光学中的矩阵方法导论

原著　A Gerrard
J.M.Burch

译编　哈流柱　闫吉祥

北京工业学院

1986.9

译编者的话

本书简称“矩阵光学”含义为用矩阵方法研究光学中若干与工程应用有关的可以线性化处理的问题，其范围甚广。本书第三章为阎吉祥所译、第四章为刘丽君、哈流柱所译、其余各章及六个附录为哈流柱所译。参加本书校对的有沙定国、王典民、梁佳沂、唐赵英等，最后全书译编责任校对是哈流柱。由于译编者水平低，肯定错误不少，敬请读者指正。

目　　录

第一章　矩阵运算导论

1-1　导言

在本书中我们介绍矩阵代数某些简单的概念，这些概念可有效地应用于包括光学成象和偏振方面一类问题。本书中讨论的内容主要针对不太熟悉矩阵和行列式的那些读者。分析是基本的并且仅包括对于理解本书其余部分是必不可少的内容。

矩阵是在1857年由数学家cayℓey 提出的，它是用于表示一个完整的线性联立方程组一种方便又简短的记号。矩阵的运算规则与通常数的运算规则稍有不同，然而这些规则很快就被人们发现并加以发展。在1920年Heisenborg 在量子力学中引入了矩阵之后，物理学家对于矩阵方法产生了极大的兴趣，尽管矩阵已被使用在多种工程计算中，但它在光学中的应用还是较新近的事。

关于行列式，我们涉及范围较窄，行列式早年即1771年由vand-ermonde 所提出。先被称之为"消元法"，因为行列式起因于解方程中使用逐一消元的方法。在我们所接触的大多数光学问题中，这些行列式值都是单位值，这样便于检验计算的结果。

让我们现在就矩阵起因引出某些符号。假设有一对线性方程组。

$$U = Ax + By$$

$$V = Cx + Dy$$

在这里A，B，C，D均是已知常数，x和y为变量。如果x和y赋于已知值，则可由这两方程计算出U和V。为了很多用途，我们把常数从变量中分出，于是我们把这一对方程记为：

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

这单个方程与原来一对方程等效，被赋于同样情确的含义而被定义之。我们把每一对方括号所包含的一组字符看为单个实体，称之为矩阵。$\begin{bmatrix} U \\ V \end{bmatrix}$和$\begin{bmatrix} x \\ y \end{bmatrix}$称为列矩阵或“列矢量”，因为每一个列矩阵仅包括一列字符。

通常矩阵是矩形阵列，其字符排列成行与列。矩阵$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$有两行两列，称为二阶方阵。稍后我们将遇到“行矩阵”（有时称“行矢量”）象〔PQ〕，其中被称为元素的字符分列着，被水平地书写成一行。仅有一个元素的矩阵就是一个普通的数或标量。

如果我们令每个矩阵为单一字符示之，我们可把上面那一对方程用更简捷方式写出来，于是

$$C_2 = SC_1$$

这里 C_1 是列矩阵 $\begin{bmatrix} x \\ y \end{bmatrix}$，$C_2$ 是列矩阵 $\begin{bmatrix} U \\ V \end{bmatrix}$，S 记为方阵 $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

现在让我们假设，U 和 V 依次转化为另一对方程变量，就是说 L 和 M 由另一对线性方程所确定，于是

$$L = PU + QV$$

$$M = RU + TV$$

在这里我们写成以下形式

$$\begin{bmatrix} L \\ M \end{bmatrix} = \begin{bmatrix} P & Q \\ R & T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix}$$

即

$$C_3 = KC_2$$

在这里 C_3 是 $\begin{bmatrix} L \\ M \end{bmatrix}$，K 是 $\begin{bmatrix} P & Q \\ R & T \end{bmatrix}$。因此我们可将 x，y 代入方程式中置换 U 和 V，从而确定 L 和 M。于是

$$L = P(Ax+By) + Q(Cx+Dy)$$

$$M = R(Ax+By) + T(Cx+Dy)$$

也就是

$$L = (PA+QC)x + (PB+QD)y$$

$$M = (RA+TC)x + (RB+TD)y$$

我们记之为

$$\begin{pmatrix} L \\ M \end{pmatrix} = \begin{pmatrix} PA+QC & PB+QD \\ RA+TC & RB+TD \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

即

$$C_3 = FC_1$$

在这里 F 是 $\begin{pmatrix} PA+QC & PB+QD \\ RA+TC & RB+TD \end{pmatrix}$。但是，另一方面，我们可写出：

$$C_3 = KC_2 = K(SC_1)$$

如果这是普通代数中的一个方程，现在我们可把它写成。

$$C_3 = KSC_1 = (KS)C_1$$

括号的位置稍有变化。KS 称为 K 和 S 的乘积。再者，比较 C_1 和 C_2 的联系方程，我们仍可写成，

$$C_3 = KSC_1$$

和 $$C_3 = FC_1$$

因此

$$F = KS$$

我们说 F 是 K 和 S 的“乘积”。

在矩阵中我们希望导出一个类似的办法，但是我们现在先须定义两个矩阵的乘积，因为仅仅简单的数乘在普通代数中早已被定义。

1-2 矩阵乘积

我们定义矩阵乘法。使得我们能够正式地越过普通代数的界线进达矩阵代数。

我们定义矩阵乘积是说K乘以S，从而给出了矩阵乘积F，也就是说

$$\begin{bmatrix} P & Q \\ R & T \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}=\begin{bmatrix} PA+QC & PB+QD \\ RA+TC & RB+TD \end{bmatrix}$$

考察右边矩阵的结构（这是积），很容易看出它的构成过程。

右上角元素是位在第一行第一列的。它是由K的第一行即〔P　Q〕和S的第一列即 $\begin{matrix} A \\ C \end{matrix}$，分别把它们相乘（行的第一个元素乘以列的第一个元素，行的第二个元素乘以列的第二个元素）形成乘积PA和QC，并将它们加在一起得到PA+QC。

F的第一行第二列元素也是用K的第一行和S的第二列元素以同样方法相乘后加在一起而得。F的第二行第一列元素是用K的第二行和S的第一列同法相乘后相加而得。最后F的第二行第二列元素用K的第二行和S的第二列同法相乘后相加而得。

对于矩阵的诸元素使用下标符号，在某些应用中确实有用。例如我们写一个列矩阵A

$$A=\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

这下标指示它在这一列元素中的位置。一个方阵S，我们可以写成

$$S=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

在这里元素的第一个下标指示该元素行位，第二个下标指示该元素的列位。如果加上下标记号来表示两个方阵K和S，就变成

$$K=\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \text{ 和 } S=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

于是乘积 $F=K\cdot S$ 变为

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \begin{bmatrix} K_{11}S_{11}+K_{12}S_{21} & K_{11}S_{12}+K_{12}S_{22} \\ K_{21}S_{11}+K_{22}S_{21} & K_{21}S_{12}+K_{22}S_{22} \end{bmatrix}$$

也就是

$$\begin{bmatrix} \sum_{i=1}^{2} K_{1i}S_{i1} & \sum_{i=1}^{2} K_{1i}S_{i2} \\ \sum_{i=1}^{2} K_{2i}S_{i1} & \sum_{i=1}^{2} K_{2i}S_{i2} \end{bmatrix}$$

这暗示出矩阵任一元素的一般公式

$$F_{RT}=\sum_{i=1}^{max} K_{Ri}S_{iT}$$

在这里 F_{RT} 是 F 的第 R 行第 T 列元素，这同样适合于 K 和 S 矩阵。（在这里使用和式符号应注意，被重复的下标 i 取下标中所有可能的值，它有时被省略。）

迄今为止，我们只限于讨论 2×2 矩阵和 2×1 列阵，但是矩阵的概念远超出这些。在本书中将用到 2×2，3×3，4×4 方阵和 2×1，3×1 和 4×1 列阵，以及 1×2，1×3，1×4 行阵。所有这些，均以同样的方式加以定义。例如，如果我们有一组四元方程，在四个未知量 B_1，B_2，B_3 和 B_4 中

$$B_1 = K_{11}A_1 + K_{12}A_2 + K_{13}A_3 + K_{14}A_4$$

$$B_2 = K_{21}A_1 + K_{22}A_2 + K_{23}A_3 + K_{24}A_4$$

$$B_3 = K_{31}A_1 + K_{32}A_2 + K_{33}A_3 + K_{34}A_4$$

$$B_4 = K_{41}A_1 + K_{42}A_2 + K_{43}A_3 + K_{44}A_4$$

我们把它们表示为矩阵形式

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & K_{24} \\ K_{31} & K_{32} & K_{33} & K_{34} \\ K_{41} & K_{42} & K_{43} & K_{44} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}$$

或者缩写形式

$$B = KA$$

在这种情况下，我们可以计算出列矩阵B的第R个元素，所用公式是

$$B_{R_1} = \sum_{i=1}^{4} K_{Ri} A_{i_1}$$

注意，构成乘积矩阵元素的法则仍适用于在S位置的A和在F位置的B，但是因为A和B现在是4×1列阵，学生应通过多做例题，完全熟悉它是十分重要的。

为说明清楚提供下列例子。

I 2.1

如果

$$A = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 6 \\ 1 & -4 \end{bmatrix}$$

则

$$AB = \begin{bmatrix} 1\times2+3\times1 & 1\times6+3\times(-4) \\ 5\times2+7\times1 & 5\times6+7\times(-4) \end{bmatrix}$$

$$= \begin{bmatrix} 2+3 & 6+(-12) \\ 10+7 & 30+(-28) \end{bmatrix}$$

$$= \begin{bmatrix} 5 & -6 \\ 17 & 2 \end{bmatrix}$$

然而

$$BA=\begin{bmatrix} 2\times1+6\times5 & 2\times3+6\times7 \\ 1\times1+(-4)\times5 & 1\times3+(-4)\times7 \end{bmatrix}$$

$$=\begin{bmatrix} 2+30 & 6+42 \\ 1+(-20) & 3+(-28) \end{bmatrix}$$

$$=\begin{bmatrix} 32 & 48 \\ -19 & -25 \end{bmatrix}$$

读者将注意到AB与BA 绝然不等。在矩阵乘法中各因子的次序必须保持：AB＝BA是不成立的。在矩阵代数学中矩阵乘法是不可交换的。

I 2.2

如果

$$C=\begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 6 \\ 1 & 3 & 4 \end{bmatrix} \qquad D=\begin{bmatrix} -5 \\ 3 \\ 1 \end{bmatrix}$$

$$CD=\begin{bmatrix} 3\times(-5)+1\times3+4\times1 \\ 2\times(-5)+1\times3+6\times1 \\ 1\times(-5)+3\times3+4\times1 \end{bmatrix}$$

$$=\begin{bmatrix} -15+3+4 \\ -10+3+6 \\ -5+9+4 \end{bmatrix}$$

$$=\begin{bmatrix} -8 \\ -1 \\ 8 \end{bmatrix}$$

如果我们求乘积DC，则需用D的第一行诸元素乘以C的第一列诸元

素。在这里D的第一行只有一个元素〔-5〕. 而C的第一列有三个元素 $\begin{bmatrix}3\\2\\1\end{bmatrix}$，因此乘法不能进行，乘积不存在。如果前乘矩阵C的列数与后乘矩阵D的行数相等，我们便可以把两个矩阵C和D乘起来。于是C和D被称为乘法相容，并且这个乘积CD是存在的。

I 2. 3

令

$$E = [3 \quad 1 \quad 4], \quad F = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 4 & -3 \\ 6 & 1 & 3 \end{bmatrix}$$

于是

$$\begin{aligned} EF &= [3\times1+1\times2+4\times6 \quad 3\times5+1\times4+4\times1 \quad 3\times9+1\times(-3)+4\times3] \\ &= [3+2+24 \quad 15+4+4 \quad 27-3+12] \\ &= [29 \quad 23 \quad 36] \end{aligned}$$

如果颠倒相乘的次序，FE不能存在——在这里F有三列，而E只有一行，FE不相容。

I 2. 4

令

$$H = [3 \quad 1 \quad 6], \quad K = \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix}$$

如果H作为前乘项，K作为后乘项，由于H的列数与K的行数均为3，则HK相容。于是

$$HK = [(3\times2)+(1\times4)+(6\times7)] = [6+4+42] = [52]$$

这是一个通常的纯数。

现在考虑如果K前乘H。这两个矩阵乘法相容，因为K只有一列，而只有一行。于是

$$KH = \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix} \begin{bmatrix} 3 & 1 & 6 \end{bmatrix} = \begin{bmatrix} 2\times3 & 2\times1 & 2\times6 \\ 4\times3 & 4\times1 & 4\times6 \\ 7\times3 & 7\times1 & 7\times6 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 2 & 12 \\ 12 & 4 & 24 \\ 21 & 7 & 42 \end{bmatrix}$$

对于这一对矩阵，HK是一纯数，而KH是一个3×3的方阵。

1－3　零矩阵

如果　$L = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$，$M = \begin{bmatrix} 7 & 1 \\ 4 & 2 \end{bmatrix}$

（或者，更确切地说M是任意二阶方阵），于是LM 和 ML 两者都等于 $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$。

矩阵 $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 称为“二阶零矩阵”，零矩阵通常记为符号0．并且在普通代数中取零的位置。

任何方形或矩形矩阵具有零矩阵的形式，就是说矩阵所有元素皆为零。零矩阵前乘或后乘任何矩阵，其矩阵乘积是一个零矩阵，这是一个法则。

I．4　单位矩阵

I．4．1

如果　$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$，　$Q = \begin{bmatrix} 3 & 5 \\ 2 & 6 \end{bmatrix}$

于是

$$PQ = QP = Q = \begin{bmatrix} 3 & 5 \\ 2 & 6 \end{bmatrix}$$

矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 具有下列性质，如果用它前乘或后乘任何一个二阶矩阵，则此二阶矩阵值不变。我们称 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 为二阶单位矩阵。

单位矩阵通常记为 I，它的另一些例子是 3×3 的矩阵 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，4×4 的矩阵 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ 等。

n 阶单位矩阵具有 n 行 n 列。除了主对角线上的元素之外，其余所有元素均为零，所谓主对角线是指从左上角至右下角的联线，这主对角线上的所有元素均为 1。

I-5 对角矩阵

单位矩阵是“对角矩阵”的特殊情况，对角矩阵一定是一个方阵，对于此方阵其非主对角线上的元素均为零，主对角线上元素可取任意值。例如 $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 和 $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$ 就是对角矩阵。

如果两个对角矩阵相乘在一起，乘积的阶数不变，其乘积非常简单并且所得的矩阵仍是一个对角矩阵。

I. 5. 1

如果

$$A = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \qquad b = \begin{bmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix}$$

则

$$AB=BA=\begin{bmatrix} a_1b_1 & 0 & 0 \\ 0 & a_2b_2 & 0 \\ 0 & 0 & a_3b_3 \end{bmatrix}$$

（读者注意，在此例中对于矩阵诸元素我们使用代数字母代替算术中的具体数。正如普通代数一样，我们可以使用代数字母记为一个未知的值或任意的数。）

I－6　矩阵连乘

如果我们求三个矩阵L，M和N的乘积，我们可采取以下两条途径：

(a)先求乘积(MN)，然后用L前乘所得的积。

(b)亦可以先求(LM)，然后用N后乘所得的积。

可以证明，只要矩阵相乘的次序保持不变，这两种方法均给出同样的结果。于是$L(MN)=(LM)N$。我们认为LMN的两种乘法与普通代数完全相同。

I.6.1

令

$$L=\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix},\quad M=\begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}$$ 而

$$N=\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

那么

$$L(MN)=\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 9 & 7 \\ 13 & 9 \end{bmatrix}$$

$$=\begin{bmatrix} 48 & 34 \\ 62 & 46 \end{bmatrix}=LMN$$

并且

$$(LM)N=\begin{pmatrix}11 & 4\\ 14 & 6\end{pmatrix}\begin{pmatrix}4 & 2\\ 1 & 3\end{pmatrix}=\begin{pmatrix}48 & 34\\ 62 & 46\end{pmatrix}=LMN$$

如上所述。

于是，虽然矩阵乘法不满足交换律，但满足“结合律”。推广至四阶或更高阶矩阵相乘均成立，$PQRS=P(QR)S=(PQR)S$ 等。

I-7 矩阵加减法

若两个矩阵 M 和 N 具有相同数目的行和相同数目的列，把两矩阵相应元素各自相加或相减就得到两矩阵的和或差。如果 $P=M+N$，于是 $P_{jk}=M_{jk}+N_{jk}$。因为零矩阵的所有元素为零，于是 $M\pm O=M$ 并且 $M-M=O$。

矩阵运算也服从“分配律”

$$A(B+C)=AB+AC$$

I.7.1 计算例

令 $A=\begin{pmatrix}2 & 0\\ 1 & 1\end{pmatrix}$，$B=\begin{pmatrix}3 & 0\\ 1 & 2\end{pmatrix}$ 和 $C=\begin{pmatrix}-1 & 1\\ 2 & 0\end{pmatrix}$，那么

$$A(B+C)=\begin{pmatrix}2 & 0\\ 1 & 1\end{pmatrix}\begin{pmatrix}2 & 1\\ 3 & 2\end{pmatrix}=\begin{pmatrix}4 & 2\\ 5 & 3\end{pmatrix}$$

而

$$AB+AC=\begin{pmatrix}6 & 0\\ 4 & 2\end{pmatrix}+\begin{pmatrix}-2 & 2\\ 1 & 1\end{pmatrix}=\begin{pmatrix}4 & 2\\ 5 & 3\end{pmatrix}$$

正如前述，如果把某一矩阵自身加 λ 次，很清楚，每一矩阵元素则要用一个普通的数（或者“标量”）λ 去乘。这运算有时称为矩阵的“标量积”。使用矩阵乘法即如果用一个对角矩阵 λI 去前乘或后乘，也可以获得同样的答案，这个对角矩阵 λI 的主对角线诸元素均为 λ。

现在我们已有了矩阵相乘，相加，或相减的法则。一个矩阵被别的矩