第 21 卷 第 1 期
2018 年 1 月

高 等 数 学 研 究
STUDIES IN COLLEGE MATHEMATICS

Vol. 21, No. 1
Jan., 2018

doi:10.3969/j.issn.1008-1399.2018.01.005

Hurwitz 定理的矩阵证明

林开亮[1]，陈见柯[2]

（1. 西北农林科技大学 理学院，陕西 杨凌 712100；
2. 中国传媒大学 理学院，北京 朝阳 100024）

摘　要　本文介绍了关于平方和乘积公式的经典 Hurwitz 定理，给出了比以往更加简洁的一个矩阵证明，并指出这个定理在几何中的一个有趣的应用.

关键词　Hurwitz 定理；Massey 定理；平方和；Euler 公式；向量积

中图分类号　O15　　**文献标识码**　A　　**文章编号**　1008-1399(2018)01-0024-04

A Matrix Approach to Hurwitz Theorem

LIN Kailiang[1] and CHEN Jianke[2]

(1. Department of Mathematics. Northwest A&T University, Yanglin 712100, PRC;
2. Department of Mathematics, Communication University of China, Beijing 100024, PRC)

Abstract　This paper introduces a matrix approach to the classical Hurwitz theorem on product of sums of squares. An interesting application in geometry is included.

Keywords　Hurwitz theorem, Massey's theorem, sum of squares, Euler's formula, vector product

1　介　绍

1898 年，德国数学家 Adolf Hurwitz 得到下述优美定理（定理 1）；本文将对此给出一个简单的矩阵论证明. 一个流行于教科书的证明是用抽象代数的方式，例如 Jacobson[4]. 本文将要给出的证明，更接近于 Hurwitz 的原始证明，它完全可以作为线性代数的一个美妙应用介绍给大一学生.

定理 1　设 $n\geqslant 1$. 则存在 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 的实双线性函数 $z_1,\cdots,z_n$ 使得

$$(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)=z_1^2+\cdots z_n^2 \tag{1}$$

对一切实变数 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 成立当且仅当 $n=1,2,4,8$.

收稿日期：2016-12-23　　**修改日期**：2017-12-06
基金项目：中国传媒大学理工科规划项目（2017XNG1726）
作者简介：林开亮（1983－），男，湖南常德人，博士，讲师，从事数学史和数学文化的研究与普及，Email：kailiang_lin@163.com.
陈见柯（1986－），男，山东兖州人，博士，讲师，研究方向代数几何，Email：jkchen003@126.com.

在 Hurwitz 之前，数学家们已经注意到，对 $n=1,2,4,8$，存在以下引人注目的平方和乘积公式：

对 $n=1$，等式平凡，即

$$x_1^2y_1^2=(x_1y_1)^2.$$

对 $n=2$，我们有

$$(x_1^2+x_2^2)(y_1^2+y_2^2)=(x_1y_1-x_2y_2)^2+(x_1y_2+x_2y_1)^2. \tag{2}$$

这不过是复数模长乘积法则的实形式，这个等式至少已经为公元 3 世纪的 Diophantus 所熟知.

类似地，由四元数范数的可乘性质可以得到 Euler 于 1748 年发现的四平方和乘积公式：

$$(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2)=\\ z_1^2+z_2^2+z_3^2+z_4^2, \tag{3}$$

其中 z_1,z_2,z_3,z_4 分别为

$$z_1=x_1y_1-x_2y_2-x_3y_3-x_4y_4,$$
$$z_2=x_1y_2+x_2y_1+x_3y_4-x_4y_3,$$
$$z_3=x_1y_3-x_2y_4+x_3y_1+x_4y_2,$$
$$z_4=x_1y_4+x_2y_3-x_3y_2+x_4y_1.$$

注　Euler 是在尝试证明数论中著名的四平方

和定理(每个正整数均为 4 个完全平方数的和,它最早由 Fermat 猜测,后来被 Lagrange 在 Euler 的基础上证明,并为 Euler 进一步简化)时,得到上述乘积公式. Gauss 从 Euler 的四平方和乘积公式已经预见到四元数(但未发表).

在 $n=8$ 时,存在一个类似的平方和乘积公式,例如,见 Jacobson[4],它最早由丹麦数学家 F. Degen 在 1818 年发现. A. Cayley 后来指出,这个公式可以从八元数(又称 Cayley 数)的范数乘法公式给出. Cayley 还进一步考虑了这样的问题,对于 $n=16$ 或者更一般的 $n=2^k$,是否存在类似的平方和乘积公式?或者,更一般地,我们可以提出下述问题:

Hurwitz 问题　对于哪些自然数 n,存在 n 个变数的平方和乘积公式?

1898 年,Hurwitz[3]对这一问题给出一个漂亮解答,这就是上述定理 1. Hurwitz 的证明是如此简单优美,以至于 20 多年以后,美国代数与数论学家 L. E. Dickson 专门撰文[2]介绍了 Hurwitz 的证明,这一经 Dickson 修改后的证明被 Curtis 收录于其线性代数教科书[1]中. 笔者这里将通过借助于 O. Veblen 和 J. von Neumann [9]中的一个引理进一步简化 Hurwitz 的原始证明.

2　Hurwitz 问题的矩阵约化

Hurwitz 解决上述平方和问题的关键在于将这一问题转化为一个矩阵问题. 我们现在就遵循 Hurwitz 的思路完成这个转化.

假定存在 n 个变数的平方和公式,设 z_i 作为 x 和 y 的实双线性函数表达如下:

$$z_i=\sum_{h,j}b_{ihj}x_hy_j,\ (i,h,j=1,2,\cdots,n),\tag{4}$$

其中 $b_{ihj}\in\mathbb{R}$. 将 z_i 以向量形式表示,即 $z_i=\boldsymbol{T}_i\boldsymbol{y}$,其中 $\boldsymbol{T}_i=(\sum_{h=1}^{n}b_{ihj}x_h)$ 为 $1\times n$ 矩阵. 记 $\boldsymbol{T}=(\boldsymbol{T}_1{}',\boldsymbol{T}_2{}'\cdots,\boldsymbol{T}_n{}')'$,这样 $z=\boldsymbol{Ty}$,进而

$$\sum_{i=1}^{n}z_i^2=z'z=(\boldsymbol{Ty})'(\boldsymbol{Ty})=\boldsymbol{y}'\boldsymbol{T}'\boldsymbol{Ty}$$

但另一方面,根据假设,我们还有:$(\sum_{i=1}^{n}x_i^2)(\sum_{i=1}^{n}y_i^2)=\boldsymbol{x}'\boldsymbol{x}\ \ \boldsymbol{y}'\boldsymbol{y}=\boldsymbol{y}'(\boldsymbol{x}'\boldsymbol{x})\boldsymbol{y}$. 因此 $\boldsymbol{T}'\ \ \boldsymbol{T}=\boldsymbol{x}'\boldsymbol{x}\boldsymbol{I}$,其中 $\boldsymbol{I}$ 为单位矩阵. 这就提示我们将不妨矩阵 $\boldsymbol{T}$ 表示为 n 个矩阵的线性组合. 不妨令

$\boldsymbol{T}=\boldsymbol{B}_1x_1+\boldsymbol{B}_2x_2+\cdots+\boldsymbol{B}_nx_n$,推导可得

$$\begin{aligned}(\sum_{k=1}^{n}\boldsymbol{B}_kx_k)'(\sum_{h=1}^{n}\boldsymbol{B}_hx_h)&=\sum_{k,h=1}^{n}\boldsymbol{B}'_k\boldsymbol{B}_hx_kx_h\\&=\sum_{h=1}^{n}\boldsymbol{B}'_h\boldsymbol{B}_hx_h^2+\sum_{h<k}(\boldsymbol{B}'_k\boldsymbol{B}_h+\boldsymbol{B}'_h\boldsymbol{B}_k)x_hx_k\\&=(\sum_{i=1}^{n}x_i^2)\boldsymbol{I}\end{aligned}$$

事实上,直接计算可知 $\boldsymbol{B}_h=(b_{ihj})$. 这样就得到原始的 Hurwitz 矩阵方程(其中,δ_{hk} 是 Kronecker 记号):

$$\boldsymbol{B}'_k\boldsymbol{B}_h+\boldsymbol{B}'_h\boldsymbol{B}_k=2\delta_{hk}\boldsymbol{I}\tag{5}$$

特别地,对于 $h=1,\cdots,n$,我们有 $\boldsymbol{B}'_h\boldsymbol{B}_h=\boldsymbol{I}$. 令 $\boldsymbol{A}_h=\boldsymbol{B}_n{}'\boldsymbol{B}_h$,这里 $h=1,\cdots,n-1$,容易验证 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_{n-1}$ 满足

$$\boldsymbol{A}_i\boldsymbol{A}_j+\boldsymbol{A}_j\boldsymbol{A}_i=-2\delta_{ij}\boldsymbol{I};\quad \boldsymbol{A}'_i=-\boldsymbol{A}_i\tag{6}$$

反之,如果 $M(n,\ \mathbb{R})$ 中存在 $n-1$ 个矩阵 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_{n-1}$ 满足(6),则容易验证这 $n-1$ 个矩阵和单位阵 $\boldsymbol{I}$ 一起满足(5). 所以 Hurwitz 问题就归结为讨论 $M(n,\mathbb{R})$ 中是否存在 $n-1$ 个矩阵 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_{n-1}$ 满足(6)?对此,Hurwitz 证明了下述结果:

定理 2　$M(n,\mathbb{R})$ 中存在 $n-1$ 个矩阵 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_{n-1}$ 满足(6)当且仅当 $n=1,2,4,8$.

Hurwitz 由此推出定理 1,这就解决了平方和乘积公式问题.

注　Hurwitz 在后来的文章中进一步考虑了这样的问题:设矩阵 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\in M(n,\ \mathbb{R})$ 满足(6),则 k 的最大值 $K(n)$ 是多少?对于这一问题,也有一个比较简单的解答,有兴趣的读者可以参考[5]或[6].

3　解的基本性质

为了证明定理 2,我们需要讨论 Hurwitz 矩阵方程的一些基本性质. 于本文而言,最关键的性质如下:

引理　设 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\in M(n,\mathbb{R})$ 满足

$$\boldsymbol{A}_i\boldsymbol{A}_j+\boldsymbol{A}_j\boldsymbol{A}_i=-2\delta_{ij}\boldsymbol{I},\tag{7}$$

则下述 2^{k-1} 个矩阵

$$\boldsymbol{I},\quad \boldsymbol{A}_i(1\leqslant i\leqslant k-1),$$

$$\boldsymbol{A}_{i_1}\boldsymbol{A}_{i_2}(1\leqslant i_1<i_2\leqslant k-1),\ \cdots,\ \boldsymbol{A}_1\cdots\boldsymbol{A}_{k-1}\tag{8}$$

线性无关.

证明　我们先来证明下述观察:

$\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k$ 中任意取 l 个($1\leqslant l\leqslant k-1$)不同矩阵作成的乘积 $\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l}$ 的迹为零.

事实上,若 l 为奇数,则对任意一个不同于 $i_1,\cdots,i_l$

的指标 j 有

$$\boldsymbol{A}_j(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l})\boldsymbol{A}_j^{-1}=(-1)^l(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l})\boldsymbol{A}_j\boldsymbol{A}_j^{-1}=-(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l}),$$

两边取迹就得到

$$\operatorname{tr}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{j_l})=-\operatorname{tr}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{j_l}),$$

从而 $\operatorname{tr}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{j_l})=0$.

若 l 为偶数,则对于 $i_1,\cdots,i_l$ 中的任意一个指标,不妨取为 i_1,有

$$\boldsymbol{A}_{i_1}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l})\boldsymbol{A}_{i_1}^{-1}=(-1)^{l-1}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l})\boldsymbol{A}_{i_1}\boldsymbol{A}_{i_1}{}^{-1}=-(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l}),$$

两边取迹就得到 $\operatorname{tr}(\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{j_l})=0$.

现在我们来证明(8)中的矩阵线性无关.假定存在实数 $c^0,c^i,c^{i_1i_2},\cdots,c^{12\cdots(k-1)}$ 使得

$$c^0\boldsymbol{I}+c^i\boldsymbol{A}_i+c^{i_1i_2}\boldsymbol{A}_{i_1}\boldsymbol{A}_{i_2}+\cdots+c^{i_1i_2\cdots i_l}\boldsymbol{A}_{i_1}\cdots\boldsymbol{A}_{i_l}+\cdots+c^{12\cdots(k-1)}\boldsymbol{A}_1\cdots\boldsymbol{A}_{k-1}=0,$$

在上式两边同乘以 $\boldsymbol{A}_{i_l}\cdots\boldsymbol{A}_{i_1}$ 并取迹,我们最终得到

$$c^{i_1i_2\cdots i_l}=0.$$

证毕.

注　这里的引理取自[9],它代替了[1]p. 321-322的引理.

4　定理2的证明

现在我们可以给出定理2的证明,因为前面我们已经给出了 $n=1,2,4,8$ 时的平方和乘积公式,因此充分性不言而喻,下面仅需证明必要性.

证明　不妨设 $n\geqslant 2$. 若 $M(n,\mathbb{R})$ 中存在 $n-1$ 个矩阵 $A_1,\cdots,\boldsymbol{A}_{n-1}$ 满足. 由此,特别地,它们满足(7),由引理可知,$M(n,\mathbb{R})$ 中存在 2^{n-2} 个线性无关的矩阵,关于维数,我们有如下不等式:

$$2^{n-2}\leqslant n^2.$$

但另一方面,注意到奇数阶反对称矩阵不可逆,所以由条件

$$\boldsymbol{A}_i^2=-I,\boldsymbol{A}_i{}'=-\boldsymbol{A}_i$$

推出 n 一定为偶数.

然而,当 $n\geqslant 10$ 时,我们可以归纳证明出

$$(2^{n+1-2}=2\cdot 2^{n-2}>2n^2>(n+1)^2):$$

$$2^{n-2}>n^2$$

所以 n 只能等于 $2,4,6,8$ 之一. 下面我们只需要排除 $n=6$ 的情形.

反证法. 假设 $A_1,\cdots,A_5\in M(6,\mathbb{R})$ 满足(6),我们将证明下面16个反对称矩阵

$$\boldsymbol{A}_i(1\leqslant i\leqslant 5),\quad \boldsymbol{A}_i\boldsymbol{A}_j(1\leqslant i<j\leqslant 5),\quad \boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\boldsymbol{A}_5$$

线性无关. 但我们知道 $M(6,\mathbb{R})$ 中的反对称矩阵构成的子空间的维数为 $\frac{6\times 5}{2}=15$,这就得到了矛盾,从而否定了 $n=6$ 这种可能.

为了证明上面的16个矩阵线性无关,我们只需再次应用引理的证明中用到的那个观察:$\boldsymbol{A}_1,\cdots,\boldsymbol{A}_5$ 中取任意的不超过4个矩阵作成的乘积的迹等于零. 事实上,设存在实数 c^i,c^{ij},c^{12345} 使得

$$\sum_{i=1}^{6}c^i\boldsymbol{A}_i+\sum_{1\leqslant i<j\leqslant 6}c^{ij}\boldsymbol{A}_i\boldsymbol{A}_j+c^{12345}\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\boldsymbol{A}_5=\boldsymbol{0}\tag{9}$$

在等式两边同乘以 A_k 得到

$$\sum c^i\boldsymbol{A}_i\boldsymbol{A}_k+\sum c^{ij}\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k+c^{12345}\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\boldsymbol{A}_5\boldsymbol{A}_k=\boldsymbol{0}$$

两边取迹得到 $c^k=0$,代入(9)式,则

$$\sum_{1\leqslant i<j\leqslant 5}c^{ij}\boldsymbol{A}_i\boldsymbol{A}_j+c^{12345}\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\boldsymbol{A}_5=\boldsymbol{0}\tag{10}$$

在等式两端右乘 $A_kA_h(k<h)$,则有

$$\sum_{1\leqslant i<j\leqslant 5}c^{ij}\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k\boldsymbol{A}_h+c^{12345}\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\boldsymbol{A}_5\boldsymbol{A}_k\boldsymbol{A}_h=\boldsymbol{0}$$

两边取迹得到 $c^{kh}=0$,代入(10)式,即可得到 $c^{12345}=0$,故此16个矩阵线性无关. 证毕.

5　推广到一般的域

事实上,这里我们并不需要将数域限制在实数域 $\mathbb{R}$,整个推理对复数域甚至是任意的特征不等于2的域 F 都适用. 换言之,本文的方法可以证明下述结论:

定理3　设 F 是一个特征不等于2的域,则 $M(n,F)$ 中存 $n-1$ 个矩阵 $A_1,\cdots A_{n-1}$ 满足(6)当且仅当 $n=1,2,4,8$.

由此可得一般的Hurwitz定理(我们有理由相信,Dickson本人已经看出这样的结果):

定理4　设 F 是一个特征不等于2的域且 $n\geqslant 1$. 则存在 $x_1,\cdots,x_n$ 与 $y_1,\cdots,y_n$ 的双线性函数 $z_1,\cdots,z_n$ 使得

$$(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)=z_1^2+\cdots z_n^2$$

对一切 $x_1,\cdots,x_n,y_1,\cdots,y_n\in F$ 成立当且仅当 $n=1,2,4,8$.

6　在几何学中的一个应用

应该指出,Hurwitz矩阵方程与代数、几何、拓

扑中的许多问题密切相关，特别值得介绍的是 1983 年 Massey[7] 给出的下述关于欧几里得空间的向量积(交叉积) 的基本结果：

定理 5　n 维欧几里得空间 $\mathbb{R}$(数量积记为 $\langle x, y\rangle$) 中存在满足以下两个条件的双线性向量积 $\times:(x,y)\mapsto x\times y$

(i) $\langle \boldsymbol{x}\times \boldsymbol{y}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}\times \boldsymbol{y}, \boldsymbol{y}\rangle = 0$.

(ii) $\|\boldsymbol{x}\times\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2\ \|\boldsymbol{y}\|^2 - \langle \boldsymbol{x}, \boldsymbol{y}\rangle^2$

当且仅当 $n=1$ 或 $n=3$ 或 $n=7$.

定理 5　可以从定理 1 得到，具体推导留给有兴趣的读者，可参见 Prasolov[8]. 注意，在三维空间中，(ii) 即我们熟知的 Lagrange 等式. Massey 定理断言，本质上，除了一维的平凡情况，只有在三维空间和七维空间才存在有意义的向量积.

参考文献

[1]　C. W. Curtis. Linear Algebra, An Introduction Approach [M]. UTM, Springer, 1996.

[2]　L. E. Dickson. On quaternions and their generalization and the history of the eight square theorem[J]. Ann. of Math. ,1919,20(2):155-171.

[3]　A. Hurwitz. Quadratischen Formen von beliebig vielen Variablen[J]. Nachrichten Ges. der Wiss. Gottingen, 1898,309-316.

[4]　N. Jacobson. 基础代数(第一卷:第二分册)[M]. 北京：高等教育出版社，1987 年.

[5]　江上鷗. Hurwitz-Radon 问题——等价和约化是处理数学问题的一种基本方法[J]. 数学通报，1964 年 04 期.

[6]　林开亮，Hurwitz-Radon 矩阵方程[J]. 数学传播，2012,36(1):48-63.

[7]　W. S. Massey. Cross products of vectors in higher dimensional Euclidean spaces [J]. The American Mathematical Monthly, 1983, 90(10): 697-701.

[8]　V. V. Prasolov. Problems and Theorems in Linear Algebra [J]. Translation of Mathematical monographs, Vol. 134, American Mathematical Society, 1994.

[9]　O. Veblen and J. von Neumann. Geometry of Complex Domain. Princeton Mimeographed Notes, Notes by W. Givens and A. H. Taub, Institute for Advanced Study, 1935-1936.

简讯

法国数学家获 2017 年度阿贝尔奖

挪威科学与文学院宣布，将 2017 年度阿贝尔奖授予法国数学家伊夫・梅耶尔，以表彰他在小波数学理论发展方面发挥的关键作用。挪威科学与文学院在颁奖词中说，梅耶尔是小波分析理论现代化发展方面“有远见的领导者”。这一理论处于数学、信息技术和计算科学交叉的发展领域，他的研究成果使小波分析发展成为一种逻辑连贯、应用广泛的理论。小波分析作为一种较新的时频分析方法，是当前应用数学和工程学科中一个迅速发展的新领域，广泛应用于计算调和分析、信号分析、数据压缩、医学成像、数字电影、计算机分类与识别以及地震勘探数据处理等领域。去年美国激光干涉仪引力波观测台探测到两个黑洞碰撞所发出的引力波，其中也应用到小波分析。

梅耶尔生于 1939 年，以一名法国公民的身份在北非的突尼斯度过了童年生活。他在 1957 年就读法国巴黎高等师范学校，并于 1966 年在法国斯特拉斯堡大学获得博士学位。他曾先后在巴黎南大学、巴黎综合理工大学、巴黎第九大学和卡尚高等师范学校任职。

阿贝尔是 19 世纪的挪威数学家，很多以他名字命名的发现已被载入教科书。2002 年在阿贝尔诞辰 200 周年时，挪威政府决定设立阿贝尔奖，意在弥补诺贝尔科学奖项中没有数学奖的遗憾。这项国际性大奖授予最杰出的数学家，奖金额为 600 万挪威克朗(约合 71 万美元)，从 2003 年起每年颁发一次。

(据新华社记者梁有昶、张淑惠)