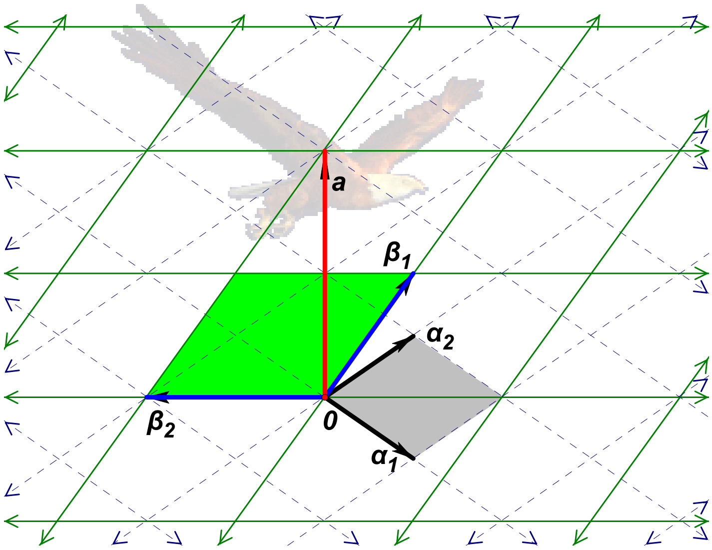
假设已知坐标 $(x_1^{'},x_2^{'},...,x_n^{'})$，直接使用上式及得到坐标 $(x_1,x_2,...,x_n)$。反之，如果已知 $(x_1,x_2,...,x_n)$ 求 $(x_1^{'},x_2^{'},...,x_n^{'})$，即公式：

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \\ \cdots \\ x_n^{'} \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

下面我们来个例子来了解具体的基过渡矩阵和坐标转换计算。看下图说话。

在 2 维平面空间的一组基为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$（没有具体值），空间一向量 $\mathbf{a}$，由图知 $\mathbf{a} = (-2, 2)$；我们再取一组基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$，基向量 $\boldsymbol{\beta}_i$ 在原来基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 上的坐标分别是 $(-\frac{1}{2}, \frac{3}{2}), (-1, -1)$，具体即：

$$\boldsymbol{\beta}_1 = -\frac{1}{2}\boldsymbol{\alpha}_1 + \frac{3}{2}\boldsymbol{\alpha}_2$$
$$\boldsymbol{\beta}_2 = -\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2$$




因此我们得到了由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 过渡到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 的过渡矩阵（由 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 的坐标按列排列得到）为：

$$\mathbf{P} = \begin{bmatrix} -\frac{1}{2} & -1 \\ \frac{3}{2} & -1 \end{bmatrix}$$

过渡矩阵对不对，我们验证一下，用过渡矩阵求向量 $\mathbf{a}$ 在基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 上的坐标，用坐标转换公式：

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \mathbf{P}^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} -\dfrac{1}{2} & -1 \\ \dfrac{3}{2} & -1 \end{bmatrix}^{-1}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ -\dfrac{3}{2} & -\dfrac{1}{2} \end{bmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

看看图上，向量 $\mathbf{a}$ 在基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 的坐标网络中的坐标正是 $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$，过渡矩阵和坐标转换公式得到验证。

## 4.2.5. 标准正交基的几何解释

标准正交基也叫规范正交基。实际上，如果这些基向量互相垂直，就叫正交基，而且每个基向量的长度等于单位 1 的话，那么这个基就叫标准正交基。

- 标准正交基的好处

为什么要得到标准正交基呢？主要原因是如果基是正交且标准的，就很容易计算向量子空间的投影和基坐标。换句类似的话是说，如果你取的坐标系是垂直的，而且取的坐标单位是 1，就很容易计算向量空间里面的向量的坐标值。

若 $\mathbf{e}_1, \ldots \mathbf{e}_r$ 是子空间的 $V$ 的一个标准正交基，则 $V$ 中任一向量 $\boldsymbol{\alpha}$ 都能有标准正交基 $\mathbf{e}_1, \ldots \mathbf{e}_r$ 线性表示：

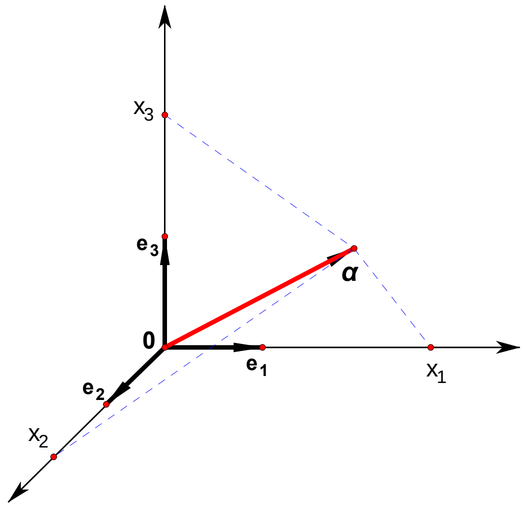
$$\boldsymbol{\alpha} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_r\mathbf{e}_r$$

其中 $(x_1, x_2, \ldots x_r)$ 就是向量 $\boldsymbol{\alpha}$ 的坐标。

如果想求坐标 $(x_1, x_2, \ldots x_r)$ 中一个 $x_i$，只要把向量 $\boldsymbol{\alpha}$ 向对应的基元 $\mathbf{e}_r$ 投影并求投影长度即可得到，求投影长度就是求内积，也就是公式

$$x_i = [\boldsymbol{\alpha}, \mathbf{e}_r] = \boldsymbol{\alpha}\mathbf{e}_r^{\mathrm{T}}$$

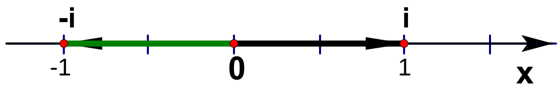
下图给出了三维向量空间的一个标准正交基的图示。向量 $\boldsymbol{\alpha}$ 分别对基坐标轴上的投影分向量分别是 $x_1\mathbf{e}_1$， $x_2\mathbf{e}_2$ 和 $x_3\mathbf{e}_3$，对应的坐标分别是 $x_1$， $x_2$ 和 $x_3$。




坐标计算公式方便、简捷，是因为向量空间的基是标准正交基。所以，我们在给出向量空间的德基时总是求取规范正交基。实际上，我们最常用的笛卡尔坐标系就是标准正交基的坐标系，只是在使用中忽略掉了 x，y 和 z 轴上的单位向量罢了。

- 一维空间的标准正交基

一维的标准正交基只有两个：$\mathbf{i}=\{(1)\}$ 和 $-\mathbf{i}=\{(-1)\}$。在 X 直线上的任意一个向量 $(x)$ 都可以表示成为 $\mathbf{i}$ 的倍数，其坐标是 $x$；任意向量 $(x)$ 也能表示成为 $-\mathbf{i}$ 的倍数，其坐标是 $-x$。




- 二维空间的标准正交基

二维欧几里德空间的单位坐标向量组 $\{(1,0),(0,1)\}$ 就是一个标准正交基。除了这个标准正交基外还有其他的标准正交基吗?

有！比如把向量组内的向量位置交换一下：$\{(0,1),(1,0)\}$，就是一个新的标准正交基。

还有！比如下面向量组也是一个标准正交基：

$$\{\left(\frac{\sqrt{3}}{2},\frac{1}{2}\right),\left(-\frac{1}{2},\frac{\sqrt{3}}{2}\right)\}$$

验证一下：

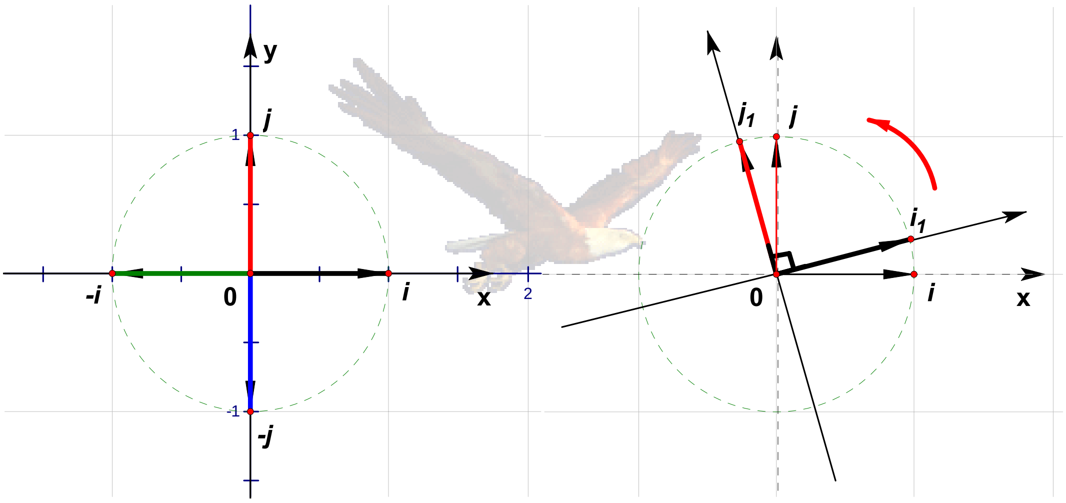
两个向量内积：$\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)\begin{pmatrix}-\dfrac{1}{2}\\ \dfrac{\sqrt{3}}{2}\end{pmatrix}=0$， 内积为 0，两个向量正交。

向量长度：$\left\|\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)\right\|=\sqrt{\left(\dfrac{\sqrt{3}}{2}\right)^2+\left(\dfrac{1}{2}\right)^2}=1$，$\left\|\left(-\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)\right\|=\sqrt{\left(-\dfrac{1}{2}\right)^2+\left(\dfrac{\sqrt{3}}{2}\right)^2}=1$；长度为 1，单位长度，因此两个向量都是单位向量。

实际上，我们可以使用这样一种方法来得到无数个包括上述的二维空间中的标准正交基，就是对二维笛卡尔单位坐标向量进行镜像或旋转即可得到。

二维平面的笛卡儿坐标空间中与坐标轴重合的单位向量有四个向量（如图）：

$$\mathbf{i}=(1,0),\quad \mathbf{j}=(0,1),\quad -\mathbf{i}=(-1,0),\quad -\mathbf{j}=(0,-1)$$

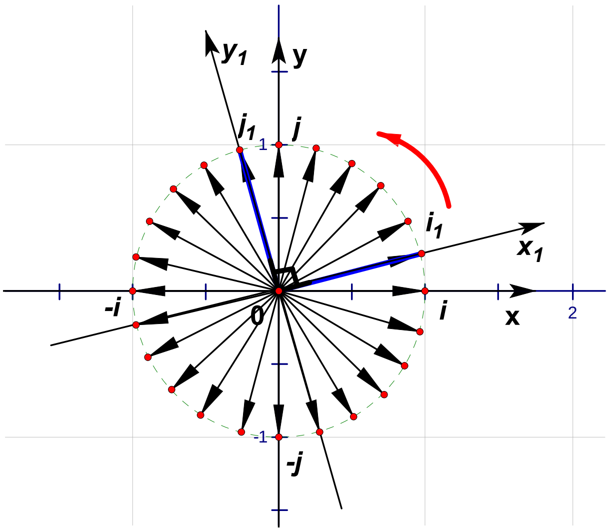



其中，$\mathbf{i}$ 和 $-\mathbf{i}$ 互为镜像，$\mathbf{j}$ 和 $-\mathbf{j}$ 也互为镜像。这四组单位向量中的任意两个向量如果满足正交要求，那么这两个向量组成的有序向量组就构成一个标准正交基。所有的八个与坐标轴重合的标准正交基列举如下

$$\{\mathbf{i},\mathbf{j}\},\ \{-\mathbf{i},\mathbf{j}\},\ \{\mathbf{i},-\mathbf{j}\},\ \{-\mathbf{i},-\mathbf{j}\},\ \{\mathbf{j},\mathbf{i}\},\ \{-\mathbf{j},\mathbf{i}\},\ \{\mathbf{j},-\mathbf{i}\},\ \{-\mathbf{j},-\mathbf{i}\}。$$

另外，要得到不与坐标轴重合的标准正交基，我们可以把上述八个标准正交基进行旋转就可以得到新的标准正交基。比如把上图中标准正交基 $\{\mathbf{i},\mathbf{j}\}$ 中的向量 $\mathbf{i}$ 和 $\mathbf{j}$，同时进行拟时针旋转到 $(\mathbf{i}_1,\mathbf{j}_1)$（保持垂直角度不变），那么 $(\mathbf{i}_1,\mathbf{j}_1)$ 也是一个标准正交基，如果旋转角等于 $\dfrac{\pi}{6}$，这个标准正交基就是前边说的（$\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)$，$\left(-\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)$）。

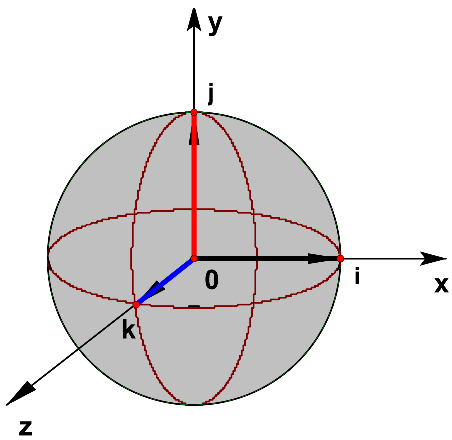
实际上，如果用一个通用的向量表达式表示出二维空间 $R^2$ 上的标准正交基就是 $\{(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\}$，其中 $\theta$ 为基向量逆向旋转的角度。这正是单位园的表示式。




- 三维空间的标准正交基

对于上述的二维标准正交基的取法，我们有一个统一的几何图形，就是在把标准正交基 $\{\mathbf{i}, \mathbf{j}\}$ 保持其角度垂直，分别进行拟时针和顺时针旋转，就可以得到全部的二维标准正交基。

类似的，三维空间的标准正交基是由单位坐标向量组 $\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ 在保持彼此的相对正交位置的同时进行全方位任意旋转，即可得到全部的标准正交基。显然，全部的标准正交基向量（无数的）的末端组成一个单位球面。




求取规范正交基的一个著名方法是施密特正交化方法，下面我们将进行讨论。

### 4.2.6. 施密特正交化的几何解释

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_r$是向量空间$V$的任意一个基，我们要在这组基上重新构造出一个新基，新的基要求是一个标准正交基。换句话讲，要找一组两两正交的单位向量组$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_r$，这个向量组与原基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_r$等价。把向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_r$化为$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_r$的过程就是规范正交化。过程及公式如下：

第一步正交化（施密特正交化）：

令

$$\begin{aligned}
&\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\\
&\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1,\\
&\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{\beta}_1,\boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1-\frac{[\boldsymbol{\beta}_2,\boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_2,\boldsymbol{\beta}_2]}\boldsymbol{\beta}_2,\\
&\ldots\ldots,\\
&\boldsymbol{\beta}_r=\boldsymbol{\alpha}_r-\frac{[\boldsymbol{\beta}_1,\boldsymbol{\alpha}_r]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1-\frac{[\boldsymbol{\beta}_2,\boldsymbol{\alpha}_r]}{[\boldsymbol{\beta}_2,\boldsymbol{\beta}_2]}\boldsymbol{\beta}_2-\cdots-\frac{[\boldsymbol{\beta}_{r-1},\boldsymbol{\alpha}_r]}{[\boldsymbol{\beta}_{r-1},\boldsymbol{\beta}_{r-1}]}\boldsymbol{\beta}_{r-1},
\end{aligned}$$

第二步单位化：

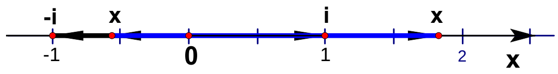
$$\begin{aligned}
&\mathbf{e}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}\\
&\mathbf{e}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}\\
&\ldots\ldots,\\
&\mathbf{e}_r=\frac{\boldsymbol{\beta}_r}{\|\boldsymbol{\beta}_r\|}
\end{aligned}$$

至此，我们得到了向量空间$V$的一个规范正交基$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_r$。

下面我们看看求取一至三维向量空间基的规范正交化几何意义是什么？

- 一维向量空间的规范正交基求法的意义：

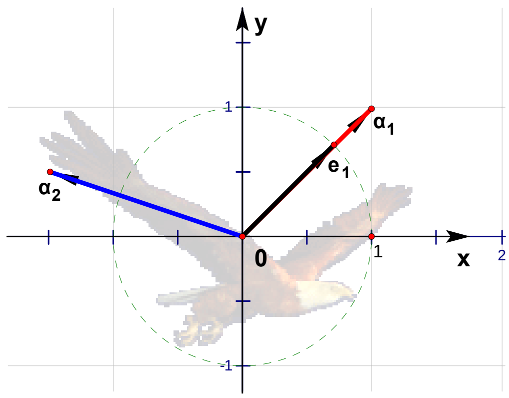
一维的向量中只有一个元素$x$，设$\{(x)\}$是一维向量空间的一个基，显然，对这个向量进行规范化即得到规范正交基：$\{(\frac{x}{|x|})\}$。




上图中画出了两个方向的 $(x)$ 向量，向量 $(x)$ 的长度被压缩或放大到单位长度 1，方向保持不变。

- 二维向量空间的规范正交基求法的几何意义：

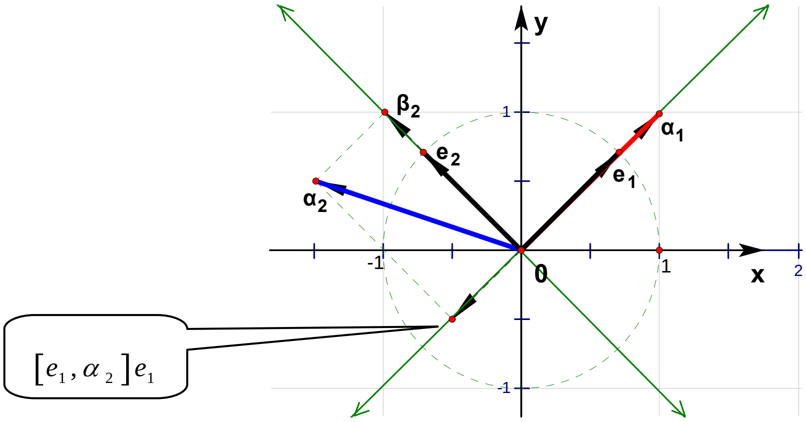
从前面的分析知道，二维向量空间的规范基集中在平面坐标系上单位圆上（自由向量全部从原点出发），设（$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$）是二维向量空间上任意的一个基（如图），设想由这个任意基通过斯密特方法得到一个规范正交基，就是想办法用向量的伸缩和加减计算把向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 变化到单位圆上的新向量 $\mathbf{e}_1,\mathbf{e}_2$，而且新向量 $\mathbf{e}_1,\mathbf{e}_2$ 要互相垂直。




再来看看施密特正交化公式的变形：

$$\beta_2=\alpha_2-\frac{[\beta_1,\alpha_2]}{[\beta_1,\beta_1]}\beta_1=\alpha_2-\frac{[\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2]}{\|\boldsymbol{\beta}_1\|^2}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\left[\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|},\boldsymbol{\alpha}_2\right]\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}=\boldsymbol{\alpha}_2-[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1,$$

在上式中，$[\mathbf{e}_1,\boldsymbol{\alpha}_2]$ 是单位向量 $\mathbf{e}_1$ 和 $\boldsymbol{\alpha}_2$ 的内积，即向量 $\boldsymbol{\alpha}_2$ 在 $\mathbf{e}_1$ 上的坐标值；因而 $[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1$ 是向量 $\boldsymbol{\alpha}_2$ 在 $\mathbf{e}_1$ 上投影分向量。向量 $\boldsymbol{\alpha}_2$ 减去自己的在 $\mathbf{e}_1$ 上投影 $[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1$，则得到一个正交于向量 $\mathbf{e}_1$ 的分向量 $\boldsymbol{\beta}_2$（如图）。




最后对正交向量 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\beta}_2$ 单位化，得到了位于单位圆上的标准正交基 $\mathbf{e}_1,\mathbf{e}_2$。

- 三维向量空间的规范正交基求法的几何意义：

类似的，三维向量空间基的规范正交化过程是二维的推广。首先给出施密特正交化的变形公式：

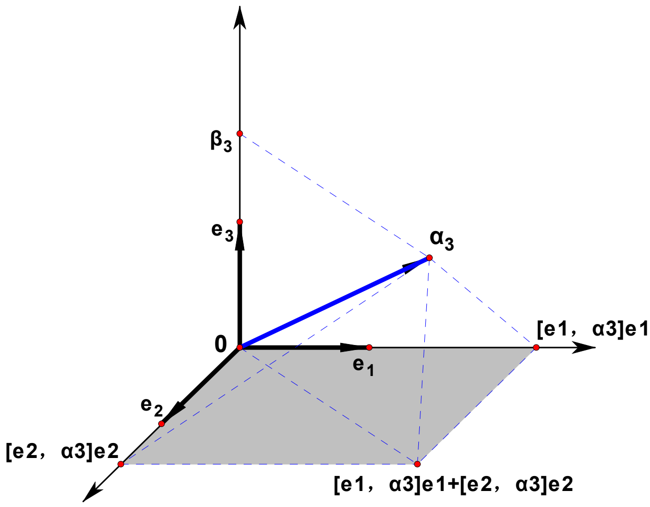
$$\begin{aligned}
&\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\\
&\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1,\\
&\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1-[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2
\end{aligned}$$

前面两个公式是二维空间的公式，我们已经讨论过。对于新增的第三个公式的几何含义，我们不难看出：

向量 $\boldsymbol{\alpha}_3$ 是减去自己在二维正交规范基上的投影得到。其中 $[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1$ 是 $\boldsymbol{\alpha}_3$ 在 $\mathbf{e}_1$ 上的投影，$[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2$ 是 $\boldsymbol{\alpha}_3$ 在 $\mathbf{e}_2$ 上的投影。$[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1+[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2$ 是 $\boldsymbol{\alpha}_3$ 在以 $\{\mathbf{e}_1,\mathbf{e}_2\}$ 为基的子空间（平面）上的投影。

这里，向量 $\boldsymbol{\alpha}_3$ 减去自己在 $\mathbf{e}_1$ 上的投影的差向量必垂直于 $\mathbf{e}_1$，同时也减去自己在 $\mathbf{e}_2$ 上的投影的差向量亦必垂直于 $\mathbf{e}_2$，因而总的差向量 $[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1+[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2$ 必然垂直于 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 张成的平面空间。

为清楚起见，下图中没有画出 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 等向量。




- r 维向量空间的规范正交基求法的几何意义：

至此，多维向量空间的规范正交几何意义显而易见。从几何作图方便性的角度出发，我们建议把施密特正交规范化公式修改一下，不再需要过度向量$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots\boldsymbol{\beta}_r$，公式如下：

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|},\\
\mathbf{e}_2 &= \frac{\boldsymbol{\alpha}_2-[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1}{\|\boldsymbol{\alpha}_2-[\mathbf{e}_1,\boldsymbol{\alpha}_2]\mathbf{e}_1\|},\\
\mathbf{e}_3 &= \frac{\boldsymbol{\alpha}_3-[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1-[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2}{\|\boldsymbol{\alpha}_3-[\mathbf{e}_1,\boldsymbol{\alpha}_3]\mathbf{e}_1-[\mathbf{e}_2,\boldsymbol{\alpha}_3]\mathbf{e}_2\|},\\
&\ldots\ldots,\\
\mathbf{e}_r &= \frac{\boldsymbol{\alpha}_r-[\mathbf{e}_1,\boldsymbol{\alpha}_r]\mathbf{e}_1-[\mathbf{e}_2,\boldsymbol{\alpha}_r]\mathbf{e}_2-\ldots-[\mathbf{e}_{r-1},\boldsymbol{\alpha}_r]\mathbf{e}_{r-1}}{\|\boldsymbol{\alpha}_r-[\mathbf{e}_1,\boldsymbol{\alpha}_r]\mathbf{e}_1-[\mathbf{e}_2,\boldsymbol{\alpha}_r]\mathbf{e}_2-\ldots-[\mathbf{e}_{r-1},\boldsymbol{\alpha}_r]\mathbf{e}_{r-1}\|}
\end{aligned}
$$

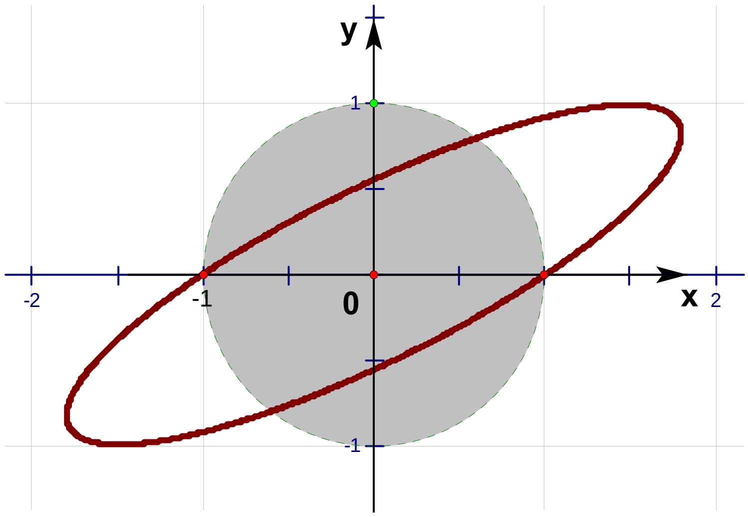
上式看起来比较复杂，其实概念清晰自然，计算方便。

## -----图解线性代数-----

# 线性代数的几何意义

# 之（5 上）

任广千　编著




2010．08．16

# 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

# 第五章　矩阵的几何意义

通过前面的章节我们初步了解到，解线性方程组的克莱姆法则使用了行列式理论，但克莱姆法则只能用于解方程个数等于未知数个数的方程组，而且系数行列式不能等于 0。即使以上条件都满足，也要计算 n+1 个 n 阶行列式。实际工程中的 n 一般很大，即使在现代计算机技术面前，计算效率也不能使人满意。

在用消元法解各种类型的线性方程组时，一系列问题出现了：当系数行列式等于 0 时 ，方程组是否有解？若有解又如何求出？当未知量个数与方程的个数不等时，线性方程组的解又如何？

要深入探讨这些问题除了向量概念外还需要引入矩阵的理论。到 1858 年，哈密尔顿（W.R.Hamilton）和凯莱（A.Cay-ley）的著作中出现了矩阵的运算，从行列式到矩阵的出现，大约经过了 100 多年的时间。

我们知道，在直角坐标系中，一个有序的实数数组 $(a,b)$ 和 $(a,b,c)$ 分别代表了平面上和空间上的一个点，这就是实数组的几何意义。类似的，在线性空间中如果确定了一个基，线性映射就可以用确定的矩阵来表示，这就是矩阵的几何意义：线性空间上的线性映射。

矩阵独立的几何意义表现为对向量的作用结果。矩阵对一个向量是如何作用的？矩阵对多个向量是如何作用的？矩阵对一个几何图形（由无数向量组成的几何图形）是如何作用的？在矩阵对一个几何图形的作用研究中，我们会发现一些有规律的东西比如特征向量、秩等等。

## 5.1.　矩阵的概念

矩阵的本质就是一个长方形的数表，在生活中的所有长方形数表都可以看成是矩阵。矩阵和行列式相似，也是以行和列组织的矩形数字阵列，因此称为矩阵。与行列式的表示法不同，矩阵是用方括号把数字块括起来，表示一个有顺序有组织的数据块；而行列式是对这些数据块进行的一个运算，是一个算式，故称为行列式。矩阵的一个三阶例子如下：

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

如果用数组来统一定义标量、向量和矩阵的话就是：标量是一维向量，向量是标量的数组，矩阵则是向量的数组。例如上面介绍的矩阵我们如果使用列向量 $\mathbf{a}=(a_1,a_2,a_3)^T$ ，$\mathbf{b}=(b_1,b_2,b_3)^T$ ，$\mathbf{c}=(c_1,c_2,c_3)^T$ 来表示它，这个矩阵就可以写作：$[\mathbf{a}\ \ \mathbf{b}\ \ \mathbf{c}]$。

当然，矩阵不只是只有几何意义，也具有现实的物理意义，矩阵的运算也都可以从实践中找到。下面有个例子：

比如某家电器公司的制造厂有几个生产线，产线在 2009 年和 2010 年的上半年的产出量的统计表如下：

| 顺序 | 产线名 | 2009 年上半年的每月产出量 |
|---|---|---|

| | | 1 月 | 2 月 | 3 月 | 4 月 | 5 月 | 6 月 |
|---|---|---|---|---|---|---|---|
| 1 | 冰箱线 | 22 | 35 | 30 | 23 | 25 | 12 |
| 2 | 吸尘器线 | 25 | 43 | 32 | 34 | 35 | 30 |
| 3 | 电视线 | 23 | 23 | 34 | 44 | 40 | 45 |

| 顺序 | 产线名 | 2010 年上半年的每月产出量 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 月 | 2 月 | 3 月 | 4 月 | 5 月 | 6 月 |
| 1 | 冰箱线 | 22 | 34 | 30 | 23 | 25 | 12 |
| 2 | 吸尘器线 | 24 | 43 | 32 | 34 | 35 | 34 |
| 3 | 电视线 | 23 | 23 | 34 | 45 | 41 | 45 |
| 4 | 手机线 | 34 | 34 | 35 | 45 | 23 | 43 |
| 5 | VCD 线 | 45 | 24 | 31 | 34 | 45 | 12 |

我们将第一个表格对应的矩阵记为 $\mathbf{A}$，第二个表格对应的矩阵记为 $\mathbf{B}$，则有：

$$\mathbf{A}=\begin{bmatrix}22&35&30&23&25&12\\25&43&32&34&35&30\\23&23&34&44&40&45\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix} \text{和} \mathbf{B}=\begin{bmatrix}22&34&30&23&25&12\\24&43&32&34&35&34\\23&23&34&45&41&45\\34&34&35&45&23&43\\45&24&31&34&45&12\end{bmatrix}。$$

那么有：

$\mathbf{A}+\mathbf{B}$ 实际意义是：2009、2010 年 1~6 月各产线每月产量的和(2001 年手机，VCD 机的产量为 0)；

$\mathbf{B}-\mathbf{A}$ 实际意义是： 2010 年 1~6 月各产线每月产量比上年同期的增产情况；

设冰箱、吸尘器、电视机、手机、VCD 机的价格分别是 l500 元台、900 元/台、300 元/台、2500 元/台、980 元/台，则有向量 $\mathbf{c}=(1500\ \ 900\ \ 3000\ \ 2500\ \ 980)$，那么矩阵乘法：

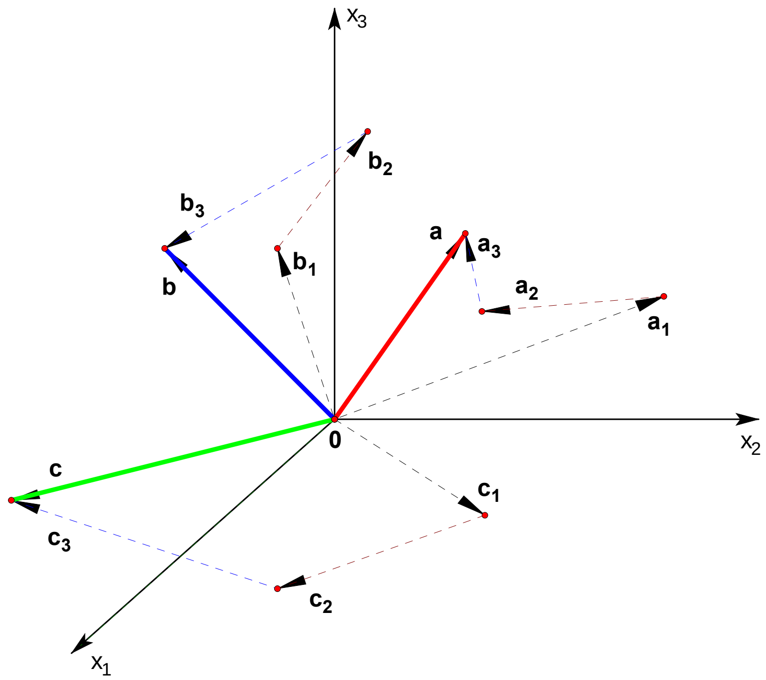
$\mathbf{cB}$ 的实际意义是：2010 年 1~6 月该制造厂每种产品的月产值。

如果每生产一件产品职工得到的奖励积分为 3 分，则数乘运算 $3\mathbf{A}$ 实际意义是 2009 年 1~6 月该厂各车间的职工月积分。

## 5.2. 矩阵加法的几何意义

矩阵的加法和乘法等简单运算可看作来自于线性方程组的简单运算，读者可以参看第七章的 7.1 节的详细介绍。在下面介绍矩阵的加法和乘法几何意义时，我们仍然不能离开向量的有力帮助。矩阵中的行向量或列向量的意义可以有效地帮助我们看清矩阵所蕴含的几何变换的意义。

多个矩阵的加法比较简单，即使不用给出几何意义，我们也能轻松掌握它。不过画出矩阵加法的几何图形，可以帮助你对多个向量所组成的几何图形的叠加有个形象的认知。




上图显示了三组向量同时连加，每组有三个向量的分量连加。把上述用矩阵表述出来就是：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ c_{11} & c_{12} & c_{13} \end{bmatrix} + \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} + \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

我们把上述每个矩阵都分解为三个行向量来给出图形的，其实因为矩阵的加法是对每个元素分别对应相加，因此对于列向量同样等效。

## 5.3. 矩阵与向量的乘法的几何意义

矩阵与向量乘积比如 $\mathbf{Ax}$ 表现为矩阵 $\mathbf{A}$ 对一个向量 $\mathbf{x}$ 作用的结果。其作用的主要过程是对一个向量进行旋转和缩放的综合过程（即线性变换的过程），一个向量就变换为另外一个向量。一个 m 行 n 列的实矩阵 $\mathbf{A}_{m\times n}$ 就是一个 $R^n \to R^m$ 上的线性变换，或者地说，矩阵 $\mathbf{A}_{m\times n}$ 把一个 $n$ 维空间的 $n$ 维向量变换为一个 $m$ 维空间的 $m$ 维向量。

## 矩阵与向量的乘积的概念

矩阵 **A** 与向量 **c** 的乘积（矩阵左乘向量，记为 **Ac** ）是一个向量，这个向量的每个分量是以矩阵 A 的每个行向量分别与列向量 **c** 的数量积作为元素的。乘式如下：

$$\mathbf{A}\cdot\mathbf{c}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}=\begin{pmatrix} (a_1\ \ a_2\ \ a_3)\cdot\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \\ (b_1\ \ b_2\ \ b_3)\cdot\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \end{pmatrix}=\begin{pmatrix} a_1c_1+a_2c_2+a_3c_3 \\ b_1c_1+b_2c_2+b_3c_3 \end{pmatrix}$$

上式 **Ac** 的乘积是把矩阵 **A** 看作两个行向量，实质上还是向量与向量的点乘积。

类似的，向量 **c** 与矩阵 **A** 的乘积（矩阵右乘向量，记为 **cA** ）也是一个向量，这个 **cA** 向量的每个分量是行向量 **c** 与矩阵 **A** 的每个列向量的点乘积。乘法公式如下：

$$\mathbf{c}\cdot\mathbf{A}=(c_1\ \ c_2\ \ c_3)\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}=\left((c_1\ \ c_2\ \ c_3)\cdot\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\quad(c_1\ \ c_2\ \ c_3)\cdot\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right)=(c_1a_1+c_2a_2+c_3a_3\quad c_1b_1+c_2b_2+c_3b_3)$$

以上的乘积运算都是“行向量·列向量”的形式。下面我们换一下思维方式，把乘积运算看成“列向量·行向量”的形式是否说得通？先从 **Ac** 的乘积开始：

如果把矩阵 A 分解为 3 个列向量的话，我们可以这样展开上式：

$$\mathbf{A}\cdot\mathbf{c}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}=\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}c_1+\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}c_2+\begin{pmatrix} a_3 \\ b_3 \end{pmatrix}c_3=\begin{pmatrix} a_1c_1+a_2c_2+a_3c_3 \\ b_1c_1+b_2c_2+b_3c_3 \end{pmatrix}。$$

上式的含义是 **Ac** 的乘积可以理解为矩阵 **A** 的列向量的线性组合，组合系数是向量 **c** 的三个分量。这个展开的实质仍然是向量点乘积的乘法。再看 $\mathbf{c}\cdot\mathbf{A}$ 的展开：

$$\mathbf{c}\cdot\mathbf{A}=(c_1\ \ c_2\ \ c_3)\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}=c_1(a_1\ \ b_1)+c_2(a_2\ \ b_2)+c_3(a_3\ \ b_3)=(c_1a_1+c_2a_2+c_3a_3\quad c_1b_1+c_2b_2+c_3b_3)$$

**这个式子可以理解为矩阵 A 的行向量的线性组合，组合系数是向量 c 的三个分量。**

实际上，在以上的各种乘法中，我们使用了矩阵和向量的分块技术（全部是“行向量·列向量”的形式），每一个分块都要看成是矩阵最基本的元素“数”进行运算。至此，我们较全面地理解了向量与矩阵乘积的展开实质。显然，左乘与右乘的结果是不同的。为什么不同，答案就在随后的章节里。

# 矩阵与向量乘积的几何意义

为了更具体的观察矩阵和向量乘积的几何意义，我们下面先考察一个矩阵与欧式空间的单位坐标向量的乘积的过程，然后再看一个矩阵与任意向量的乘积的几何意义。

## 矩阵与单位坐标向量的乘积的几何解释

三维的单位坐标向量就是 $\mathbf{i}=(1,0,0)$，$\mathbf{j}=(0,1,0)$，$\mathbf{k}=(0,0,1)$。我们取 $x$ 坐标轴上的单位向量 $\mathbf{i}$ 与 $\mathbf{A}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ 相乘，得乘式如下：

$$\mathbf{iA}=(1,0,0)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(a_1\ \ a_2\ \ a_3),$$

$$\mathbf{Ai}^T=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$$

$\mathbf{iA}$ 的结果是 $(a_1\ \ a_2\ \ a_3)$，这恰是矩阵 $\mathbf{A}$ 的第一行。$\mathbf{Ai}^T$ 的结果是 $\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$，这恰是矩阵 $\mathbf{A}$ 的第一列。

为了更明了，下面把 $\mathbf{j}$ 和 $\mathbf{k}$ 与矩阵 $\mathbf{A}$ 的乘积也一并列出来：

$$\mathbf{jA}=(0,1,0)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(b_1\ \ b_2\ \ b_3)$$

$$\mathbf{kA}=(0,0,1)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(c_1\ \ c_2\ \ c_3)$$

$$\mathbf{Aj}^T=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}=\begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}$$

$$\mathbf{Ak}^T = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$$

从向量对矩阵的作用方面上，我们可以这样理解上述的乘式给出的操作意义上的内涵：

1、 $x$单位坐标向量$\mathbf{i}$左乘一个矩阵就是把矩阵$x$轴行向量（第一行）给取出来；类似的，$y$轴单位坐标向量$\mathbf{j}$左乘一个矩阵就是把矩阵$y$轴行向量（第二行）给取出来；$z$坐标单位向量$\mathbf{k}$左乘一个矩阵就是把矩阵$z$轴行向量（第三行）给取出来；

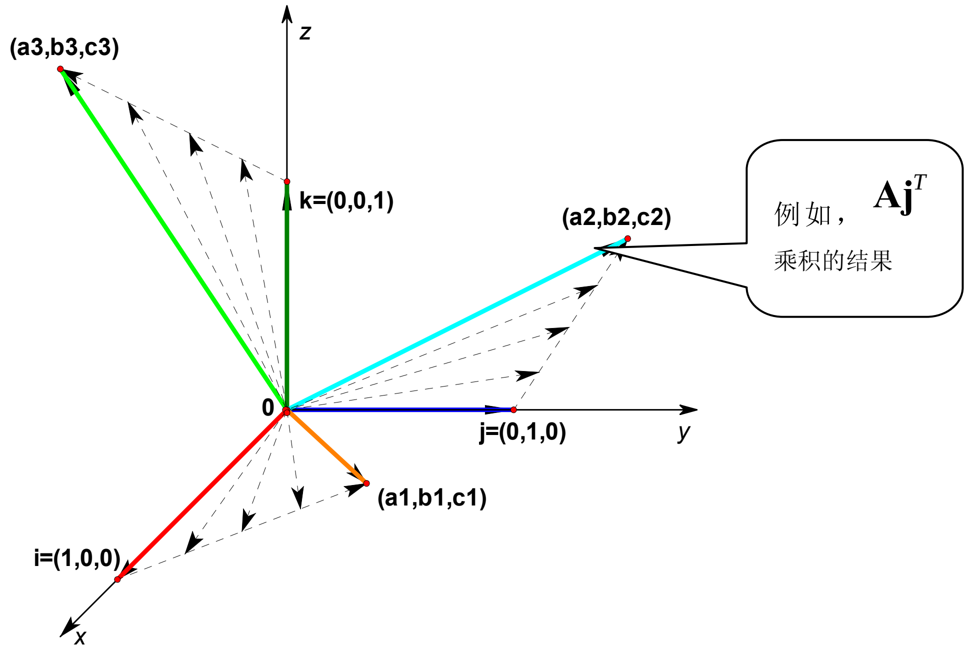
2、 $x$坐标单位向量$\mathbf{i}$右乘一个矩阵就是把矩阵$x$轴列向量（第一列）给取出来；类似的，$y$坐标单位向量$\mathbf{j}$右乘一个矩阵就是把矩阵$y$轴列向量（第二列）给取出来；$z$坐标单位向量$\mathbf{k}$右乘一个矩阵就是把矩阵$z$轴列向量（第三列）给取出来；

另外，从矩阵对向量的作用上，我们又可以从几何图形上这样理解其作意义上的内涵：

3、 一个矩阵右乘$x$坐标单位向量$\mathbf{i}$，就是把向量$\mathbf{i}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$x$轴上的行向量（第一行）；类似的，一个矩阵右乘$y$坐标单位向量$\mathbf{j}$，就是把向量$\mathbf{j}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$y$轴上的行向量（第二行）；一个矩阵右乘$z$坐标单位向量$\mathbf{k}$，就是把向量$\mathbf{k}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$z$上的行向量（第三行）；

4、 一个矩阵左乘$x$坐标单位向量$\mathbf{i}$，就是把向量$\mathbf{i}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$x$轴上的列向量（第一列）；类似的，一个矩阵左乘$y$坐标单位向量$\mathbf{j}$，就是把向量$\mathbf{j}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$y$轴上的列向量（第二列）；一个矩阵左乘$z$坐标单位向量$\mathbf{k}$，就是把向量$\mathbf{k}$的图形缩放旋转变换，变换后的向量就是这个矩阵的$z$轴上的列向量（第三列）；

任何一个矩阵和单位向量$\mathbf{i}$相乘得到矩阵的第一列向量。相类似的，任何一个向量和单位向量$\mathbf{i}$相乘（内积），是这个向量的第一个分量。




上图中，我们给出了三阶矩阵 $\mathbf{A}$ 分别右乘 $x, y, z$ 坐标单位向量 $\mathbf{i}^T$， $\mathbf{j}^T$， $\mathbf{k}^T$ 后的变化几何图形，单位向量 $\mathbf{i}^T$， $\mathbf{j}^T$， $\mathbf{k}^T$ 分别**缩放旋转**变化成了向量 $\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$、向量 $\begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}$ 和向量 $\begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$。虚线向量表示一种变化的过程。

## 矩阵与任意向量的乘积的几何解释

在前面的章节中我们讲过，一个向量可以拆分为单位坐标向量的线性表示，或者讲是单位坐标向量的伸缩变换后的和。我们再次列出这个表达式如下：

$$\mathbf{d} = (d_1, d_2, d_3) = d_1\mathbf{i} + d_2\mathbf{j} + d_3\mathbf{k} = d_1(1,\ 0,\ 0) + d_2(0,\ 1,\ 0) + d_3(0,\ 0,\ 1)$$

或

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = d_1\mathbf{i}^T + d_2\mathbf{j}^T + d_3\mathbf{k}^T = d_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + d_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + d_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

所以，矩阵 $\mathbf{A}$ 对任意向量 $\mathbf{d}$ 的乘积式就是

$$\mathbf{A}\cdot\mathbf{d}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}(d_1\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}+d_2\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}+d_3\begin{pmatrix}0\\ 0\\ 1\end{pmatrix})$$

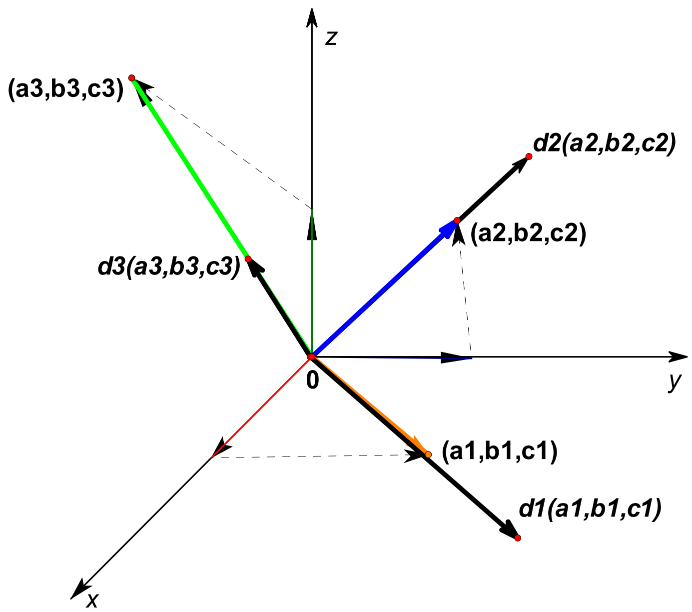
$$=d_1\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}+d_2\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}+d_3\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

$$=d_1\mathbf{A}\cdot\mathbf{i}^T+d_2\mathbf{A}\cdot\mathbf{j}^T+d_3\mathbf{A}\cdot\mathbf{k}^T$$

上式揭示了矩阵 $\mathbf{A}$ 对任意向量 $\mathbf{d}$ 的乘积的几何解释可以分解为几个操作过程：首先，矩阵 $\mathbf{A}$ 分别对单位向量 $\mathbf{i}^T$， $\mathbf{j}^T$， $\mathbf{k}^T$ 进行伸缩旋转变换后得到三个向量（列向量，上节的内容），然后对这三个列向量 $\mathbf{Ai}^T$， $\mathbf{Aj}^T$， $\mathbf{Ak}^T$ 分别进行伸缩变换得到了 $d_1\mathbf{Ai}^T$， $d_2\mathbf{Aj}^T$， $d_3\mathbf{Ak}^T$，再把此三向量相加，得到的和向量就是 $\mathbf{A}\cdot\mathbf{d}$ 的乘积。

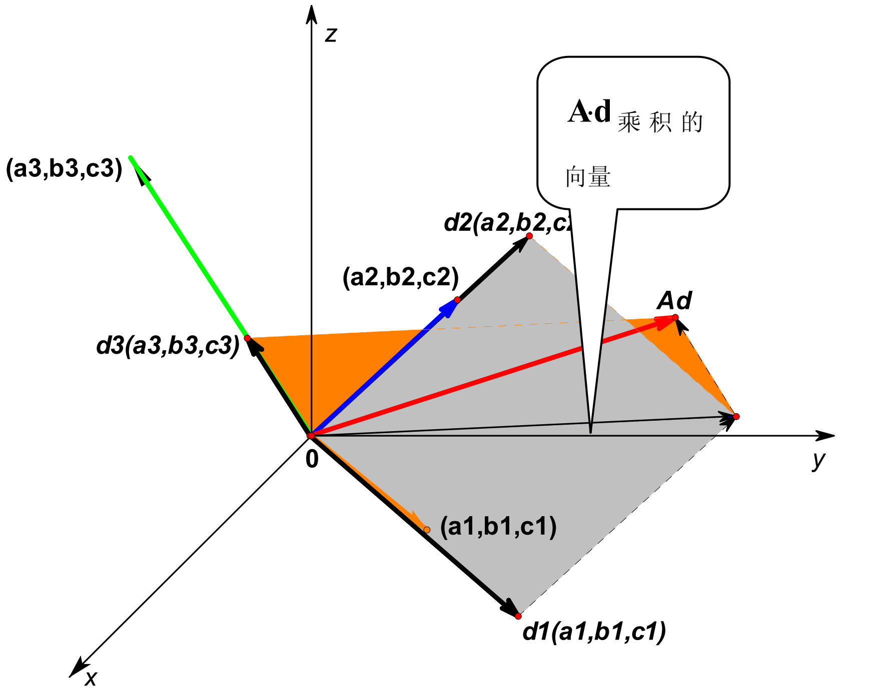
绘出上式的几何图像如下。

下图得到了中间结果三个向量 $d_1\mathbf{Ai}^T=d_1(a_1,b_1,c_1)$， $d_2\mathbf{Aj}^T=d_2(a_2,b_2,c_2)$， $d_3\mathbf{Ak}^T=d_3(a_3,b_3,c_3)$ 的几何图形。




下图把中间结果三个向量 $d_1\mathbf{Ai}^T=d_1(a_1,b_1,c_1)$， $d_2\mathbf{Aj}^T=d_2(a_2,b_2,c_2)$ 和 $d_3\mathbf{Ak}^T=d_3(a_3,b_3,c_3)$，依照平行四边形法则或多边形法则连加，得到最终结果

$\mathbf{A}\cdot\mathbf{d}=(a_1d_1+a_2d_2+a_3d_3, b_1d_1+b_2d_2+b_3d_3, c_1d_1+c_2d_2+c_3d_3)$。




呵，不太直观是吧。好，把上式简化一下，利用线性表示的概念（在矩阵与向量乘积概念一节中讲过）我们可能得到一个更直接的易于理解的几何图像，调整乘式如下：

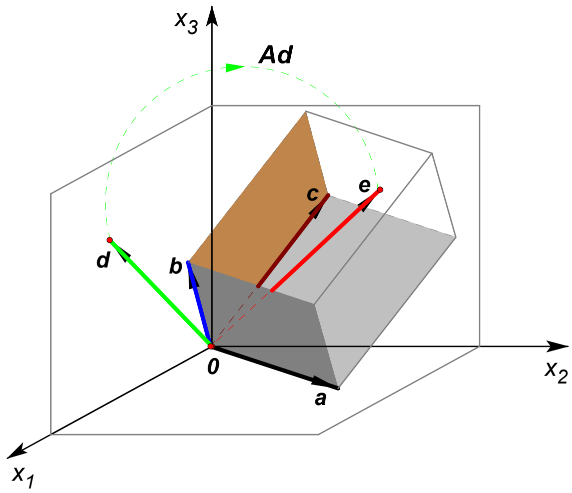
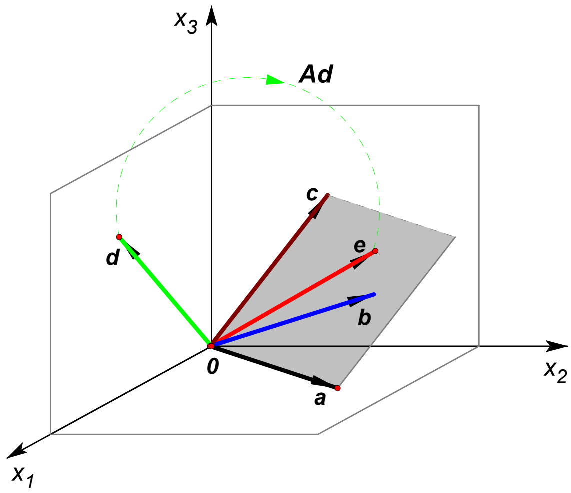
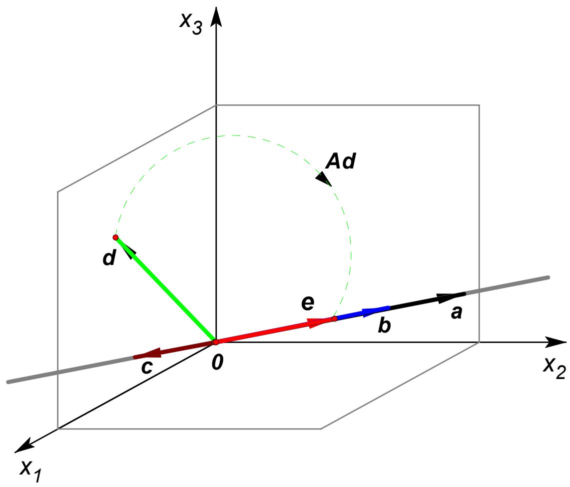
$$\mathbf{A}\cdot\mathbf{d}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}(d_1\begin{pmatrix}1\\0\\0\end{pmatrix}+d_2\begin{pmatrix}0\\1\\0\end{pmatrix}+d_3\begin{pmatrix}0\\0\\1\end{pmatrix})$$

$$=d_1\begin{pmatrix}a_1\\ b_1\\ c_1\end{pmatrix}+d_2\begin{pmatrix}a_2\\ b_2\\ c_2\end{pmatrix}+d_3\begin{pmatrix}a_3\\ b_3\\ c_3\end{pmatrix}$$

上式的几何意义就是向量组的线性表示的几何意义（后面的章节中还要讨论），几何解释就是把矩阵 $\mathbf{A}$ 的列向量进行伸缩变换（比例变换，也可能改变方向）后首尾相连（即向量的和）得到了一个新向量，这个向量就是 $\mathbf{A}\cdot\mathbf{d}$。这个几何解释校前面的解释更简洁一点，但没有大的差别。如果我们知道矩阵的列空间的概念，我们可以得到一个新的几何意义上的理解：

矩阵 $\mathbf{A}$ 的列向量空间是 $\mathbf{A}$ 的所有列向量所张成 $R^n$ 中的一个子空间。尽管这些列向量可能是线性相关的，那么 $\mathbf{A}\cdot\mathbf{d}$ 的表达式（见上式）则说明，矩阵 $\mathbf{A}$ 把 $R^n$ 中的向量 $\mathbf{d}$ 映射到列向量子空间里的一个向量上去了。一个 m 行 n 列的实矩阵 $\mathbf{A}_{m\times n}$ 就是一个 $R^n\to R^m$ 上的线性变换，或者地说，矩阵 $\mathbf{A}_{m\times n}$ 把一个 $n$ 维空间的 $n$ 维向量变换为一个 $m$ 维空间的 $m$ 维向量。

看一个综合性的表格如下：

| 顺序 | 图形解释 | 对应的几何图形 |
|---|---|---|
| 1 | $\mathbf{A}_1\cdot\mathbf{d}_1=\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}$<br>$=d_1\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}+d_2\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}+d_3\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$<br>$=\begin{pmatrix}e_1\\e_2\\e_3\end{pmatrix}$<br>$3\times3$矩阵 **A** 把一个三维向量 **d** 映射到一个三维向量 **e** 。 |  |
| 2 | $\mathbf{A}_2\cdot\mathbf{d}_2=\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}$<br>$=d_1\begin{pmatrix}a_1\\a_2\end{pmatrix}+d_2\begin{pmatrix}b_1\\b_2\end{pmatrix}+d_3\begin{pmatrix}c_1\\c_2\end{pmatrix}$<br>$=\begin{pmatrix}e_1\\e_2\end{pmatrix}$<br>$2\times3$矩阵 **A** 把一个三维向量 **d** 映射到一个二维向量 **e** 。 |  |
| 3 | $\mathbf{A}_3\cdot\mathbf{d}_3=\begin{bmatrix}a_1&b_1&c_1\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}$<br>$=d_1a_1+d_2b_1+d_3c_1$<br>$=e$<br>$1\times3$矩阵 **A** 把一个三维向量 **d** 映射到一个一维向量 **e** （或实数 $e$ ）。 |  |

下节我们讨论一个特殊的矩阵---旋转矩阵。因为通过进一步了解旋转矩阵会深化大家对矩阵几何意义的理解。

## 旋转矩阵对向量的乘积的几何解释

大家已经知道，一个矩阵乘以一个向量，一般将会对向量的几何图形进行旋转和伸缩变化。在教科书中，我们常见的一个例子就是旋转矩阵，旋转矩阵只对向量进行旋转变化而没有伸缩变化。例如二阶旋转矩阵 **A**:

$$\mathbf{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

显然，二阶的所有的某一特定角度的旋转矩阵都分布在单位园上。对$\theta$取几个不同的弧度，就会得到几个旋转矩阵：

| $\theta$ | $0$ | $\frac{\pi}{4}$ | $\frac{\pi}{2}$ | $\frac{3\pi}{4}$ |
|---|---|---|---|---|
| $\mathbf{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ | $\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$ | $\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$ | $\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}$ | $\begin{pmatrix}-\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$ |

比如单位矩阵 $\mathbf{A}_1=\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$ 实际上就是对一个向量 **c** 旋转的角度是 0，也就是 $\mathbf{Ac}=\mathbf{c}$；而矩阵

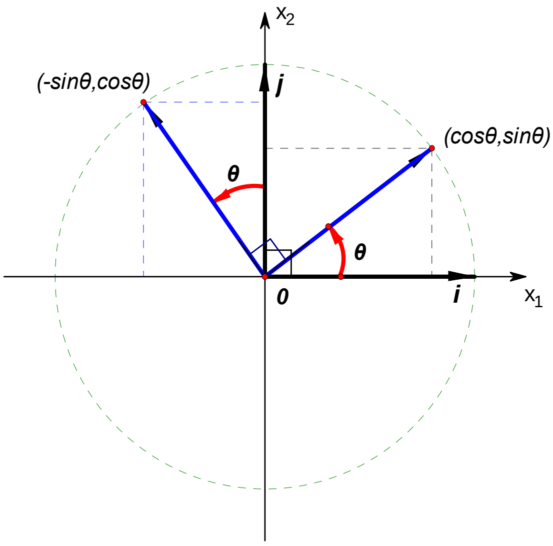
$\mathbf{A}_2=\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$ 就是对一个向量 **c** 旋转的角度是 $\frac{\pi}{4}$。

是不是这样的呢？首先看一下旋转矩阵 **A** 对单位向量 **i, j** 的作用效果。

$$\mathbf{Ai}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},$$

$$\mathbf{Aj}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}=\begin{pmatrix}\cos(\theta+\pi/2)\\ \sin(\theta+\pi/2)\end{pmatrix}$$

从上式和下图看出，旋转矩阵对单位向量 **i, j** 确实分别逆时针旋转了一个$\theta$角度。旋转后的两个向量 **Ai** 和 **Aj** 保持长度不变和夹角不变。或者说向量 $\mathbf{i}\to\mathbf{Ai}$ 长度不变，角度逆时针增加了$\theta$度；向量 $\mathbf{j}\to\mathbf{Aj}$ 长度不变，同时同向角度增加了$\theta$度。显然，向量的和 $\mathbf{i}+\mathbf{j}\to\mathbf{Ai}+\mathbf{Aj}$ 长度不变，角度逆时针增加了$\theta$度。



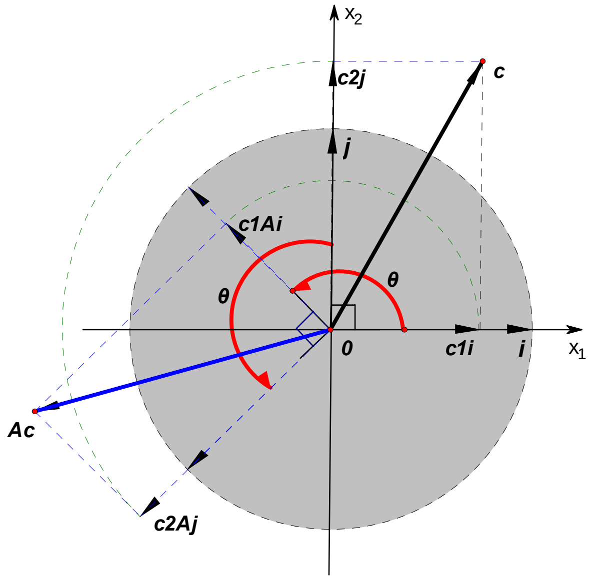

对于任意向量$\mathbf{c}$，我们知道，可以分解为单位向量的线性表示：

$$\mathbf{c}=\begin{pmatrix}c_1\\c_2\end{pmatrix}=c_1\begin{pmatrix}1\\0\end{pmatrix}+c_2\begin{pmatrix}0\\1\end{pmatrix}=c_1\mathbf{i}+c_2\mathbf{j}$$

那么，旋转矩阵作用于向量$\mathbf{c}$的式子为：

$$\mathbf{A}\begin{pmatrix}c_1\\c_2\end{pmatrix}=\mathbf{A}(c_1\begin{pmatrix}1\\0\end{pmatrix}+c_2\begin{pmatrix}0\\1\end{pmatrix})\ =c_1\mathbf{A}\begin{pmatrix}1\\0\end{pmatrix}+c_2\mathbf{A}\begin{pmatrix}0\\1\end{pmatrix}=c_1\mathbf{Ai}+c_2\mathbf{Aj}$$

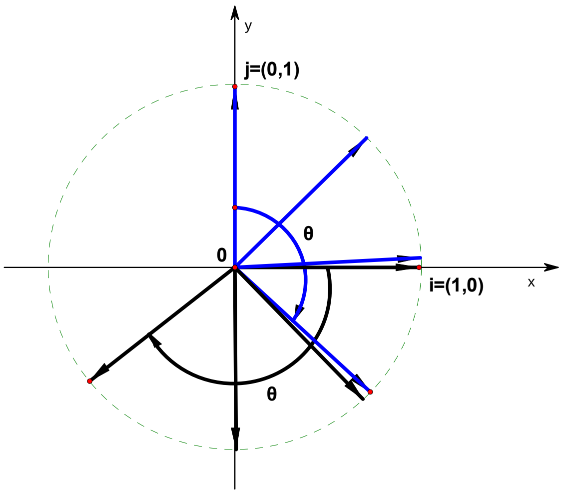
类似的，我们考察旋转矩阵$\mathbf{A}$把任意向量$\mathbf{c}$变换到$\mathbf{Ac}$时，对比上述两式，向量$c_1\mathbf{i}\to c_1\mathbf{Ai}$长度不变，角度逆时针增加了$\theta$度；向量$c_2\mathbf{j}\to c_2\mathbf{Aj}$长度不变，同时同向角度增加了$\theta$度。那么，向量的和$c_1\mathbf{i}+c_2\mathbf{j}\to c_1\mathbf{Ai}+c_2\mathbf{Aj}$长度不变，角度逆时针增加了$\theta$度，即$\mathbf{c}\to\mathbf{Ac}$长度不变，角度逆时针增加了$\theta$度。如下图所示。




另外，对于不同的旋转矩阵之间的比较，也有一个小小的规律。

| $\theta$ | 0 | $\frac{\pi}{4}$ | $\frac{\pi}{2}$ | $\frac{3\pi}{4}$ |
|---|---|---|---|---|
| $A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ | $\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$ | $\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$ | $\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}$ | $\begin{pmatrix}-\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$ |

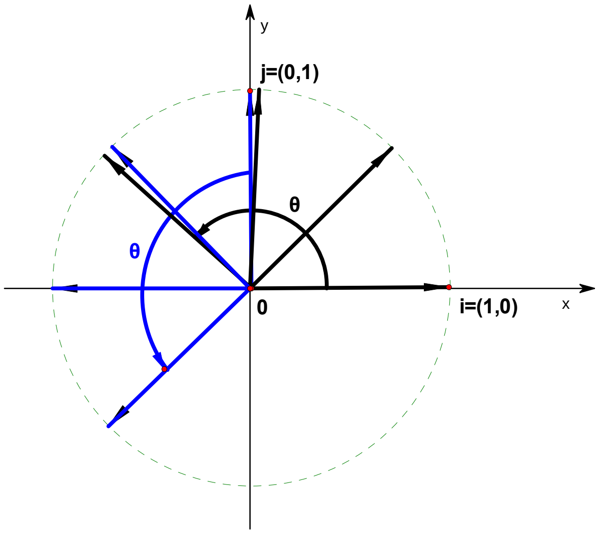
我们给出矩阵几组具体的数值来观察其**行向量**的图形的变化。




由上表格和图形知道，当 θ =0 时，旋转矩阵的行向量就是 x 和 y 轴的单位向量 **i** 和 **j**，随着 θ 角度的增大，单位向量在单位园上同时进行顺时针等角度旋转。

如果我们从列向量的角度看旋转矩阵，又会得到什么图形呢？把上述矩阵的几组具体的数值重新列在下面，来观察其**列向量**的图形的变化。

| $\theta$ | $0$ | $\frac{\pi}{4}$ | $\frac{\pi}{2}$ | $\frac{3\pi}{4}$ |
|---|---|---|---|---|
| $A=\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$ | $\begin{pmatrix}1, & 0 \\ 0, & 1\end{pmatrix}$ | $\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$ | $\begin{pmatrix}0, & -1 \\ 1, & 0\end{pmatrix}$ | $\begin{pmatrix}-\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$ |




显然，从上图看出列向量在随着 θ 的角度的增加而逆时针旋转，刚好与行向量同时反向旋转！这正是正交矩阵的特征。也就是说，旋转矩阵就是正交矩阵。

## 5.4. 矩阵与矩阵的乘法几何意义

两个矩阵相乘如 **AB** 的几何意义可以从多个角度来了解。如果把矩阵 **A** 看作一个几何图形，那么乘以 **B** 就是把 **A** 的图形进行了有规律的变换，这个规律的变换就是线性变换（这里实际上把矩阵 **A** 看成了多个向量的组合）；如果把两个矩阵看作等同的，那么 **AB** 的结果是把两个线性变换进行了叠加或复合（嗨，至于矩阵乘法有什么代数上的意义，你可以先看看 6.9 节“方程组和矩阵、向量组的关系”会得到合理的解释）。下面我们具体的讨论一下。

### 矩阵与矩阵的乘法的意义

矩阵与矩阵的乘法可以从矩阵与向量的乘法得到，因为一个矩阵与多个向量相乘，这多个向量就可以组成一个矩阵（会有些限制）。或者说，矩阵本身就是一个有排列顺序要求的向量组，所以矩阵与矩阵相乘可以看作矩阵乘以列向量（或者行向量乘以矩阵的）的组合。例如：

$$\mathbf{Ac}=\begin{bmatrix}a_1 & a_2 \\ b_1 & b_2\end{bmatrix}\begin{pmatrix}c_1 \\ c_2\end{pmatrix} \quad , \quad \mathbf{Ad}=\begin{bmatrix}a_1 & a_2 \\ b_1 & b_2\end{bmatrix}\begin{pmatrix}d_1 \\ d_2\end{pmatrix}$$

如果把列向量 $\mathbf{c}$ 和 $\mathbf{d}$ 可以组合成一个矩阵 $\mathbf{B}=(\mathbf{c},\mathbf{d})$，那么上述的乘积可以用两个矩阵的乘积来表示：

$$\mathbf{AB} = \mathbf{A}(\mathbf{c},\mathbf{d}) = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \end{bmatrix}$$

当然，如果有更多的向量组合起来，可以形成这样的矩阵乘式：

$$\mathbf{AC} = \mathbf{A}(\mathbf{c},\mathbf{d},\mathbf{e},\mathbf{f},\mathbf{g}) = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} c_1 & d_1 & e_1 & f_1 & g_1 \\ c_2 & d_2 & e_2 & f_2 & g_2 \end{bmatrix}$$

类似的，通过向量乘以矩阵的定义，我们同样可以定义反顺序的矩阵乘式，这里向量为行向量：

$$\mathbf{CA} = \begin{pmatrix} \mathbf{c} \\ \mathbf{d} \\ \mathbf{e} \\ \mathbf{f} \\ \mathbf{g} \end{pmatrix} \mathbf{A} = \begin{bmatrix} c_1 & c_2 \\ d_1 & d_2 \\ e_1 & e_2 \\ f_1 & f_2 \\ g_1 & g_2 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

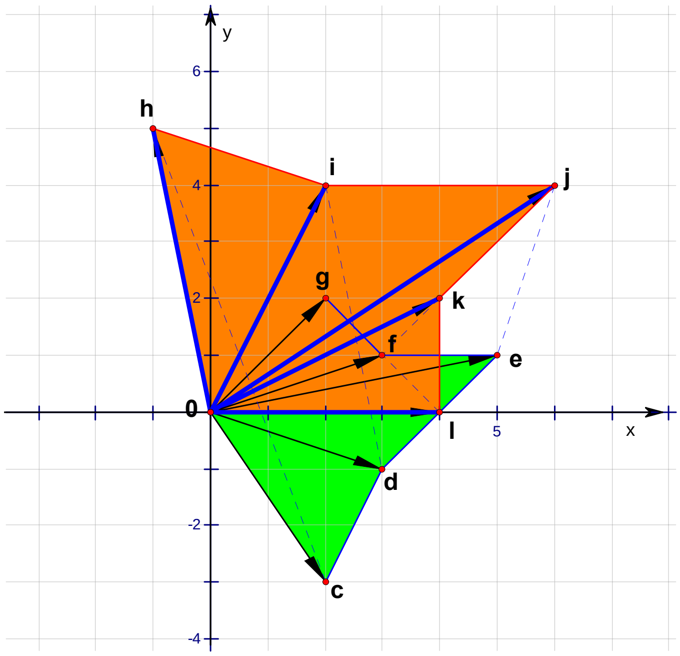
因此，矩阵与矩阵的相乘的几何意义，可以从矩阵与多个向量相乘的几何意义得到，只是多个向量被按照顺序组合成了另一个矩阵。

矩阵乘以矩阵如 $\mathbf{AB} = \mathbf{C}$ 的一般几何意义就是把其中一个矩阵如 $\mathbf{B}$ 的数个行向量或列向量构成的几何图形进行旋转、缩放、镜像等变换（另外一个矩阵 $\mathbf{A}$ 起到的作用）得到了数个新向量，这些新向量作为行向量或者列向量组成一个新的矩阵 $\mathbf{C}$，这个新矩阵 $\mathbf{C}$ 会构成新的几何图形。对于下面的乘式：

$$\mathbf{C} = \mathbf{AB} = \mathbf{A}\begin{pmatrix} \mathbf{c} & \mathbf{d} & \mathbf{e} & \mathbf{f} & \mathbf{g} \end{pmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} c_1 & d_1 & e_1 & f_1 & g_1 \\ c_2 & d_2 & e_2 & f_2 & g_2 \end{bmatrix} = \begin{bmatrix} h_1 & i_1 & j_1 & k_1 & l_1 \\ h_2 & i_2 & j_2 & k_2 & l_2 \end{bmatrix} = \begin{pmatrix} \mathbf{h} & \mathbf{i} & \mathbf{j} & \mathbf{k} & \mathbf{l} \end{pmatrix}$$

我们给出具体的数据例子并画出这个变换的图形：

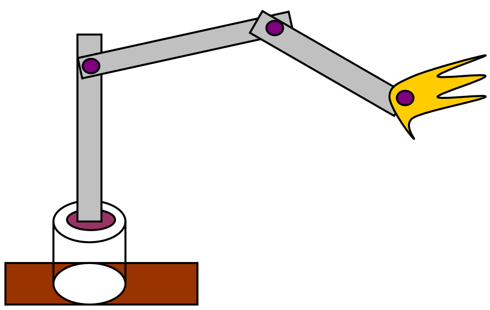
$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 5 & 3 & 2 \\ -3 & -1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 6 & 5 & 4 \\ 5 & 4 & 4 & 1 & 0 \end{bmatrix}$$




一片绿色的小枫叶，经过夏天的风雨洗礼的自然变换，终于成了一片红彤彤的大枫叶。

其实，两个矩阵相乘，我们可以有两种理解。一种理解主要是考察一个矩阵对另一个矩阵所起的变换作用。其作用的矩阵看作是动作矩阵，被作用的矩阵可以看作是由行或列向量构成的几何图形。这个理解就是上面给出的几何解释。

另外一个理解就是两个矩阵都被看作是作用矩阵，两个矩阵的乘积被看作是两个矩阵的和作用。一连串的矩阵相乘如$\mathbf{S}=\mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C}\cdot\mathbf{D}\cdot\mathbf{E}\cdot\mathbf{F}$，会把所有的矩阵线性变化的作用力传递并积累下去，最终得到一个和作用力$\mathbf{S}$。工业上的例子就是机器人的手臂，机械臂上的每个关节就是一个旋转矩阵（比如可以是一个$4\times4$矩阵），机械臂末端的位置或动作是所有关节运动的综合效果。这个综合效果可以用旋转矩阵的乘法得到。在下面的矩阵乘法运算律一节里面，我们将给个几何图形的例子。



## 矩阵左乘与右乘的不同

一个矩阵左乘或右乘另一个矩阵的几何图形变化不会相同，因此，矩阵相乘不能满足交换律。也即

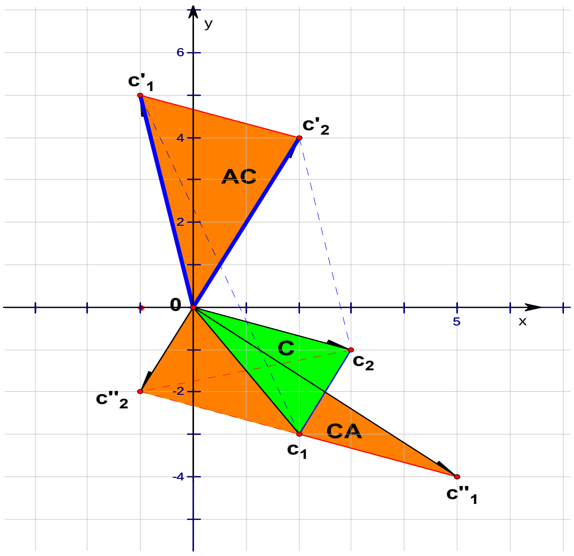
$$\mathbf{AC}\neq\mathbf{CA}$$

下面大家看看矩阵的乘积交换后他们的列向量所构成的三角图形的变化：

$$\mathbf{AC}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}2 & 3\\ -3 & -1\end{bmatrix}=\begin{bmatrix}-1 & 2\\ 5 & 4\end{bmatrix}$$

$$\mathbf{CA}=\begin{bmatrix}2 & 3\\ -3 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}=\begin{bmatrix}5 & -1\\ -4 & -2\end{bmatrix}$$

矩阵 **C** 的图形标示为三角形 **C**，矩阵 **C** 在矩阵 **A** 的左乘和右乘的作用下分别得到了三角形 **AC** 和 **CA** 两个图形，显然，**AC** 和 **CA** 是不同的。图形如下：




：矩阵相乘 $\mathbf{AB}\neq\mathbf{BA}$ 不能满足交换律，是由矩阵相乘的定义决定的，$\mathbf{A}$ 乘以 $\mathbf{B}$ 定义为 $\mathbf{A}$ 的行向量逐个点乘 $\mathbf{B}$ 的列向量； $\mathbf{B}$ 乘以 $\mathbf{A}$ 定义为 $\mathbf{B}$ 的行向量逐个点乘 $\mathbf{A}$ 的列向量；一个矩阵的行向量与其列向量是不同的，因此 $\mathbf{AB}\neq\mathbf{BA}$。

一般情况下，矩阵乘法不能交换，但有两个例外要注意：一是单位矩阵可以交换，即 $\mathbf{IA}=\mathbf{AI}=\mathbf{A}$；另一个是逆矩阵可以交换，即 $\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$。可以交换的主要原因之一是单位矩阵的行向量等于列向量。

下面我们给个旋转矩阵相乘积的例子。

## 矩阵乘幂的几何及物理解释

矩阵乘幂运算就是对一个矩阵 **A** 进行一定次数的联乘运算 $\mathbf{A}^n$。有实际物理意义的可乘幂运算的矩阵常见的有：Markov(马尔科夫)链中的转移概率矩阵、旋转矩阵和图的邻接矩阵等。在天气的马尔科夫链中，天气转移概率矩阵的 $n$ 次幂 $\mathbf{A}^n$ 表示第 $n$ 天后天气是晴、阴或雨的概率；逆时针旋转 $\theta$ 角的矩阵的 $n$ 次幂 $\mathbf{A}^n$ 表示对一个对象旋转了 $n$ 个 $\theta$ 角度。邻接矩阵可以表明一个电路网络、电话网络、道路网络以及机构

网络的连接关系，而邻接矩阵的乘幂可以得到任意两个网络节点之间的连接长度及是否有连接的情况。比如邻接矩阵的 2 次幂 $\mathbf{A}^2$ 的结果告诉我们任意两个节点之间是否存在长度为 2 的道路，而且还告诉了我们每一对联结长度为 2 的道路有多少。

对于三个矩阵乘幂的例子中，下面我们较详细的给出了旋转矩阵乘幂的几何解释。

前面讲过，二阶逆时针旋转矩阵表示为

$$\mathbf{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

那么，两个及至 n 个旋转矩阵的乘积表达式可以由三角公式得到：

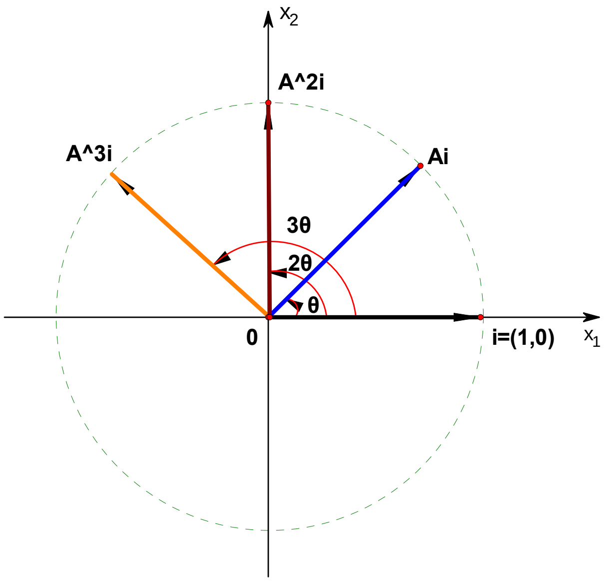
$$\mathbf{A}^2=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{pmatrix}$$

$$\mathbf{A}^3=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{pmatrix}=\begin{pmatrix}\cos 3\theta & -\sin 3\theta\\ \sin 3\theta & \cos 3\theta\end{pmatrix}$$

......................

$$\mathbf{A}^n=\begin{pmatrix}\cos n\theta & -\sin n\theta\\ \sin n\theta & \cos n\theta\end{pmatrix}$$

由上式看到，旋转矩阵的连乘积等效于旋转角的连加和。下图给出了旋转矩阵 $\mathbf{A},\mathbf{A}^2,\mathbf{A}^3$ 对单位向量 $\mathbf{i}$ 的旋转效果。矩阵的乘法变成了角度的加法。



# 5.5. 矩阵与线性变换的关系的几何意义

## 矩阵与线性变换的关系

从前面矩阵的乘法可以知道，任意一个矩阵其本身蕴含着一个变换。这个变换我们可以称为一个矩阵变换。前面章节里我们也讲过线性变换的概念：数域 **F** 上线性空间 **V** 中的变换 **T** 称为 **V** 中的**线性变换，**那么当且仅当满足条件：

$$T(\boldsymbol{\alpha}+\boldsymbol{\beta})=T(\boldsymbol{\alpha})+T(\boldsymbol{\beta}) \quad (\boldsymbol{\alpha},\boldsymbol{\beta}\in V)$$
$$T(k\boldsymbol{\alpha})=kT(\boldsymbol{\alpha}) \quad (k\in F,\boldsymbol{\alpha}\in V)$$

这里，我们把变换 **T** 的符号替换为矩阵 **A**，线性变换就变成了矩阵变换。

实际上，从 $R^n \to R^m$ 上的线性变换都可以表述为一个矩阵变换；反过来，**一个矩阵变换也必然是一个线性变换。两者具有一一对应的关系**。这个对应关系笼统地表述如下：

- 线性变换的和对应着矩阵的和；
- 线性变换的乘积对应着矩阵的乘积；
- 线性变换的数量乘积对应着矩阵的数量乘积；
- 线性变换的逆对应着矩阵的逆；

因此，对于一个变换我们需要弄懂两个问题：这个变换是否线性变换？如果是线性变换如何求对应的矩阵？

判断是否线性变换只要确定这个变换是否满足加法和数乘法则就可以了。线性变换一般可以用文字描述出来，如“一个旋转角度为 $\pi/8$ 的旋转变换”，或者“一个关于 $x_2=-x_1$ 的镜像变换”等，但如何知道旋转变换的所对应的矩阵呢？这个问题有幸有一个定理给出了回答。作为一个例子，下面的定理给出了如何把一个 $R^2$ 空间上的线性变换转换成一个对应的 2 阶矩阵的办法：

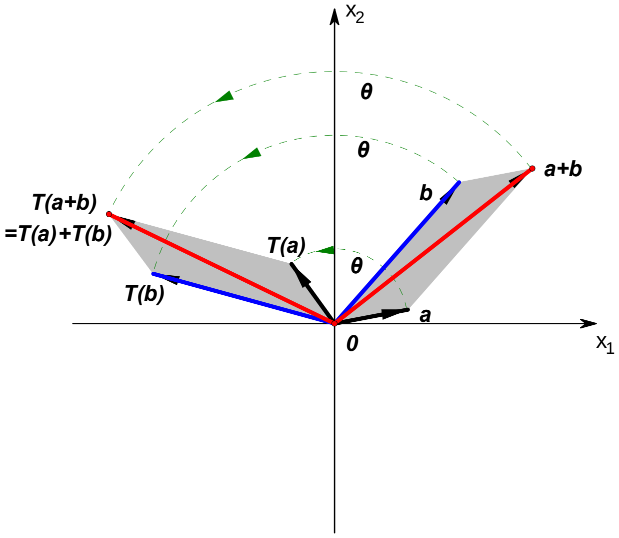
**定理：设 $T:R^2\to R^2$ 是任意一个线性变换，那么 $T$ 的矩阵的列向量为 $T(\mathbf{e}_1)$ 和 $T(\mathbf{e}_2)$。**

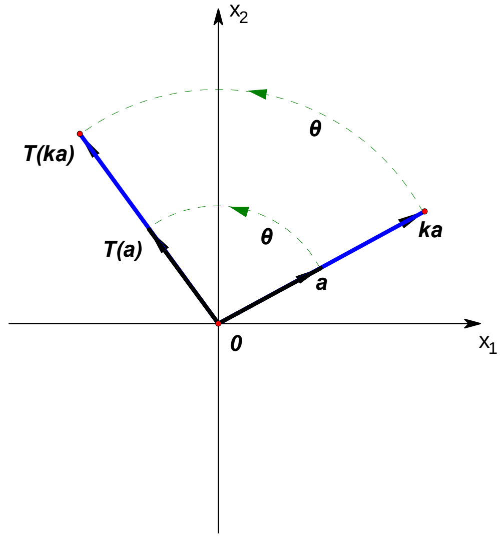
OK，两个问题得到了解答。那我们就验证两个例子加深一下印象。

我们先看一个在平面上的旋转变换的例子（呵，这个例子举了 N 多次了）。

设变换 $T:R^2\to R^2$ 是 $R^2$ 中把每个向量按逆时针方向绕原点 0 转动 $\theta$ 角的旋转变换。这里我们要弄清两个问题，一是旋转变换是否是线性变换？二是旋转变换对应的矩阵如何求（注：这个矩阵就是前面多次介绍的平面旋转矩阵）？

判断是否线性变换就是确认这个变换是否满足加法和数乘。我们看下图所示。



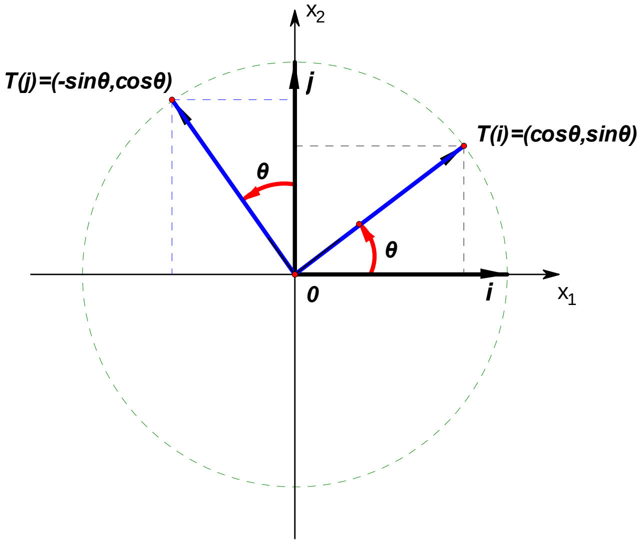




左图中向量$\mathbf{a}$和$\mathbf{b}$同时逆时针旋转了一个$\theta$角，因此两个向量之间的夹角没变，长度也没变，显然，构成的平行四边形的外形没变；或者说，平行四边形也同时逆时针旋转了同样的角度。因此旋转前后的平行四边形的对角线长度没变同时也旋转了同样的$\theta$角，换句话说，旋转前的平行四边形对角线向量$\mathbf{a}+\mathbf{b}$经过旋转变换$T()$后等于旋转后平行四边形的对角线$T(\mathbf{a})+T(\mathbf{b})$，亦即$T(\mathbf{a}+\mathbf{b})=T(\mathbf{a})+T(\mathbf{b})$。这说明满足加法原则。

右图是满足数乘$T(k\mathbf{a})=kT(\mathbf{a})$的图示。为了不侮辱读者的智慧，解释就免了。

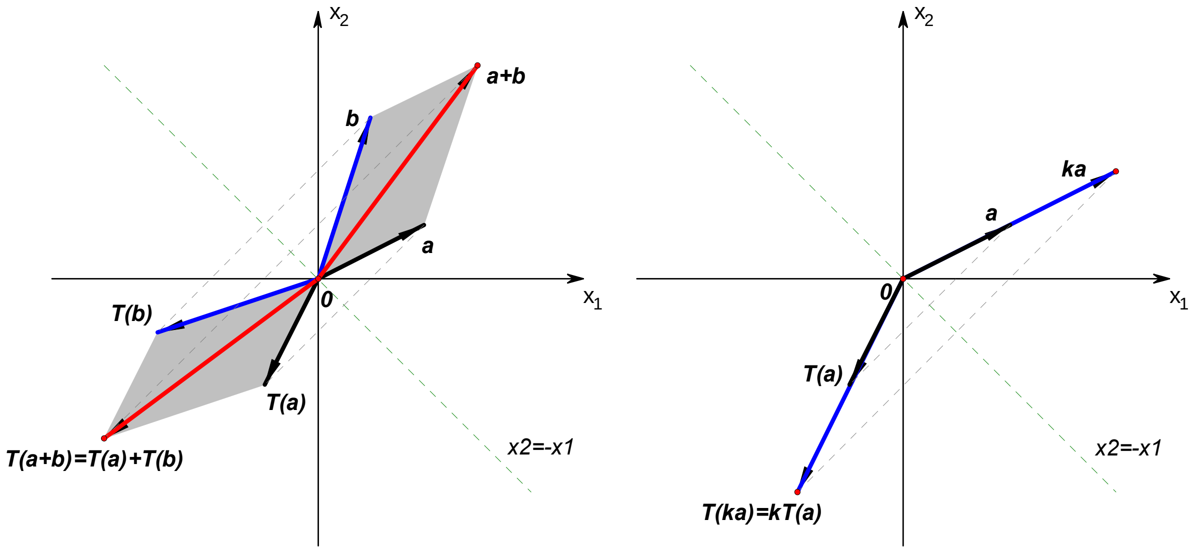
现在我们已经知道了平面空间中的逆时针旋转矩阵的通用矩阵是$\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$，如何利用定理得到这个旋转矩阵的？实际上我们在前面的多处章节里比如“标准正交基的几何解释”和“旋转矩阵对向量的乘积的几何解释”中已经给出了回答。这里只是套一下定理，给与重新确认：

把坐标向量$\mathbf{e}_1$（有时写作$\mathbf{i}$）逆时针旋转$\theta$角度，即$T(\mathbf{e}_1)=T(\mathbf{i})=\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$；同理，把坐标向量$\mathbf{e}_2$（有时写作$\mathbf{j}$）逆时针旋转$\theta$角度，即有$T(\mathbf{e}_2)=T(\mathbf{j})=\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}$。这两个新向量刚好就是旋转矩阵的两个列向量，把他们按顺序组合起来就是旋转矩阵$\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$。




因为这个求矩阵的定理很实用，我们再举一个镜像变换的例子：

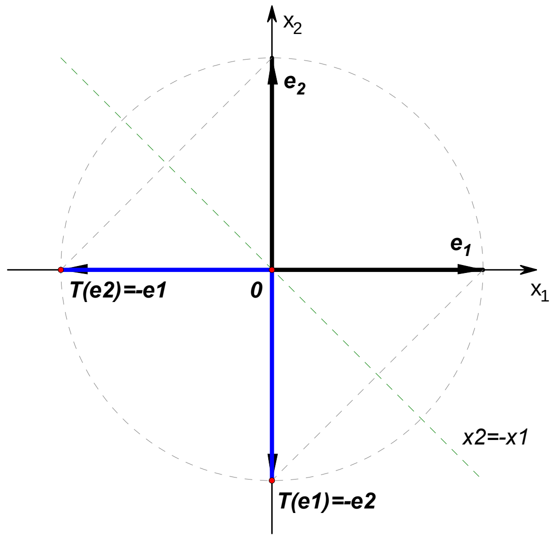
设$T:R^2\to R^2$是把向量空间$R^2$中把每个向量对于直线$x_2=-x_1$作反射的变换。容易看出，这个变换是保持了向量的加法和数乘的原则（如下图，我们看到向量$\mathbf{a}$和$\mathbf{b}$作向量加法的平行四边形在镜像变换$T()$下对于直线$x_2=-x_1$的反射）。




因此镜像变换是一个线性变换。

那么对应镜像变换的矩阵的第一个列向量是$T(\mathbf{e}_1)=T(\mathbf{i})=-\mathbf{j}=\begin{pmatrix}0\\-1\end{pmatrix}$,第二个列向量是

$T(\mathbf{e}_2)=T(\mathbf{j})=-\mathbf{i}=\begin{pmatrix}-1\\0\end{pmatrix}$（见下图）。




所以其镜像矩阵为$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$。

所以啊，由于线性变换与矩阵之间有一一对应（在给定基的前提下），而且保持线性运算关系不变（线性变换的加法和数乘分别对应于在某一个基下的矩阵的加法和数乘），因此，可以用矩阵来研究线性变换，也可以用线性变换来研究矩阵。

常见的线性变换有初等变换、等价变换、相似变换、合同变换等。我们也常常听到正交变换的名字，但由于正交变换包括平移、旋转和镜像，我们知道平移变换不是线性变换，因此不是所有的正交变换是线性变换。

下面或以后的章节我们将讨论对应着各种线性变换的各种类型矩阵的几何意义。首先看看对应着初等变换的称为初等矩阵的几何意义。

## 基本初等矩阵/初等变换的几何意义

线性变换与矩阵之间有一一对应关系，因此就有初等矩阵所对应的所谓初等变换。一组向量的线性相关性必然会体现在这组向量的数组之间的关系上，因此我们定能通过向量之间的线性运算把这个关系揭示出来，由于线性运算是在同一分量上进行，人们把一组向量并列在一起，便于对同一分量进行运算，这就是矩阵的初等变换。由此看来，初等变换只对表示向量组的矩阵才能看到其几何意义。在数学中，数域 R 上的矩阵的基本初等变换有 3 种：

（1）交换某两行的位置；

（2）把某一行乘以一个非零数$k(k \in R)$；

（3）把某一行的$k(k \in R)$倍加到另一行上。

对于数域$R$上的一个单位矩阵分别实施上述3种基本初等变换，所得矩阵分别称为基本初等矩阵(1)、(2)、(3)。基本初等矩阵与矩阵的基本初等变换也是一一对应的；任何一个可逆矩阵都可以分解成基本初等矩阵（1)、(2)、(3）的乘积。同时基本初等矩阵（1）可以由（2）和（3）导出。所以，任何一个可逆矩阵都可以分解成（2）和（3）这两种基本初等矩阵的乘积。从变换的角度来说，一个可逆的线性

变换是连续实施若干次（2）和（3）两种基本初等变换的结果。

：注意：初等变换（2）就是线性变换的数乘变换；初等变换（3）就是线性变换的数乘后的加法，或者讲是线性组合。他们之所以基础就是因为他们是线性变换的数乘和加法的法则。

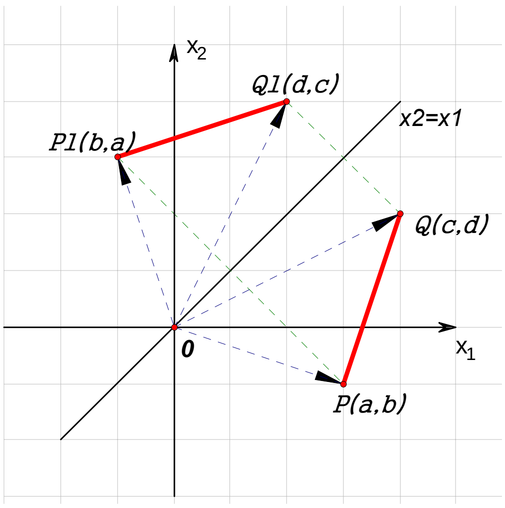
一个二阶矩阵作用在一个二维向量上得到一个新的向量。因此，一个二阶矩阵把平面上的每一个点都变成唯一的点，从而它是平面到平面的映射（即变换），可以刻画平面上的几何变换。同理，任意一个三阶矩阵可以刻画空间中的几何变换。由于任何一个可逆矩阵都可以分解成基本初等矩阵（1）、（2）、（3）的乘积，因此，基本初等矩阵（1）、（2）、（3）的几何意义是理解一般矩阵所表示的变换的几何意义的基础。从几何的观点理解矩阵，把矩阵视为一种几何变换，赋予矩阵一种直观意义。

基本初等矩阵（1）的几何意义是：关于某一“标准轴（面）”的镜像反射（对称）变换；基本初等矩阵（2）的几何意义是：在某一坐标轴方向的伸缩变换；基本初等矩阵（3）的几何意义是：在某一坐标轴方向的切变变换。

- **基本初等矩阵（1）的几何意义**

在二维和三维几何空间中，基本初等矩阵（1）所表示的几何变换是：对图形实施关于某一“标准轴（面）”的镜像反射（对称）变换。其中，标准轴是指在二维几何平面上的直线 $x_2 = x_1$；标准面是指在 3 维几何空间中的平面 $x_2 = x_1, x_2 = x_3, x_3 = x_1$。

例如：设 $P(a,b)$ 是二维几何平面 $x_2 0 x_1$ 上的任意一点（等同于任意向量），点 $P$ 经基本初等矩阵 $T(1,2)=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 变换后的结果为 $P_1 = T(1,2)P = (b,a)$（如下图），

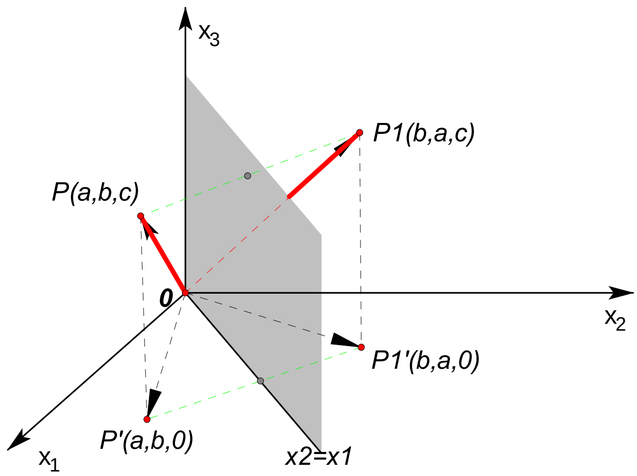



$P_1$ 是点 $P$ 关于直线 $x_2 = x_1$ 的对称点。同样，$Q_1$ 是点 $Q$ 关于直线 $x_2 = x_1$ 的对称点，由此可类推知线

段 $P_1Q_1$ 是 $PQ$ 关于直线 $x_2 = x_1$ 的对称图形。

设 $P(a,b,c)$ 是 3 维几何空间 $0x_1x_2x_3$ 中任意一点，点 $P$ 经基本初等矩阵 $T(1,2),T(2,3),T(1,3)$ 分别变换后，所得结果为：$P_1 = T(1,2)P = (b,a,c)$，$P_2 = T(2,3)P = (a,c,b)$，$P_3 = T(1,3)P = (c,b,a)$。点 $P_1$ 是点 $P$ 关于平面 $x_2 = x_1$ 的对称点、点 $P_2$ 是点 $P$ 关于平面 $x_2 = x_3$ 的对称点、点 $P_3$ 是点 $P$ 关于平面 $x_3 = x_1$ 的对称点。

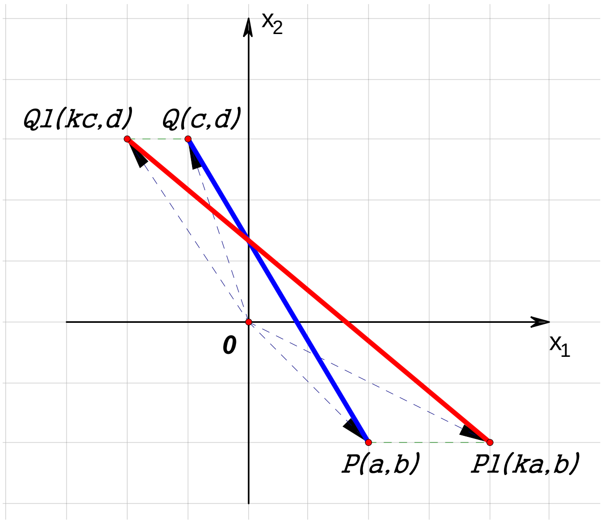
下图绘出了点 $P_1$ 是点 $P$ 关于平面 $x_2 = x_1$ 的对称点。




这个结论可以从二维的例子演推到，点 $P'(a,b,0)$ 和 $P_1'(b,a,0)$ 在二维平面 $x_2 0 x_1$ 上关于直线 $x_2 = x_1$ 镜像对称；扩展到三维空间后，就得到了 $P(a,b,c)$ 和 $P_1(b,a,c)$ 关于平面 $x_2 = x_1$ 镜像对称。

- **基本初等矩阵（2）的几何意义**

在二维和三维几何空间中，基本初等矩阵（2）所表示的几何变换是：对图形实施的在某一坐标轴方向的伸缩变换。例如：设 $P(a,b)$ 是平面 $x_1 0 x_2$ 上的任意一点，点 $P$ 经基本初等矩阵 $T(1(k)) = \begin{pmatrix} k & o \\ 0 & 1 \end{pmatrix}$ 和 $T(2(k)) = \begin{pmatrix} 1 & o \\ 0 & k \end{pmatrix}$ 分别实施变换后结果为：$P_1 = T(1(k))P = (ka,b)$ 和 $P_2 = T(2(k))P = (a,kb)$。点 $P_1,P_2$ 分别是把点 $P$ 在 $x_1$ 轴方向的坐标伸缩 $k$ 倍在 $x_2$ 轴方向的坐标不变，在 $x_2$ 轴方向的坐标伸缩 $k$ 倍在 $x_1$ 轴方向的坐标不变而得到的。




上图给出了两点 $P(a,b), Q(c,d)$ 在 $T(1(2))$ 的作用下沿 $x_1$ 轴伸缩的情况。同时，线段 $PQ$ 经在 $T(1(2))$ 的作用下变化为 $P_1Q_1$。

设 $P(a,b,c)$ 是三维几何空间 $0x_1x_2x_3$ 中任意一点，点 $P$ 经三阶基本初等矩阵 $T(1(k))$，$T(2(k))$，$T(3(k))$（其中 $k \neq 0$）实施变换后，所得结果分别为：$P_1 = T(1(k))P = (ka,b,c)$，$P_2 = T(2(k))P = (a,kb,c)$，$P_3 = T(3(k))P = (a,b,kc)$。$P_1$ 是把点 $P$ 在 $x_1$ 轴方向的坐标伸缩 $k$ 倍，而在 $x_2, x_3$ 轴方向的坐标不变而得到的；$P_2$ 是把点 $P$ 在 $x_2$ 轴方向的坐标伸缩 $k$ 倍，而在 $x_1, x_3$ 轴方向的坐标不变而得到的；$P_3$ 是把点 $P$ 在 $x_3$ 轴方向的坐标伸缩 $k$ 倍，而在 $x_1, x_2$ 轴方向的坐标不变而得到的。若 $k > 0$，其伸缩方向与原坐标方向相同；若 $k < 0$，其伸缩方向与原坐标方向相反。作为特例，当 $k = 0$ 时，基本初等矩阵（2）表示的几何变换是：对图形实施的在某一坐标轴（面）上的投影变换；当 $k = -1$ 时，基本初等矩阵（2）表示的几何变换是：对图形实施的关于某一坐标轴（面）的镜像反射（对称）变换。

例如：矩阵 $T(1(0)), = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ 和 $T(1(-1)), = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$，对点 $P(a,b)$ 实施变换后分别为：

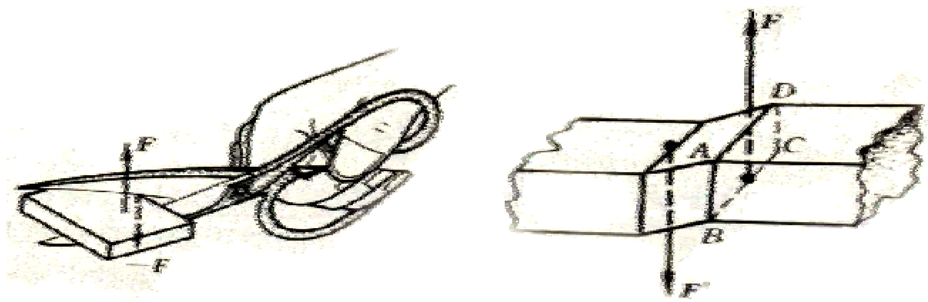
$P_1 = T(1(0))P = (0,b)$，是点 $P(a,b)$ 在 $x_2$ 轴上的投影点；$P_2 = T(1(-1))P = (-a,b)$ 是点 $P(a,b)$ 关于 $x_2$ 轴反射点。

3 阶矩阵 $T(1(0)), T(1(-1))$ 对点 $P(a,b,c)$ 实施变换后分别为：$P_1 = T(1(0))P = (0,b,c)$，是点 $P(a,b,c)$ 在 $x_2 0 x_3$ 平面上的投影点；$P_2 = T(1(-1))P = (-a,b,c)$ 是点 $P(a,b,c)$ 关于坐标平面 $x_2 0 x_3$ 的反射点。

- **基本初等矩阵（3）的几何意义**

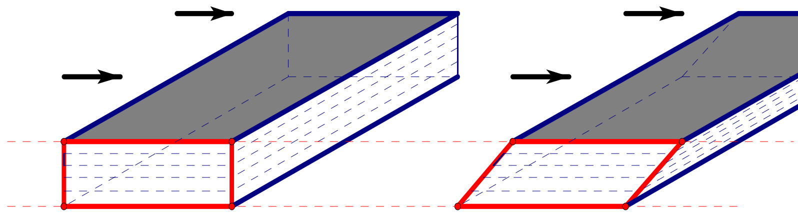
在二维和三维几何空间中，基本初等矩阵（3）所表示的几何变换是：对图形实施的在某一坐标轴方向的切变变换。切变在物理学中属于物体最基本的形变之一，在高等代数和解析几何中很少使用切变这一概念，因此，对于基本初等矩阵（3）的几何意义需要结合物理中的切变来加以说明。

切变，在物理学中指由剪应力作用造成的扭曲（或变形）。如下图 1，当物体受到力偶作用使物体两个平行截面间发生相对平行移动时，在弹性力学中把这种形变叫做剪切形变，简称切变。切变的主要特征为：（1）平行截面间相对滑移；（2）只有纯粹的形状变化，而没有大小（面积或体积）的变化。




剪切形变图

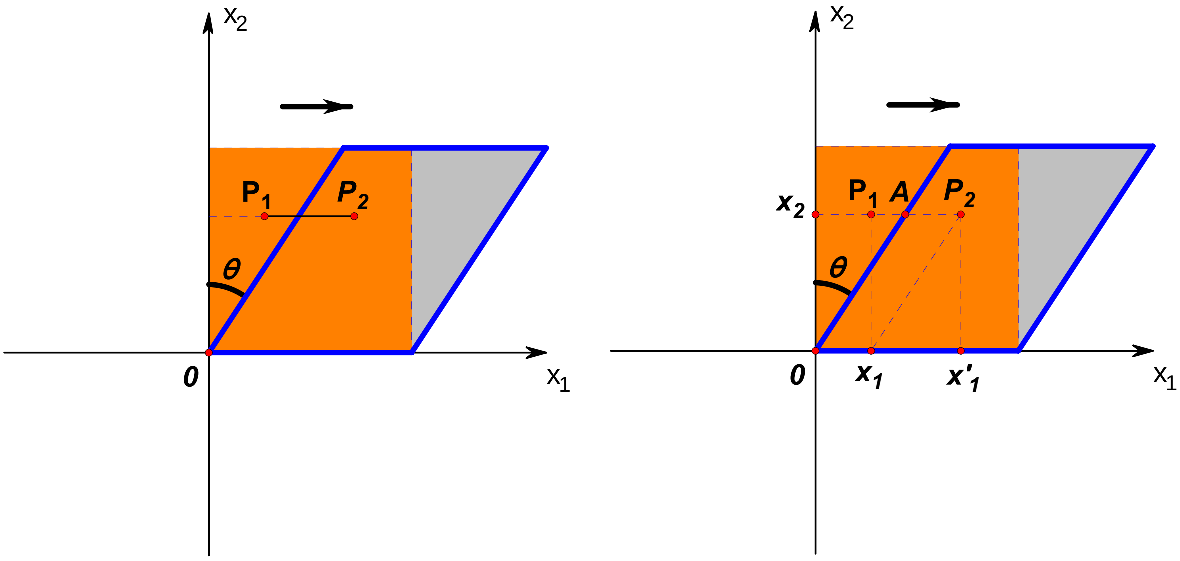
切变的事例在日常生活中有很多。如下图，将一本厚书放在桌面上，推动它的封面，使书页发生滑动，这时在书页的两头边缘上画出的一个矩形变成了平行四边形，这本书受到的就是切变。其形状发生了变化，而体积不变（这本书的宽度和厚度均没有发生变化）。



厚书切变图

二维几何空间中的切变变换矩阵

如下图左，在平面上，一个正方形受到了一个 $x_1$ 轴方向的水平切变力，切变表现为，其力偶作用由正方形变成一个等底、等高的平行四边形。其中，点 $P_1(x_1,x_2)$ 为正方形的任意一点，点 $P_2(x_1^{'},x_2^{'})$ 是切变后 $P_1$ 点所对应的平移点。




由切变的意义，在从点 $P_1$ 到点 $P_2$ 的切变过程中，纵坐标 $x_2$（高度）不会发生变化，即 $x_2^{'}=x_2$；横坐标 $x_1$ 会向右（力的方向）平移一段距离 $P_1P_2$。如果把切变的变化量用一个变化的夹角来度量，就可以得到 $P_1P_2=x_2 tg\theta$（见上图右，因为两个三角形全等：$\Delta 0Ax_2 \cong \Delta x_1P_2P_1$），因此横坐标的变化关系为 $x_1^{'}=x_1+x_2 tg\theta$。把 $tg\theta$ 简写为一个变量 $k$ $(k\in R)$，那么有：

$$\begin{cases} x_1^{'}=x_1+kx_2 \\ x_2^{'}=x_2 \end{cases},$$

改写成矩阵/向量的表达式为：

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix}=\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}。$$

类似的，如果正方形切变的方向是竖直 $x_2$ 轴的方向，那么变化前后的坐标表达式：

$$\begin{cases} x_1^{'}=x_1 \\ x_2^{'}=kx_1+x_2 \end{cases},$$

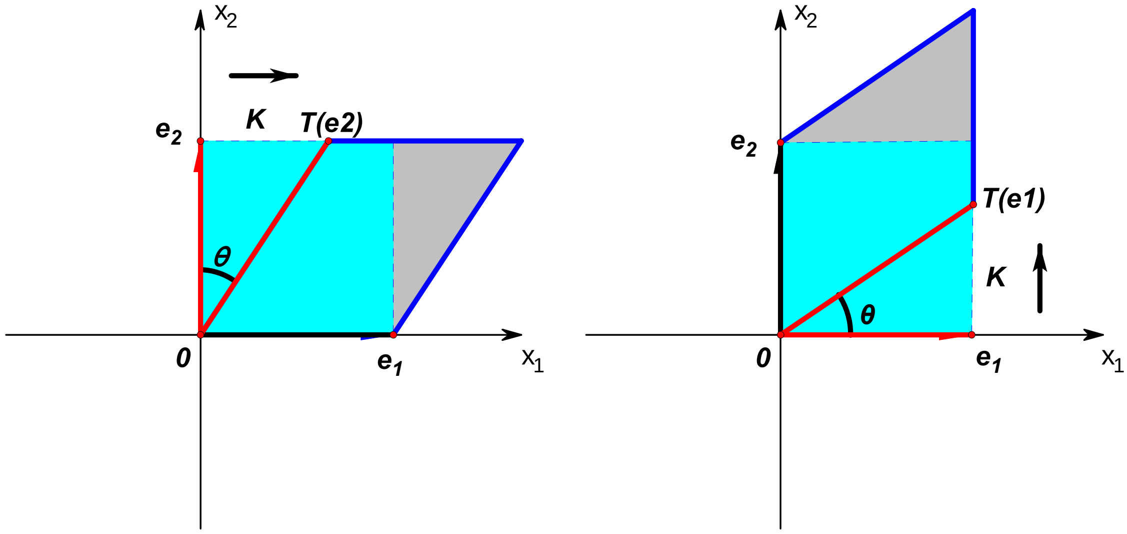
改写成矩阵/向量的表达式为：

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix}=\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}。$$

可见，矩阵 $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ 和 $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ 是刻画平面上切变的数学模型，我们可以称其为平面上的切变变换矩阵。

它们是把二阶单位矩阵的某一行的 k 倍加到另一行的基本初等矩阵$T(i, j(k))$。

实际上，如果我们利用上节刚刚知道的定理来快速推算这个基本的初等矩阵，可能让你更清晰简单。即使你忘掉了也可以随时推导出来。下面我们试一下：




先看上图左，设切变前的正方形为单位长度的，其两个边分别是单位向量$\mathbf{e}_1, \mathbf{e}_2$。那么切变变换后，

单位向量$\mathbf{e}_1, \mathbf{e}_2$分别变换为新向量$T(\mathbf{e}_1)$和$T(\mathbf{e}_2)$。由图易看出 ：向量$\mathbf{e}_1$没有变化，所以$T(\mathbf{e}_1)=\mathbf{e}_1=\begin{pmatrix}1\\0\end{pmatrix}$；

向量$\mathbf{e}_2$的高度没变化，但在$x_1$方向移动了距离$k$，那么有$T(\mathbf{e}_2)=\begin{pmatrix}k\\1\end{pmatrix}$。因此有$x_1$方向的切变矩阵

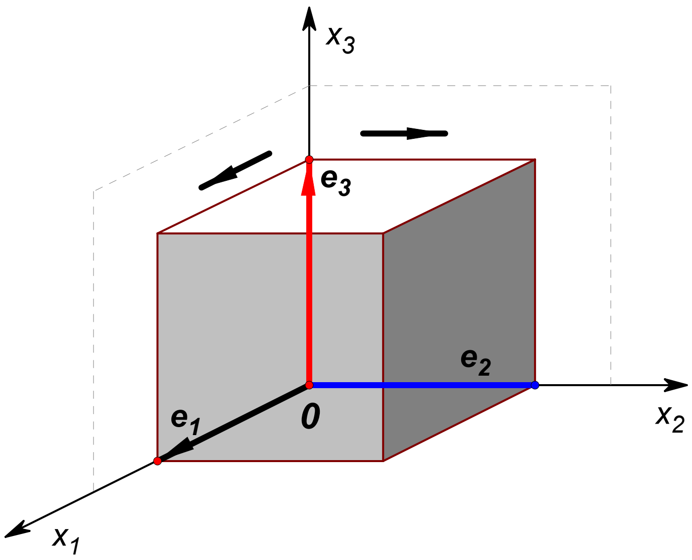
$$\left(T(\mathbf{e}_1), T(\mathbf{e}_2)\right)=\begin{pmatrix}1 & k\\0 & 1\end{pmatrix}。$$

类似的，由上图右可得到$x_2$方向的切变矩阵$\left(T(\mathbf{e}_1), T(\mathbf{e}_2)\right)=\begin{pmatrix}1 & 0\\k & 1\end{pmatrix}$。

三维几何空间中的切变变换矩阵

在 3 维几何空间中，一个正方体的切变表现为，其受到力偶作用由正方体变成一个等底、等高的平行六面体。

在下面的例子中，我们假设正方体受到了平行于$x_1$或者$x_2$轴方向的切变作用力两种情况（见下图）。




当受到平行于 $x_1$ 轴方向的切变作用力时（如下图左），正方体内部的任意一点的坐标中 $x_2$ 和 $x_3$ 不会变，只有 $x_1$ 坐标会发生变化。$x_1$ 坐标的变化情况与二维平面上的切变类同。

当受到平行于 $x_2$ 轴方向的切变作用力时（如下图右），正方体内部的任意一点的坐标中 $x_1$ 和 $x_3$ 不会变，只有 $x_2$ 坐标会发生变化。$x_2$ 坐标的变化情况与二维平面上的切变类同。

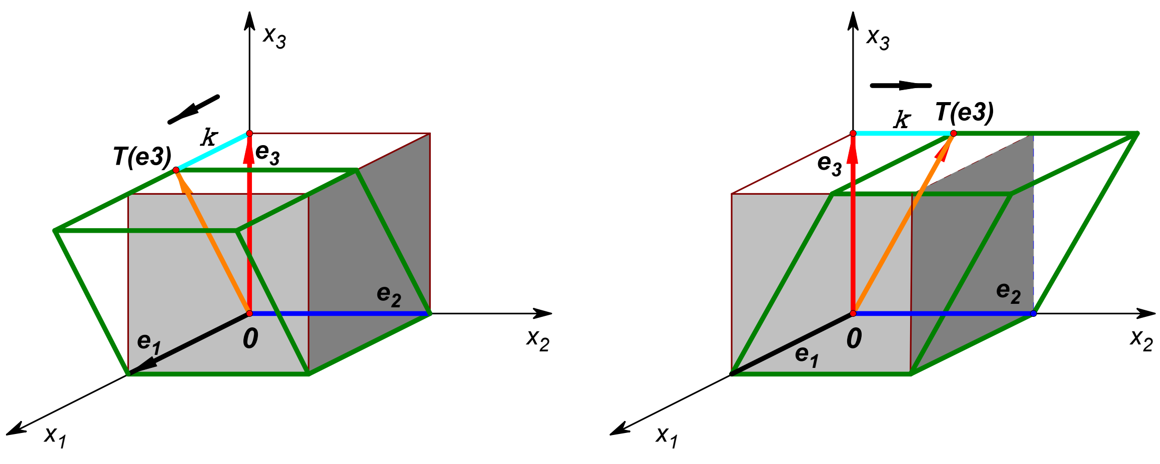



图 5  正方体的切变示意图

当我们应用求变换矩阵的定理求解其矩阵时，就会发现这两种情况都是单位向量 $\mathbf{e}_3$ 发生了变化，而 $\mathbf{e}_1, \mathbf{e}_2$ 没有变化。或者说，单位向量 $\mathbf{e}_3$ 只有两个坐标方向可以切变，要么 $\mathbf{e}_1$ 方向要么 $\mathbf{e}_2$ 方向。因此不难得到左图所示的切变换矩阵为：

$$[T(\mathbf{e}_1),T(\mathbf{e}_2),T(\mathbf{e}_3)]=\left[\mathbf{e}_1,\mathbf{e}_2,\begin{pmatrix}k\\0\\1\end{pmatrix}\right]=\begin{bmatrix}1&0&k\\0&1&0\\0&0&1\end{bmatrix},$$

其中，$k$ 是单位向量 $\mathbf{e}_3$ 在 $x_1 0 x_3$ 平面上沿 $x_1$ 轴方向的移动距离。

同理，右图所示的切变换矩阵为：

$$[T(\mathbf{e}_1),T(\mathbf{e}_2),T(\mathbf{e}_3)]=\left[\mathbf{e}_1,\mathbf{e}_2,\begin{pmatrix}0\\k\\1\end{pmatrix}\right]=\begin{bmatrix}1&0&0\\0&1&k\\0&0&1\end{bmatrix},$$

其中，$k$ 是单位向量 $\mathbf{e}_3$ 在 $x_2 0 x_3$ 平面上沿 $x_2$ 轴方向的移动距离。

至此，还有四种切变的情况没有得到矩阵：因为单位向量 $\mathbf{e}_1$ 也有两个坐标方向可以切变，要么 $\mathbf{e}_2$ 方向要么 $\mathbf{e}_3$ 方向；单位向量 $\mathbf{e}_2$ 亦有两个坐标方向可以切变，要么 $\mathbf{e}_1$ 方向要么 $\mathbf{e}_3$ 方向。如果要把其余四种切变的情况罗列出来就是：

$\mathbf{e}_1$ 沿 $\mathbf{e}_2$ 方向切变：$[T(\mathbf{e}_1),\mathbf{e}_2,\mathbf{e}_3]=\left[\begin{pmatrix}1\\k\\0\end{pmatrix},\mathbf{e}_1,\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\k&1&0\\0&0&1\end{bmatrix}$；

$\mathbf{e}_1$ 沿 $\mathbf{e}_3$ 方向切变：$[T(\mathbf{e}_1),\mathbf{e}_2,\mathbf{e}_3]=\left[\begin{pmatrix}1\\0\\k\end{pmatrix},\mathbf{e}_1,\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\0&1&0\\k&0&1\end{bmatrix}$；

$\mathbf{e}_2$ 沿 $\mathbf{e}_1$ 方向切变：$[\mathbf{e}_1,T(\mathbf{e}_2),\mathbf{e}_3]=\left[\mathbf{e}_1,\begin{pmatrix}k\\1\\0\end{pmatrix},\mathbf{e}_2\right]=\begin{bmatrix}1&k&0\\0&1&0\\0&0&1\end{bmatrix}$；

$\mathbf{e}_2$ 沿 $\mathbf{e}_3$ 方向切变：$[\mathbf{e}_1,T(\mathbf{e}_2),\mathbf{e}_3]=\left[\mathbf{e}_1,\begin{pmatrix}0\\1\\k\end{pmatrix},\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\0&1&0\\0&k&1\end{bmatrix}$；

因此，以上六个变换矩阵是把三阶单位矩阵的某一行的 $k$ 倍加到另一行的基本初等矩阵 $T(i,j(k))$。作为特例，若 $k=0$，则基本初等矩阵（3）为单位矩阵，其几何意义是图形到其自身的恒等变换。

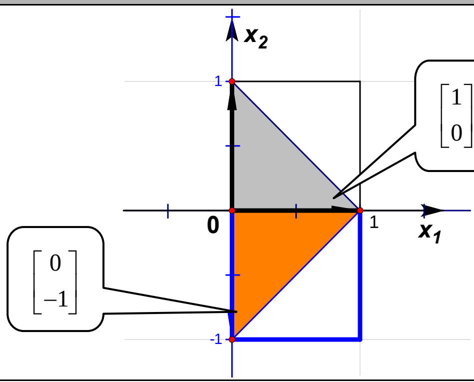
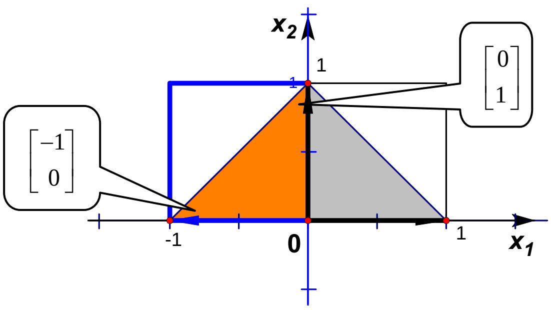
## 矩阵及其对应线性变换的几何图形

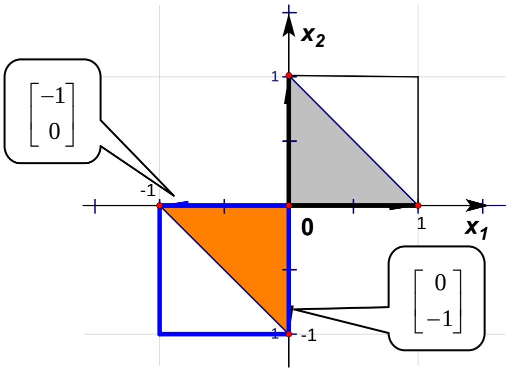
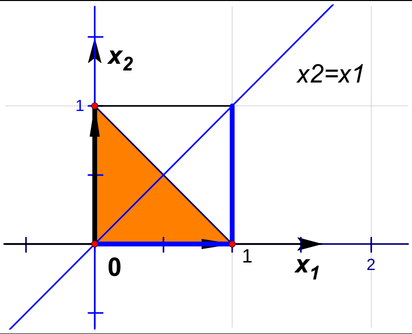
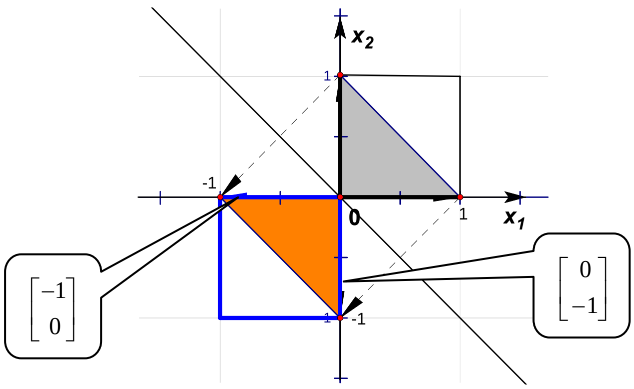
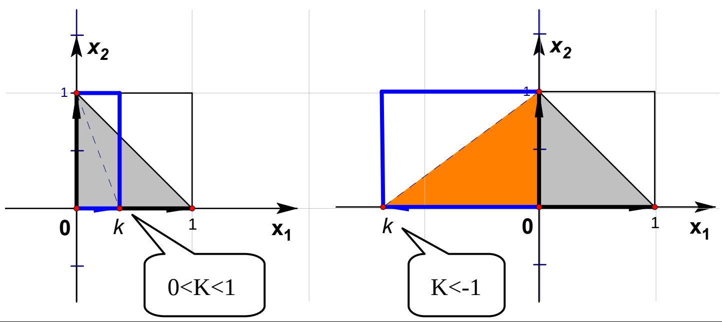
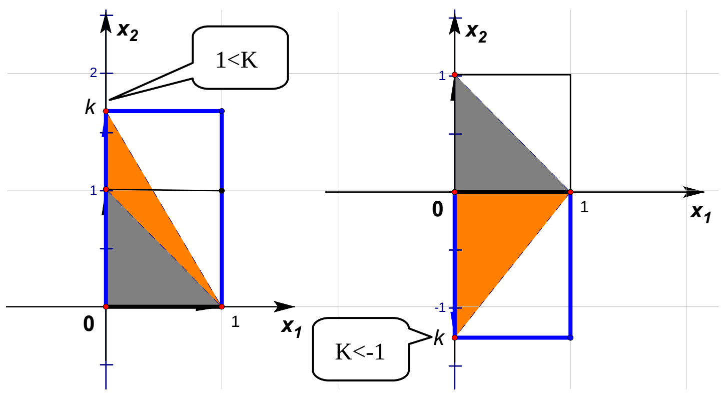
如果给你一个矩阵，如何马上理解其几何意义上的线性变换呢？其实很简单，你只要把前面已知变换求矩阵的的定理反着用就可以了。也就是说，把一个矩阵的列向量所张成平行多面体和同阶的单位矩阵的单位向量所张成的正多面体比对，就可以发现其对应的变换关系。一般情况下，我们只能对二阶和三阶矩阵的几何图形进行明显的对比。从前面的例子我们知道，二阶和三阶矩阵的正多面体就是正方形和正立方体。

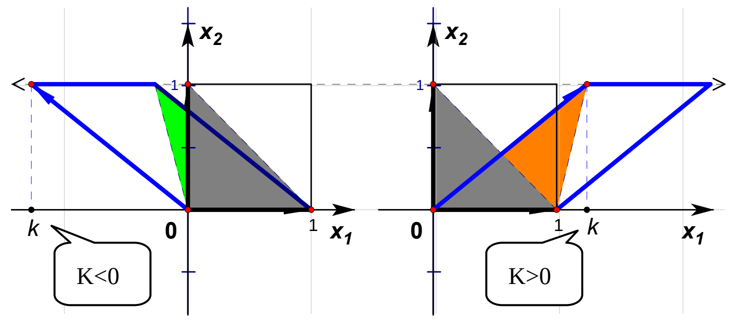
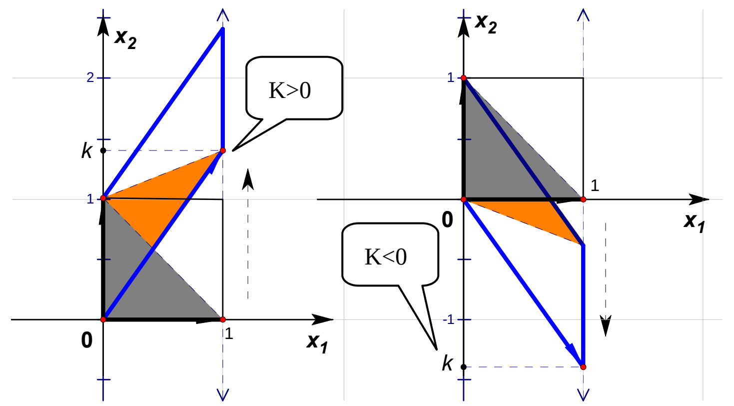
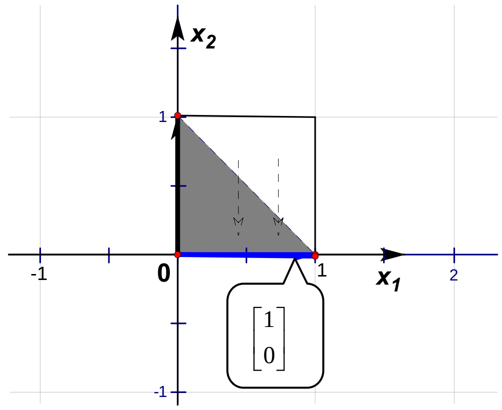
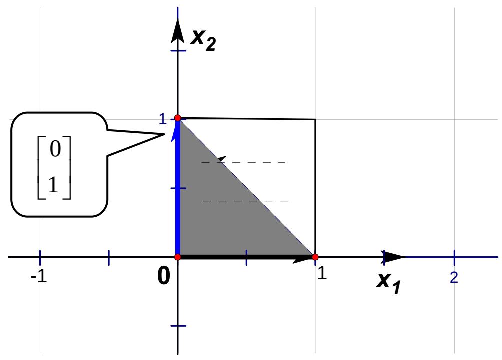
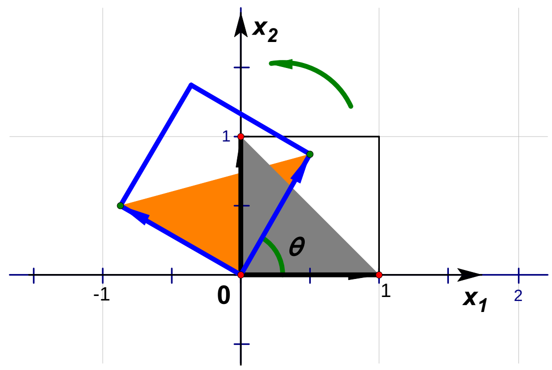
一个线性几何图形乘以一个矩阵，就会对这个图形进行一个线性变换或矩阵变换。几何图形是一个向量或者是一个三角形或者是一个四边形或者一个园，这些几何图形都被认为是一个向量集合所构成。当一个矩阵对众多不同的几何图形进行变换时就会显现不同的变换特征，同时也会加深我们对一个具体矩阵的意义的了解。比如我们常常见到二阶矩阵对一个单位正方形的图形进行变换，实际上就是把单位矩阵 **I** 的行向量或列向量所张成几何图形进行的一个变换。

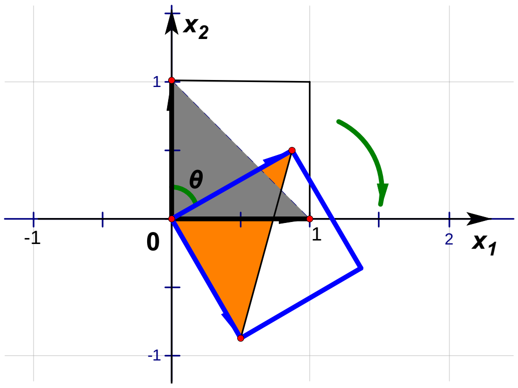
另外，如果被变换的几何图形是一个向量园，那么就会容易发现矩阵的特征值和特征向量的特质如果被变换的几何图形是一个正方形其特征向量的几何特点就不太明显。因此在下节的矩阵的特征向量和秩的几何意义的讨论中，我们总是把被变换的几何图形设为单位园或单位球。

下面我们给出各种二阶矩阵所对应的线性变换的平面几何图形，这里假设被变换的图形是一个位于第一象限内的三角形或正方形，他们由单位矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 所生成。这个汇总可帮助大家有个系统的认识。另外，在图标的最右列，给出了线性变换对应矩阵的行列式值，由行列式的值我们可以看出变换后的平面图形的面积的变化比率。

| 矩阵 | 变换前后的图像 | 对应的线性变换 | 矩阵的行列式 |
|---|---|---|---|
| $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ |  | 关于 $x_1$ 轴的镜像对称变换 | -1 |
| $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ |  | 关于 $x_2$ 轴的镜像对称变换 | -1 |

| | | | |
|---|---|---|---|
| $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ |   | 关于原点的对称变换 | 1 |
| $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ |   | 关于直线 $x_2 = x_1$ 的对称变换 | -1 |
| $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ |   | 关于直线 $x_2 = -x_1$ 的对称变换 | -1 |
| $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ |   | $x_1$ 轴（水平）方向的伸缩变换 | k |
| $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$ |   | $x_2$ 轴（垂直）方向的伸缩变换 | k |

| | | | |
|---|---|---|---|
| $\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$ |  | $x_1$ 轴（水平）方向的剪切变换 | 1 |
| $\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$ |  | $x_2$ 轴（垂直）方向的剪切变换 | 1 |
| $\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$ |  | $x_1$ 轴（水平）方向的投影变换 | 0 |
| $\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$ |  | $x_2$ 轴（垂直）方向的投影变换 | 0 |
| $\begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}$ |  | 逆时针旋转变换 | 1 |

| | | | |
|---|---|---|---|
| $\begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix}$ |   | 顺时针旋转变换 | 1 |

前面的行列式一章中知道，二阶行列式的几何意义就是两个列向量或行向量所构成的平行四边形的有向面积，实际上，二阶行列式的几何意义也可以看作是两个列向量或行向量所构成三角形的有向面积的 2 倍（如上表中的图形）。看作平行四边形或者是三角形都可以，这不会影响我们对矩阵的理解。

在上表中，我们得到一些结论：

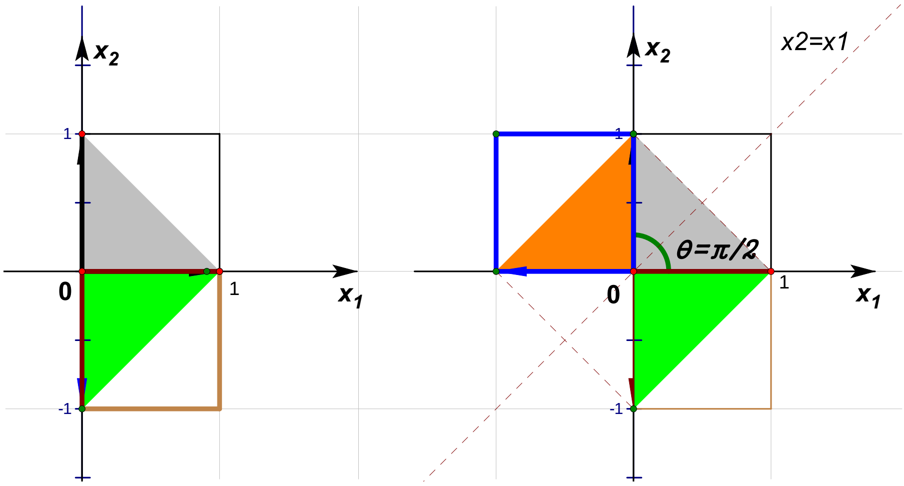
- 行列式等于 1，表示变换后图形的面积乘以 1，面积大小不会变化，图形的有向性也不会变化（通俗的讲，图面的正反面不会反转）；
- 行列式等于-1，表示变换后图形的面积乘以-1，面积大小不会变化，但图形的有向性会变化（通俗的讲，图面的正反面发生反转）；
- 行列式等于 0，表示变换后图形的面积乘以 0，面积收缩为 0，图形缩变为一维图形；
- 行列式等于 k，表示变换后图形的面积会变化 k 倍，K 大于 1，图形面积会放大，k 小于 1，图形面积会缩小，图形的有向性也会因 k 是否负值而发生变化；

## 矩阵的乘法运算律的几何意义

与前面的章节比较起来，本节利用了线性变换的概念来再次回顾矩阵乘积运算的几何意义，由此可更深入地认识矩阵运算性质。

- **两个矩阵相乘是两个线性变换的复合**

例如，对于平面坐标系 $x_1 0 x_2$ 上由单位矩阵所张成的第一象限的正方形实施关于 $x_1$ 轴的反射变换 $\mathbf{A}=\begin{bmatrix}1 & 0 \\ 0 & -1\end{bmatrix}$（见下图左）再实施关于标准轴 $x_2 = x_1$ 的反射变换 $\mathbf{B}=\begin{bmatrix}0 & 1 \\ 1 & 0\end{bmatrix}$，得到的图形正好是将原来的正方形逆时针旋转 90° 得到的（见下图右）。



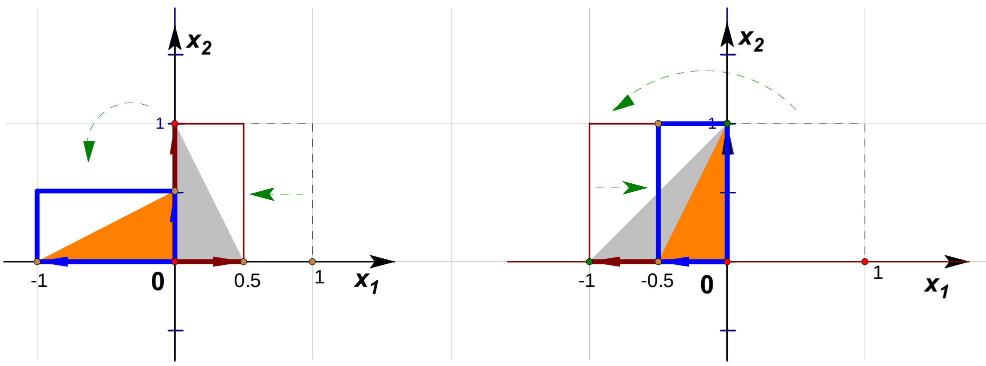

这一结果用矩阵运算表示为 $\mathbf{BA}=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&0\\0&-1\end{bmatrix}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$，即两个矩阵相乘得到一个新矩阵，新矩阵所表示的变换正是原来两个矩阵表示的变换的复合。这正是矩阵乘法的意义，即两个矩阵相乘表示连续两次实施变换。这一特例同时也可以帮助大家认识“旋转变换可以由连续两次反射变换来实现”这一性质。

- 矩阵的乘法不满足交换律

根据上述对矩阵乘法的理解，容易直观得出矩阵运算满足结合律。是否满足交换律呢？仍然看上面的例子。

如果先对第一象限的正方形进行实施关于标准轴 $x_2=x_1$ 的反射变换 $\mathbf{B}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，再实施关于 $x_1$ 轴的反射变换 $\mathbf{A}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$，得到的图形将是将原来的正方形顺时针旋转 90° 的结果。这与前面的结论不同，因为其中 $\mathbf{B}$ 的变换并没有对图形的外形造成变化，因而 $\mathbf{BA}\neq\mathbf{AB}$。

我们再看个例子。同样的，对在第一象限的由单位矩阵列向量所张成的正方形图形，先向靠近 $x_2$ 轴的方向压缩一半，再逆时针旋转 90° 得到的结果（见下图左）与先逆时针旋转 90°，再向靠近 $x_2$ 轴的方向压缩一半得到的结果（见下图右）的比较，大家应确信交换变换的顺序得到的结果一般来说是不同的。



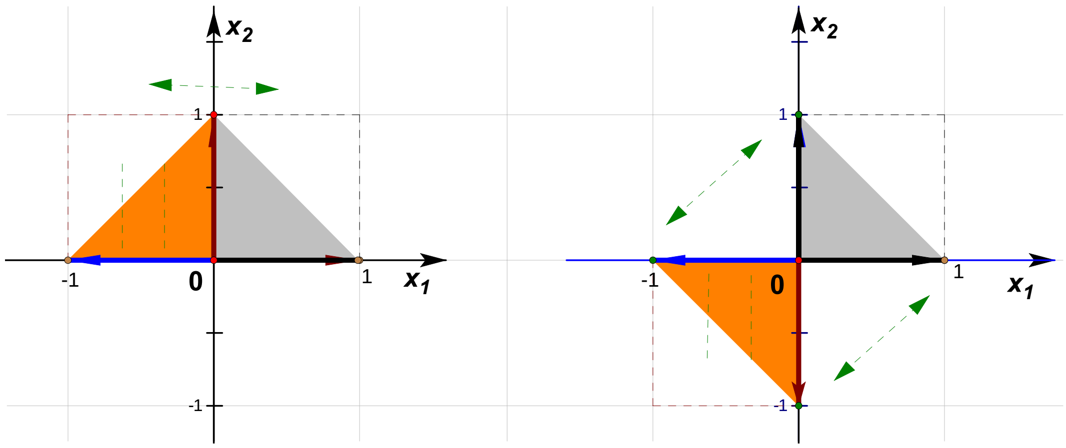

即：

$$\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 0.5\end{bmatrix}\neq\begin{bmatrix}1 & 0\\0 & 0.5\end{bmatrix}\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$$

因此，矩阵的乘法一般不满足交换律。但需要注意的是，不过有些情形矩阵的乘法可以满足交换律，例如连续多次旋转或连续两多次压缩变换是可以交换顺序的。

- 矩阵的乘法不满足消去律

根据投影变换的特点，把一个图形先作关于 $x_2$ 轴的反射变换，再向 $x_1$ 轴作投影变换的结果（见图左），与先作关于坐标原点的对称变换，再向 $x_1$ 轴作投影变换得到的结果（见图右）是一样的。




然而，虽然有：$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$，但 $\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}\neq\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$。因此，矩阵的乘法不满足消去律。待续。。。