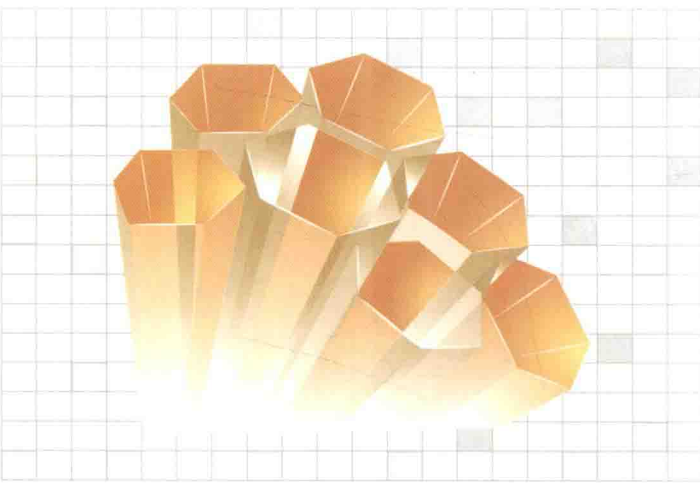
信息环境下大学数学课程改革系列教材
高等学校应用型创新型人才培养系列教材

线性代数疑难释义

张鹏鸽　高淑萍　编著



西安电子科技大学出版社
http://www.xduph.com

信息环境下大学数学课程改革系列教材
高等学校应用型创新型人才培养系列教材

线性代数疑难释义

张

西安电子科技大学出版社

内 容 简 介

全书共分为 6 章，第 1 章讲述了线性代数的相关概念在二维、三维空间中的几何解释；第 2 章针对矩阵的初等变换在本课程中的应用展开讨论；第 3 章介绍了分块矩阵及其应用；第 4 章重点讨论了线性方程组与矩阵、向量组及向量空间的关系；第 5 章给出了矩阵的三种标准形以及可逆矩阵的等价条件；第 6 章介绍了矩阵的四个子空间及其关系.

本书可供高等院校理工科、经管类等专业学生选用，也可作为相关工作人员的学习参考书.

图书在版编目(CIP)数据

线性代数疑难释义/张鹏鸽，高淑萍主编. —西安：西安电子科技大学出版社，2015.9

高等学校应用型创新型人才培养系列教材

ISBN 978-7-5606-3852-2

Ⅰ.①线… Ⅱ.①张… ②高… Ⅲ.①线性代数—高等学校—教材

Ⅳ.①O151.2

中国版本图书馆 CIP 数据核字(2015)第 216123 号

策　　划　毛红兵
责任编辑　王　飞　毛红兵
出版发行　西安电子科技大学出版社(西安市太白南路 2 号)
电　　话　(029)88242885　88201467　　邮　　编　710071
网　　址　www.xduph.com　　电子邮箱　xdupfxb001@163.com
经　　销　新华书店
印刷单位　陕西华沐印刷科技有限责任公司
版　　次　2015 年 9 月第 1 版　2015 年 9 月第 1 次印刷
开　　本　787 毫米×1092 毫米　1/16　印张 4.875
字　　数　110 千字
印　　数　1～6000 册
定　　价　10.00 元
ISBN 978-7-5606-3852-2/O

XDUP 4144001-1

前　言

理论上线性代数是一门高度抽象的数学学科，从现代科学的观点看来，它也是一门应用性很强的学科，尤其是计算机科学的发展更是扩大了该学科的应用领域.

随着西安电子科技大学牵头承担的教育部教改项目——“用 MATLAB 和建模实践改造工科线性代数课程”(目前已结题)的实施，以及与兄弟院校之间的交流、合作，编者发现线性代数中依然存在一些困扰着任课教师和学生的问题，大家普遍反映线性代数难教、难学.

为此，编者在实施、总结项目的基础上，编写了本书. 本书从不同的角度分析了线性代数的重点、难点. 比如第 1 章重点从几何角度给出了线性代数中一些抽象概念的直观解释，以方便学生理解、掌握. 矩阵的初等变换是解决线性代数中大多数问题的普遍方法，本书第 2 章对其进行了归纳、总结. 线性方程组是线性代数的一条主线，第 4 章针对线性方程组与矩阵、向量组以及向量空间的关系展开讨论，不但让学生能体会到主线的作用，更重要的是帮助学生理解矩阵的秩、向量组的线性相关性、向量空间的基、维数、解空间的基础解系等抽象概念的由来，以达到高屋建瓴的效果. 第 5 章重点给出了矩阵的三种标准形及可逆矩阵的等价条件. 第 3 章分块矩阵及其应用及第 6 章矩阵的四个基本子空间读者可作为扩充知识进行研读. 其中第 1 章由高淑萍教授撰写，第 2～6 章由张鹏鸽副教授撰写。

本书具有以下特色：

(1) 直观性. 对线性代数中的重要概念，如向量组的线性相关性、行列式、方程组、线性变换、特征值与特征向量等给出了直观的几何解释，便于学生掌握.

(2) 整体性. 注重同一个方法在不同问题中的应用，如第 2 章中矩阵的初等变换贯穿着全书内容，本书将其归纳、总结，并讨论了该方法的具体应用.

(3) 分类性. 如对二次型的分类、矩阵的标准形的分类等进行了讨论.

本书可能存在疏漏或不妥之处，敬请广大读者不吝指正.

编者

2015 年 9 月

目 录

第 1 章　线性代数中相关概念的几何解释

编者从多年的教学实践和学习过程中发现，几何直观是领悟数学的一种有效途径. 线性代数是研究线性空间及其空间上的线性变换的学科，它所具有的高度抽象性和逻辑性使得学生望而生畏. 事实上，要掌握一个概念，就必须弄清这个概念的内涵和外延. 由于线性代数中概念的抽象程度较高，其内涵变小，相对而言，其外延就大，因此，学生对它的掌握和理解容易发生障碍. 下面从行列式、线性方程组、向量组的线性相关性、线性变换、特征值与特征向量及二次型这几个概念入手，讲解其在几何中的“原始”意义，便于学生理解和掌握这些抽象概念.

1.1　行列式的几何意义

在大多数线性代数教材中，行列式的概念是首先介绍的，虽然学生在中学已经学过二、三阶行列式，但实际上大多数学生对行列式的概念知之甚少，或者只是机械地背算式.

行列式是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，这一点要与矩阵区别开. 矩阵只是一个数表，行列式则是对一个方形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.

1.1.1　二阶行列式的几何意义

这里首无介绍一个一阶行列式的意义.

一阶行列式$|a_1|=a_1$，即 a_1 的一阶行列式就是数 a_1，也可说是向量 $\boldsymbol{a}_1$ 本身，这个数 a_1 的本身是一维坐标轴上的有向长度(如图 1－1 所示). 这里强调的是“有向”，因为长度是有向的，是一个向量，这是个很重要的概念.

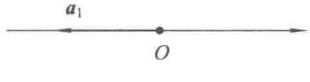


图 1－1

下面继续讨论二阶行列式.

二阶行列式 $\boldsymbol{D}=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$ 的几何意义是 xoy 平面上以行向量 $\boldsymbol{a}=(a_1, a_2)$，$\boldsymbol{b}=(b_1, b_2)$ 为邻边的平行四边形的有向面积. 为什么? 下面推导这一结论.

我们在二维几何空间 $\mathbf{R}^2$ 中取定一个直角坐标系$[\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2]$，设 $\boldsymbol{a}=a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2$，$\boldsymbol{b}=b_1\boldsymbol{\varepsilon}_1+$

$b_2\boldsymbol{\varepsilon}_2$. 考察以 $\boldsymbol{a}$、$\boldsymbol{b}$ 为邻边的平行四边形的面积与构成它的两个向量之间的关系（如图 1-2 所示）：

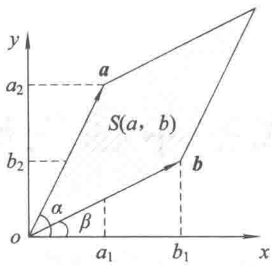


图 1-2

则以 $\boldsymbol{a}$、$\boldsymbol{b}$ 为边的平行四边形的面积为

$$S(\boldsymbol{a},\ \boldsymbol{b}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \sin < \boldsymbol{a} \cdot \boldsymbol{b} >$$

其中，$|\boldsymbol{a}| = \sqrt{a_1^2 + a_2^2}$，$|\boldsymbol{b}| = \sqrt{b_1^2 + b_1^2}$，$\sin < \boldsymbol{a},\ \boldsymbol{b} >$ 为向量 $\boldsymbol{a}$、$\boldsymbol{b}$ 之间的夹角的正弦.

$$\sin < \boldsymbol{a},\ \boldsymbol{b} > = \sin(\alpha - \beta) = \sin\alpha\ \cos\beta - \cos\alpha\ \sin\beta$$

参照图 1-2 中的关系将三角函数用坐标值表示出来，即

$$\sin < \boldsymbol{a},\ \boldsymbol{b} > = \frac{a_2}{|\boldsymbol{a}|} \cdot \frac{b_1}{|\boldsymbol{b}|} - \frac{a_1}{|\boldsymbol{a}|} \cdot \frac{b_2}{|\boldsymbol{b}|} = \frac{a_2 b_1 - a_1 b_2}{|\boldsymbol{a}| \cdot |\boldsymbol{b}|}$$

将上式整理，得

$$S_{平行四边形} = |\boldsymbol{a}| \cdot |\boldsymbol{b}|\ \sin < \boldsymbol{a},\ \boldsymbol{b} > = a_2 b_1 - a_1 b_2 = -\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

若以 $\boldsymbol{a}$、$\boldsymbol{b}$ 为邻边的平行四边形与构成它的两个向量之间的关系如图 1-3 所示，则

$$S(\boldsymbol{a},\ \boldsymbol{b}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}|\ \sin < \boldsymbol{a},\ \boldsymbol{b} > = a_1 b_2 - a_2 b_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

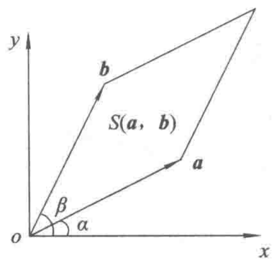


图 1-3

1.1.2 三阶行列式的几何意义

三阶行列式的几何意义是其行向量或列向量所张成的平行六面体的有向体积.

行列式的展开定理是线性代数中的一个重要定理，在行列式计算、矩阵计算中十分有用. 如果记三阶行列式的三个行向量分别为 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_3$，且 $\boldsymbol{\beta} = (A_{31},\ A_{32},\ A_{33})$，其中 A_{31}、A_{32}、A_{33} 表示代数余子式，$\boldsymbol{\beta} = \boldsymbol{\alpha}_1 \times \boldsymbol{\alpha}_2$（$\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 的向量积），它的模等于以向量 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 为邻边的平行四边形的面积，三阶行列式的值等于 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_3$ 的内积，即 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_3$ 的混合积（$\boldsymbol{\alpha}_1 \times$

$\boldsymbol{\alpha}_2 \cdot \boldsymbol{\alpha}_3$，即

$$\begin{vmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33}$$

$$= (\boldsymbol{\beta}, \boldsymbol{\alpha}_3) = (\boldsymbol{\alpha}_1 \times \boldsymbol{\alpha}_2) \cdot \boldsymbol{\alpha}_3$$

即以向量 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_3$ 为相邻棱的平行六面体的体积 V(如图 1-4 所示)，可以表示为

$$\begin{aligned} V &= S \cdot h \\ &= \| \boldsymbol{\alpha}_1 \times \boldsymbol{\alpha}_2 \| \cdot (\| \boldsymbol{\alpha}_3 \| \cdot \cos\theta) \\ &= (\boldsymbol{\alpha}_1 \times \boldsymbol{\alpha}_2) \cdot \boldsymbol{\alpha}_3 \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \end{aligned}$$

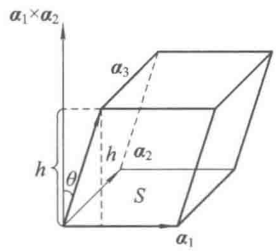


图 1-4

其中，S 表示以 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 为邻边的平行四边形的面积 $S = \| \boldsymbol{\alpha}_1 \times \boldsymbol{\alpha}_2 \|$，$h$ 表示以 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 为邻边的底面上的高，θ 表示平行六面体与 β 的夹角.

1.2 线性方程组的几何意义

线性方程组理论是线性代数中最经典的理论，历史上，线性代数的第一个问题也是关于线性方程组的问题，下面以三元方程组

$$\begin{cases} x_1 + x_2 - x_3 = 5 \\ 2x_1 - 3x_2 + x_3 = 5 \\ -5x_1 + 2x_2 - 2x_3 = 0 \end{cases} \tag{1}$$

为例加以说明. 教材上利用高斯消元法可以求三元一次线性方程组(1)，即对方程组的增广矩阵作初等行变换化为阶梯形矩阵，获得同解方程组，进而求得原方程组的解. 高斯消元法的优点是：把中学学过的解线性方程组的初等方法推广到利用初等行变换求解方程组的高等方法上来，即实现了由初等数学方法到高等数学的过渡，这一方法完全可以机械操作，因此学生很少了解其背后的几何意义.

实际上，方程组(1)具有直观的几何意义，而且可以利用几何方法求得其解. 方程组(1)的 3 个三元一次方程可以看成是三维空间中的 3 个平面. 三元线性方程组(1)解的情况可以用这些平面上直线的交点个数和三维空间中平面的交线条数来判断.

在解析几何中，平面可以通过一个三元一次方程来描述；反过来，任意一个三元一次方程可以表示三维空间中的一个平面，设 m 个三元一次方程构成一个三元线性方程组.

1) 如果方程组有解，就意味着这 m 个方程代表 m 个平面有公共点

如果方程组有唯一解，则 m 个平面交于一点；如果方程组有无穷多个解，且基础解系中只含有一个向量，则平面组交于一条直线；若基础解系中含有两个向量，则平面组重合(即为同一个平面).

2) 若方程组无解，则平面不相交(平行)

下面从向量组秩的角度分析上面的结论，不妨设 $m=3$.

设有下列线性方程组：

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

记

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

$$\boldsymbol{\gamma}_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \end{bmatrix}$$

$$\boldsymbol{\gamma}_2 = \begin{bmatrix} a_{21} & a_{22} & a_{23} & b_2 \end{bmatrix}$$

$$\boldsymbol{\gamma}_3 = \begin{bmatrix} a_{31} & a_{32} & a_{33} & b_1 \end{bmatrix}$$

$$\boldsymbol{\beta}_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix}$$

$$\boldsymbol{\beta}_2 = \begin{bmatrix} a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$\boldsymbol{\beta}_3 = \begin{bmatrix} a_{31} & a_{32} & a_{33} \end{bmatrix}$$

方程组中三个方程分别代表的平面记为 π_1，π_2，π_3，其对应法向量分别为 $\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，$\boldsymbol{\beta}_3$.

(1) 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)=3$ 时，$R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)=3$，则方程组有唯一解.

(2) 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)=2$ 时，$R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=2$ 或 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)=3$.

① 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=2$(即增广矩阵的秩＝系数矩阵的秩)时，方程组有无穷多个解，且基础解系只含有一个解向量.

几何意义：三个平面相交于一条直线.

进一步分析可知：

(a) 当 $\boldsymbol{\gamma}_1$，$\boldsymbol{\gamma}_2$，$\boldsymbol{\gamma}_3$ 中有两个线性相关时，其几何意义为：有两个平面重合，与第三个平面交于一条直线.

(b) 当 $\boldsymbol{\gamma}_1$，$\boldsymbol{\gamma}_2$，$\boldsymbol{\gamma}_3$ 中两两线性无关时，其几何意义为：三个平面交于一条直线.

② 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=3$(即增广矩阵的秩≠系数矩阵的秩)时，方程组无解.

进一步分析可知：

(a) 当 $\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，$\boldsymbol{\beta}_3$ 中有两个线性相关时，其几何意义为：其中两个平面平行，分别与第三个平面相交.

(b) 当 $\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，$\boldsymbol{\beta}_3$ 中两两线性无关时，其几何意义为：三个平面两两相交.

(3) 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)=1$ 时，$R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3)=1$(即行秩＝列秩)，此时三个平面的法向量平行，三个平面平行.

① 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=1$ 时，方程组有无穷多个解，且基础解系中含有两个解向量. 其几何意义为：三个平面相交于一个平面，三个平面重合.

② 当 $R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta})=2$ 时，方程组无解.

(a) 当 $\boldsymbol{\gamma}_1$，$\boldsymbol{\gamma}_2$，$\boldsymbol{\gamma}_3$ 中有两个线性相关时，其几何意义为：其中两个平面重合，与第三个平面平行.

(b) 当 $\boldsymbol{\gamma}_1$，$\boldsymbol{\gamma}_2$，$\boldsymbol{\gamma}_3$ 中两两线性相关时，其几何意义为：三个平面平行且互不重合.

1.3 向量组的线性相关性的几何意义

在线性代数中，关于向量组的线性相关、线性无关、极大无关组和秩的概念是教学中的一个难点，但借助于三维空间的几何直观，将会有助于抽象概念的理解.

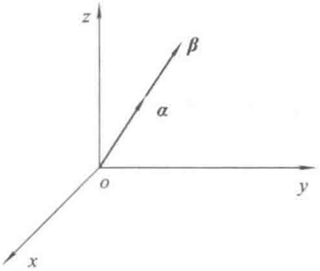


图 1-5

我们从三维几何空间中向量之间的关系入手.

已知 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}\in\mathbf{R}^3$，且 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 共线，如图 1-5 所示.

$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 共线$\Leftrightarrow\boldsymbol{\beta}=k\boldsymbol{\alpha}$

$\Leftrightarrow k\boldsymbol{\alpha}+(-1)\boldsymbol{\beta}=\mathbf{0}$($\mathbf{0}$ 为零向量)

$\Leftrightarrow$存在一组不全为零的数 k、-1，使得 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 的线性组合为零向量.

$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 共线亦称为 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 线性相关.

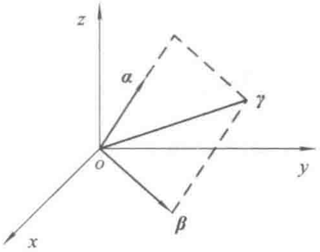


图 1-6

已知 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}\in\mathbf{R}^3$，且 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 共面，如图 1-6 所示.

$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 共面$\Leftrightarrow\boldsymbol{\gamma}=k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}$

$\Leftrightarrow k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}+(-1)\boldsymbol{\gamma}=\mathbf{0}$($\mathbf{0}$ 为零向量)

$\Leftrightarrow$存在一组不全为零的数 k_1、k_2、-1，使得 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 的线性组合为零向量.

$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 共面亦称为 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 线性相关.

于是，可得三维空间中向量组的线性相关性与其对应的几何意义，如表 1-1 所示.

表 1-1 三维空间中向量组的线性相关性及其对应的几何意义

线性相关性	几何意义
$\boldsymbol{\alpha}$ 线性相关	$\boldsymbol{\alpha}$ 为非零向量
$\boldsymbol{\alpha}$ 线性无关	$\boldsymbol{\alpha}$ 为零向量
$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 线性相关	$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 共线
$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 线性无关	$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 可张成一平面
$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 线性相关	$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 共面
$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 线性无关	$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 可张成一空间

自然地，上述三维几何空间中两个向量的共线和三个向量的共面可以推广到 n 维空间中，进而得到线性组合、线性表示及组合系数等概念，同时可给出向量组的线性相关和线性无关的定义，并能很好地理解相关结论，如 $n+1$ 个 n 维向量必线性相关，因为 n 维向量必属于 n 维空间，而 n 维空间意味着基向量所含向量的个数为 n，即 n 维空间中最大线性

无关组中所含的向量个数必为 n，则再添一个向量必线性相关，所以，$n+1$ 个 n 维向量必线性相关.

1.4 线性变换的几何意义

首先回顾一下变换(映射)、线性映射和线性变换的概念.

1) 变换

设有两个非空集合 U、V，若对于 V 中任一元素 α，按照一定法则，总有 V 中一个确定的元素 β 和它对应，则这个对应规则称为从集合 U 到集合 V 的变换(或映射)，记作 $\beta=T(\alpha)$，$(\alpha\in U)$.

2) 线性映射

设 U_m、V_n 分别是实数域上 m 维和 n 维线性空间，T 是一个从 U_m 到 V_n 的映射，且满足：

(ⅰ) $\forall\alpha_1,\alpha_2\in U_m$，有 $T(\alpha_1+\alpha_2)=T(\alpha_1)+T(\alpha_2)$；

(ⅱ) $\forall\alpha\in U_m$，$k\in\mathbf{R}$，有 $T(k\alpha)=kT(\alpha)$.

则称 T 为 U_m 到 V_n 的线性映射.

3) 线性变换

若 $U_m=V_n$，即线性映射发生在同一个线性空间上，则称之为线性变换.

因此，线性变换可作为特殊的线性映射，而在线性代数中，主要讨论由矩阵所决定的线性变换的各种特性.

下面通过实例介绍线性变换的几何意义.

例 1 定义线性变换 $T:\mathbf{R}^2\to\mathbf{R}^2$ 为

$$T(\boldsymbol{\alpha})=\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-y\\x\end{bmatrix},\quad \boldsymbol{\alpha}\in\mathbf{R}^2$$

求 $\boldsymbol{\alpha}_1=\begin{bmatrix}4\\1\end{bmatrix}$，$\boldsymbol{\alpha}_2=\begin{bmatrix}2\\3\end{bmatrix}$和 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2=\begin{bmatrix}6\\4\end{bmatrix}$在变换 T 下的像，并说明其几何意义.

解 $T(\boldsymbol{\alpha}_1)=\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}=\begin{bmatrix}-1\\4\end{bmatrix}$

$$T(\boldsymbol{\alpha}_2)=\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}-3\\2\end{bmatrix}$$

$$T(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)=\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}6\\4\end{bmatrix}=\begin{bmatrix}-4\\6\end{bmatrix}=T(\boldsymbol{\alpha}_1)+T(\boldsymbol{\alpha}_2)$$

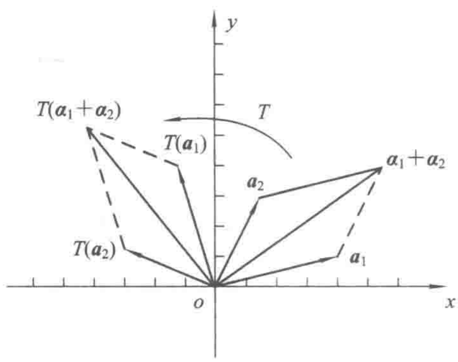


图 1-7

图 1-7 表明变换 T 把 $\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 和 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 绕原点逆时针旋转了 90°. 实际上，T 把 α_1、α_2 确定的平行四边形变成了 $T(\boldsymbol{\alpha}_1)$、$T(\boldsymbol{\alpha}_2)$确定的平行四边形.

下面介绍几类简单的线性变换及其几何直观.

(1) 旋转变换(见图 1-8～图 1-10)

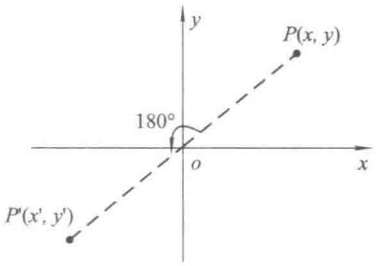


图 1－8

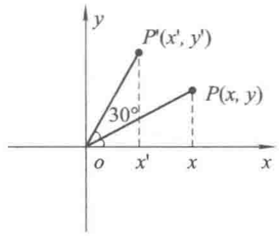


图 1－9

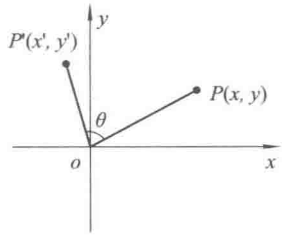


图 1－10

$$P(x,\ y) \xrightarrow[\text{的旋转变换}]{\text{旋转角为 }180^\circ} P'(x',\ y')$$

$$\begin{cases} x' = -x \\ y' = -y \end{cases}$$

即

$$\begin{cases} x' = -x + 0 \cdot y \\ y' = 0 \cdot x - y \end{cases} \leftrightarrow \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$P(x,\ y) \xrightarrow[\text{的旋转变换}]{\text{旋转角为 }30^\circ\text{度}} P'(x',\ y')$$

$$\begin{cases} x' = \dfrac{\sqrt{3}}{2}x - \dfrac{1}{2}y \\ y' = \dfrac{1}{2} + \dfrac{\sqrt{3}}{2}y \end{cases} \leftrightarrow \begin{bmatrix} \dfrac{\sqrt{3}}{2} & -\dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \end{bmatrix}$$

$$P(x,\ y) \xrightarrow[\text{的旋转变换}]{\text{旋转角为 }\theta} P'(x',\ y')$$

$$\begin{cases} x' = \cos\theta \cdot x - \sin\theta \cdot y \\ y' = \sin\theta \cdot x + \cos\theta \cdot y \end{cases} \leftrightarrow \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

(2) 反射变换(见图 1－11～图 1－13)

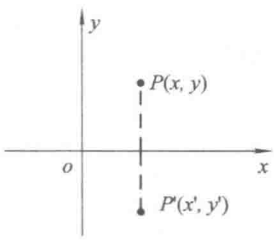


图 1－11

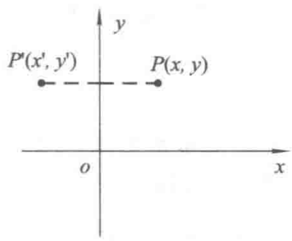


图 1－12

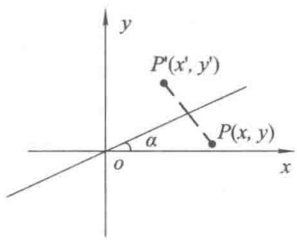


图 1－13

$$P(x,\ y) \xrightarrow[\text{反射变换}]{\text{关于 }x\text{ 轴的}} P'(x',\ y')$$

$$\begin{cases} x' = x = x + 0 \cdot y \\ y' = y = 0 \cdot x + y \end{cases} \leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$P(x,\ y)\xrightarrow[\text{反射变换}]{\text{关于 } y \text{ 轴的}}P'(x',\ y')$$

$$\begin{cases}x'=-x=-x+0\cdot y\\ y'=y=0\cdot x+y\end{cases}\leftrightarrow\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

$$P(x,\ y)\xrightarrow[\text{的反射变换}]{\text{关于直线 } l(\text{斜率 } k=\tan\alpha)}P'(x',\ y')$$

$$\begin{cases}x'=\cos2\alpha\cdot x-\sin2\alpha\cdot y\\ y'=\sin2\alpha\cdot x-\cos2\alpha\cdot y\end{cases}\leftrightarrow\begin{bmatrix}\cos2\alpha & \sin2\alpha\\ \sin2\alpha & -\cos2\alpha\end{bmatrix}$$

(3) 伸缩变换(见图 1-14):

$$y=\sin x\xrightarrow{\text{伸缩变换}}y=2\ \sin2x$$

$$\begin{cases}x'=\dfrac{1}{2}x=\dfrac{1}{2}x+0\cdot y\\ y'=2y=0\cdot x+2\cdot y\end{cases}\leftrightarrow\begin{bmatrix}\dfrac{1}{2} & 0\\ 0 & 2\end{bmatrix}$$

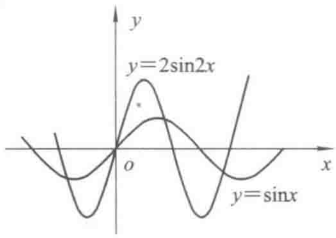


图 1-14

一般地，在直角坐标系内，将每一个点的横坐标变为原来的 k_1 倍，纵坐标变为原来的 k_2 倍(k_1、k_2 均为非零常数)的线性变换，对应的变换公式为

$$\begin{cases}x'=k_1x=k_1x+0\cdot y\\ y'=k_2y=0\cdot x+k_2y\end{cases}\leftrightarrow\begin{bmatrix}k_1 & 0\\ 0 & k_2\end{bmatrix}$$

1.5 特征值与特征向量的几何意义

定义 设 $\boldsymbol{A}$ 是 n 阶方阵，如果存在常数 λ 和 n 维非零列向量 $\boldsymbol{x}$ 使关系式 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 成立，则数 λ 称为方阵 $\boldsymbol{A}$ 的特征值，非零向量 $\boldsymbol{x}$ 称为 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量.

将 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 改写为 $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，则对于方阵 $\boldsymbol{A}$ 的特征值和特征向量的求解问题就转化为齐次线性方程组 $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 有无非零解的判断及求非零解的过程.

由于工程技术中一些问题，如振动问题和稳定性问题，常归结为求一个方阵的特征值和特征向量的问题，然而一般教材对特征值与特征向量的引入均如上定义所述，该定义本身具有高度的抽象性，学生对这一知识只能机械地记住计算过程，而定义本身及其内涵理解依然是相当模糊地. 然而，学生若能理解该定义的几何内涵和本质特征，则掌握起来便轻松自然，并在后续课程中使用时也会得心应手. 下面针对特征值与特征向量的几何意义作一介绍.

根据定义 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，方阵 $\boldsymbol{A}$ 与非零向量 $\boldsymbol{x}$ 的乘积实质是对非零向量 $\boldsymbol{x}$ 做了一个线性变换(其变换矩阵为 $\boldsymbol{A}$)；用常数 λ 对非零向量做数乘其实质为对向量 $\boldsymbol{x}$ 做同方向上的伸缩变换或者是反方向上的伸缩变换. 综上，方阵 $\boldsymbol{A}$ 的特征值与特征向量的定义可以理解为：对于方阵 $\boldsymbol{A}$ 对应的空间中的线性变换，存在某些向量 $\boldsymbol{x}$ 在其变换下，$\boldsymbol{x}$ 的主方向不发生改变，只是在同方向或者是反方向上进行伸缩，这些向量则是方阵 $\boldsymbol{A}$ 对应于某个特征值 λ 的特征向量，而 $\boldsymbol{x}$ 对应的那个特征值 λ 即为该向量 $\boldsymbol{x}$ 做伸缩变换时的伸缩系数.

于是，容易判定一个给定的向量是否为一个方阵的特征向量以及一个指定的向量是否为该方阵的特征值.

例 2 设 $\boldsymbol{A}=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}-1\\ 1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}2\\ 1\end{bmatrix}$，判断 $\boldsymbol{u}$、$\boldsymbol{v}$ 是否为 $\boldsymbol{A}$ 的特征向量.

解

$$\boldsymbol{Au}=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix}=\begin{bmatrix}-5\\ -1\end{bmatrix}\neq\lambda\begin{bmatrix}-1\\ 1\end{bmatrix}$$

$$\boldsymbol{Av}=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}\begin{bmatrix}2\\ 1\end{bmatrix}=\begin{bmatrix}4\\ 2\end{bmatrix}=2\begin{bmatrix}2\\ 1\end{bmatrix}=2\boldsymbol{v}$$

易见，$\boldsymbol{v}$ 是 $\boldsymbol{A}$ 的特征值 2 所对应的特征向量，$\boldsymbol{u}$ 不是，且 $\boldsymbol{A}$ 拉伸了 $\boldsymbol{v}$.

如图 1－15 所示，显然，$\boldsymbol{u}$ 不是 $\boldsymbol{A}$ 的特征向量，$\boldsymbol{v}$ 是 $\boldsymbol{A}$ 的特征值 2 的特征向量.

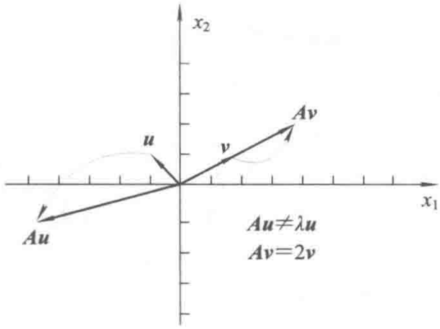


图 1－15

1.6 二次型的几何意义

二次型理论起源于解析几何中二次曲线、二次曲面的问题. 一般的 n 元二次型化为标准形问题在很多工程中有广泛的应用.

二次型是指含有 n 个变量 x_1，x_2，…，x_n 的二次齐次多项式：

$$\begin{aligned}f(x_1, x_2, \cdots, x_n) &= a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2\\ &\quad +2a_{12}x_1x_2+2a_{13}x_1x_3+\cdots+2a_{1n}x_1x_n\\ &\quad +2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+2a_{(n-1)n}x_{n-1}x_n\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\\ &= [x_1, x_2, \cdots, x_n]\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \ddots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}\\ &= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\end{aligned}$$

其中，$\boldsymbol{X}=[x_1, x_2, \cdots, x_n]^{\mathrm{T}}$，$\boldsymbol{A}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \ddots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}$（$a_{ij}=a_{ji}$），即 $\boldsymbol{A}$ 为实对称矩阵，并

称 $\boldsymbol{A}$ 为二次型 f 的矩阵.

在某些情况下，当二次型不含交叉乘积项，即当二次型的矩阵是对角阵时，二次型更易于应用. 庆幸地是，交叉乘积项可以通过适当的变量替换消去.

假设对二次型 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 施行变量替换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，则有

$$f(\boldsymbol{X}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} = (\boldsymbol{P}\boldsymbol{Y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Y}) = \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P})\boldsymbol{Y} \triangleq f(\boldsymbol{Y})$$

即通过变量替换将二次型 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 化为新二次型 $f(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Y}$，其中 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为新二次型的矩阵.

若 $\boldsymbol{P}$ 为正交矩阵，则 $\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}^{-1}$，且 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角矩阵，此时二次型 $f(\boldsymbol{Y})$仅含平方项，不含交叉乘积项.

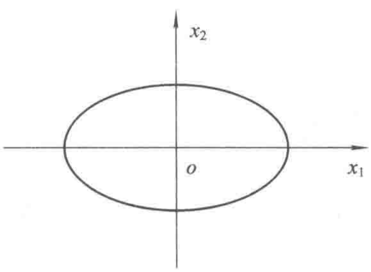
为此，下面介绍主轴定理.

定理　设 $\boldsymbol{A}$ 是一个 $n\times n$ 阶对称矩阵，则存在一个正交变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，将二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 化成仅含平方项的二次型 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Y}$，其中 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$.

定理中 $\boldsymbol{P}$ 的列称为二次型的主轴.

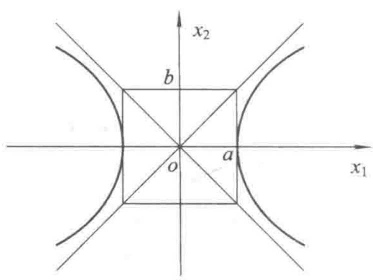
主轴的几何意义：设 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$，其中 $\boldsymbol{A}$ 是一个 2×2 可逆对称矩阵，令 C 为一常数，可以证明在 $\mathbf{R}^2$ 中满足 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=C$ 的全体 $\boldsymbol{X}$ 所构成的集合或者是一个椭圆(或圆)(如图 1-16 所示)——正平方项；双曲线(如图 1-17 所示)——一正一负平方项.

若 $\boldsymbol{A}$ 是对角阵，则图像位于标准位置.



$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1,\quad a>b>0$$

图 1-16



$$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1,\quad a>b>0$$

图 1-17

若 $\boldsymbol{A}$ 不是对角阵，则 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=C$ 的图像旋转到标准位置以外.

若经过一正交变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，将二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 化为只含平方项的新二次型 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Y}$，则在新的主轴下为标准位置.

例如：

$$Q(\boldsymbol{X}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} = 5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$$

通过求 $\boldsymbol{A}=\begin{bmatrix}5 & -2\\ -2 & 5\end{bmatrix}$的特征值 $\lambda_1=7$，$\lambda_2=3$ 及其对应的特征向量并作单位正交化，得

$$\boldsymbol{p}_1=\begin{bmatrix}\dfrac{1}{\sqrt{2}}\\[2mm] \dfrac{1}{\sqrt{2}}\end{bmatrix},\quad \boldsymbol{p}_2=\begin{bmatrix}-\dfrac{1}{\sqrt{2}}\\[2mm] \dfrac{1}{\sqrt{2}}\end{bmatrix}$$

记 $\boldsymbol{P}=[\boldsymbol{p}_1, \boldsymbol{p}_2]$.

令 $\boldsymbol{X}=\boldsymbol{PY}$，则

$$Q(\boldsymbol{X}) = 3y_1^2 + 7y_2^2 = 48$$

在新坐标系 y_1oy_2 下该图像位于标准位置(如图 1-18 所示).

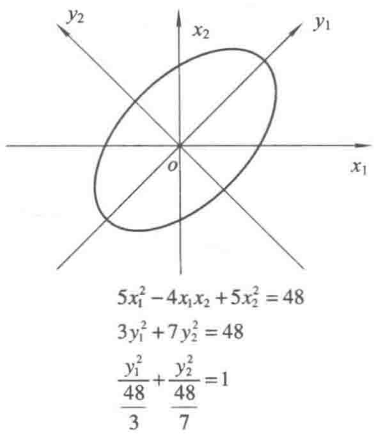


图 1-18

再如，二次型

$$Q(\boldsymbol{X}) = x_1^2 - 8x_1x_2 - 5x_2^2 = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$$

其中，$\boldsymbol{X}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$，$\boldsymbol{A}=\begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$.

由于含有交叉乘积项，故该二次型的图像是非标准图形，但通过对 $\boldsymbol{A}$ 施行正交变换，使得 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 可化为新坐标系下的标准图形，并可计算出在原坐标系下 $\boldsymbol{X}=\begin{bmatrix} 2 \\ -2 \end{bmatrix}$时 $Q(\boldsymbol{X})$的值.

与上述方法一致，可得正交矩阵

$$\boldsymbol{P} = \begin{bmatrix} \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \\ -\dfrac{1}{\sqrt{5}} & \dfrac{2}{\sqrt{5}} \end{bmatrix}$$

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{D} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\boldsymbol{X} = \boldsymbol{PY}$$

于是

$$\begin{aligned} Q(\boldsymbol{X}) &= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX} \xlongequal{\boldsymbol{X}=\boldsymbol{PY}} (\boldsymbol{PY})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{PY}) \\ &= \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APY} \\ &= \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{DY} = 3y_1^2 - 7y_2^2 \end{aligned}$$