矩阵李群李代数

耿修瑞

中国科学院空天信息创新研究院

gengxr@sina.com.cn

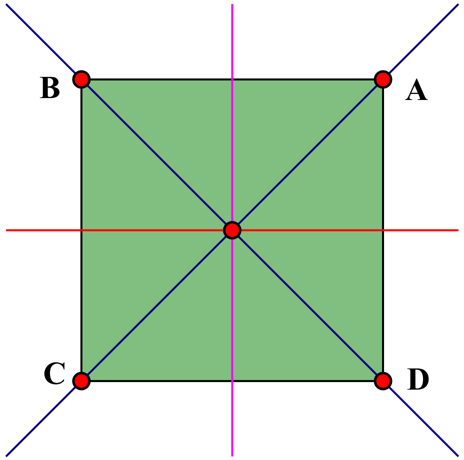
2025.6

提纲

- 群
- 置换群
- 矩阵李群
- 李群
- 矩阵指数
- 矩阵李代数
- 李代数
- 矩阵李群同态定理
- 矩阵李群的应用

群

群的概念源自于对自然界中的对称性



$$\mathbf{I}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\quad \mathbf{Q}_1=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\quad \mathbf{Q}_2=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\quad \mathbf{Q}_3=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$$

$$\mathbf{R}_1=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\quad \mathbf{R}_2=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\quad \mathbf{R}_3=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\quad \mathbf{R}_4=\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}$$

群

定义1（**群**）一个非空集合G上定义了一个二元运算$*$，满足：

（1）封闭性：对于 G 中任意的元素 g_1, g_2，总有 $g_1 * g_2 \in G$ 。

（2）结合律：对于 G 中任意的元素 g_1, g_2, g_3，都有 $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$ 。

（3）单位元的存在性：对于G中的任意元素 g，总存在一个元素$e \in G$，使得 $g * e = e * g = g$。

（4）逆元的存在性：对于G中的任意元素 g，总存在一个元素$h \in G$，使得 $g * h = h * g = e$。

则称$(G, *)$为一个群。在不引起歧义的情况下，$(G, *)$也可以简称为G 。此外，为了方便起见，群乘积符号$*$也可以省略，即 $g_1 * g_2$ 可以记为$g_1 g_2$ 。

群

定义2（**阿贝尔群**）若群的运算满足交换律，即对于任意的$g,h\in G$，都有$gh=hg$，则称群G为阿贝尔群，也叫交换群。

定义3（**子群**）一个群 G 的子集 H，如果满足下列条件：

（1）单位元的存在性：H中包含G的单位元。

（2）逆元的存在性：如果 $h\in G$，则 $h^{-1}\in G$。

（3）封闭性：如果 $h_1,h_2\in G$，则 $h_1h_2\in G$。

则称H为G的子群。

定义4（**正规子群**）若H为G的子群，如果对于任意的$g\in G$，都有$gH=Hg$，则称H为G的正规子群。正规子群也称不变子群。

其中，$gH=\{gh\mid h\in H\}$称为子群H的一个左陪集，$Hg=\{hg\mid h\in H\}$称为子群H的一个右陪集。

群

$$\mathbf{I}=\begin{bmatrix}1&0\\0&1\end{bmatrix}\qquad \mathbf{Q}_1=\begin{bmatrix}0&-1\\1&0\end{bmatrix}\qquad \mathbf{Q}_2=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}\qquad \mathbf{Q}_3=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

$$\mathbf{R}_1=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\qquad \mathbf{R}_2=\begin{bmatrix}-1&0\\0&1\end{bmatrix}\qquad \mathbf{R}_3=\begin{bmatrix}0&1\\1&0\end{bmatrix}\qquad \mathbf{R}_4=\begin{bmatrix}0&-1\\-1&0\end{bmatrix}$$

其中，$G=\{\mathbf{I},\mathbf{Q}_1,\mathbf{Q}_2,\mathbf{Q}_3,\mathbf{R}_1,\mathbf{R}_2,\mathbf{R}_3,\mathbf{R}_4\}$ 构成群。该群有八个子群，分别是$G_1=\{\mathbf{I}\}$，$G_2=\{\mathbf{I},\mathbf{Q}_2\}$，$G_3=\{\mathbf{I},\mathbf{Q}_1,\mathbf{Q}_2,\mathbf{Q}_3\}$，$G_4=\{\mathbf{I},\mathbf{R}_1\}$，$G_5=\{\mathbf{I},\mathbf{R}_2\}$，$G_6=\{\mathbf{I},\mathbf{R}_3\}$，$G_7=\{\mathbf{I},\mathbf{R}_4\}$，$G_8=G=\{\mathbf{I},\mathbf{Q}_1,\mathbf{Q}_2,\mathbf{Q}_3,\mathbf{R}_1,\mathbf{R}_2,\mathbf{R}_3,\mathbf{R}_4\}$。其中 G_1 只包含 G 的单位元是 G 的最小子群。G_2,G_3 为 G 的两个旋转子群。G_4,G_5,G_6,G_7 是 G 的四个镜面反射子群。G_8 是 G 本身，是 G 的最大子群。G_2,G_3 均为 G 的非平凡正规子群。G_2,G_3,G_4,G_5,G_6,G_7 则为 G 的非平凡阿贝尔子群。

群

常用的矩阵群的例子

1. 一般线性群：对于任意一个正整数n，所有的$n\times n$可逆实矩阵的集合在矩阵乘法操作下构成一个群，称为一般线性群，记为$\mathrm{GL}(n;\mathbb{R})$。（请尝试验证）

2. 特殊线性群：所有行列式为1的$n\times n$可逆实矩阵的集合显然是$\mathrm{GL}(n;\mathbb{R})$的一个子群，该群称为特殊线性群，记为$\mathrm{SL}(n;\mathbb{R})$。（请尝试验证）

3. 正交群和特殊正交群：所有的$n\times n$实正交矩阵的集合，称为正交群，记为$\mathrm{O}(n)$，所有的行列式为1的$n\times n$实正交矩阵构成的集合，称为特殊正交群，记为$\mathrm{SO}(n)$。（请尝试验证）

4. 循环移位群：所有奇数阶循环移位矩阵的任意次幂构成的集合，称为循环移位群，记为$CS(n)=\left\{\mathbf{Q}_n^x \middle| x\in\mathbb{R}, n\in 2\mathbb{Z}+1\right\}$。（请尝试验证）

$$\mathbf{Q}_n=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

置换群

定义5（**置换群**）集合 $\{1,2,\cdots,n\}$ 上所有的一一到上的映射构成的集合在复合运算下构成一个群，该群称之为置换群，记为 S_n 。

假设 σ 是 S_n 中的一个元素，他将 $1,2,\cdots,n$ 映射为 $i_1,i_2,\cdots,i_n$ ，则 σ 可以表示为

$$\sigma=\begin{bmatrix}1 & 2 & \cdots & n\\ i_1 & i_2 & \cdots & i_n\end{bmatrix}$$

当 $n=3$ 时, S_n 包含6个元素，分别为

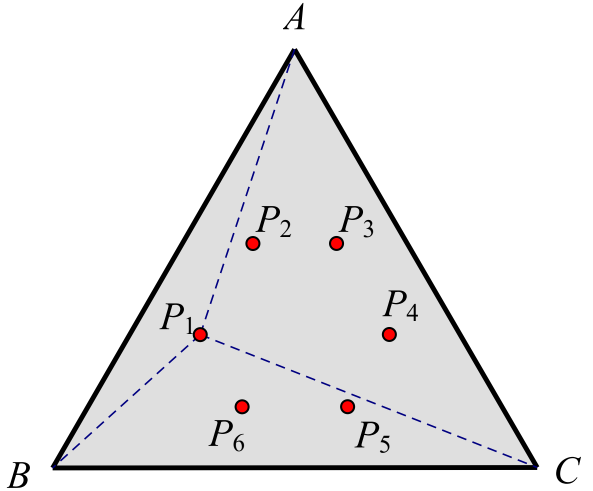
$$\sigma_1=\begin{bmatrix}1&2&3\\1&2&3\end{bmatrix}\sigma_2=\begin{bmatrix}1&2&3\\1&3&2\end{bmatrix}\sigma_3=\begin{bmatrix}1&2&3\\3&2&1\end{bmatrix}\sigma_4=\begin{bmatrix}1&2&3\\2&1&3\end{bmatrix}\sigma_5=\begin{bmatrix}1&2&3\\2&3&1\end{bmatrix}\sigma_6=\begin{bmatrix}1&2&3\\3&1&2\end{bmatrix}$$

也可以用矩阵表示出上述6个元素

$$\mathbf{I}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\quad \mathbf{R}_1=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\quad \mathbf{R}_2=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\quad \mathbf{R}_3=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\quad \mathbf{Q}_1=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}\quad \mathbf{Q}_2=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$$

置换群

置换群可以反映单形体内部各个点的重心坐标的对称性。



将恒等变换σ_1作用于$P_1:(p_{1A},p_{1B},p_{1C})$ ，仍得到$P_1:(p_{1A},p_{1B},p_{1C})$

将σ_2作用于$P_1:(p_{1A},p_{1B},p_{1C})$，可得到$P_4:(p_{1A},p_{1C},p_{1B})$

将σ_3作用于$P_1:(p_{1A},p_{1B},p_{1C})$，可得到$P_6:(p_{1C},p_{1B},p_{1A})$

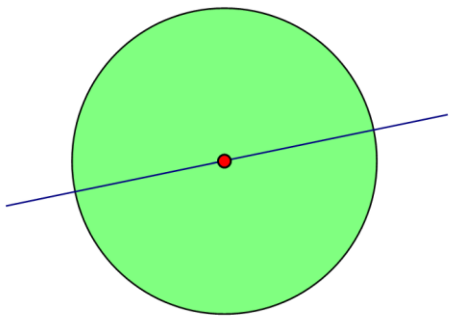
将σ_4作用于$P_1:(p_{1A},p_{1B},p_{1C})$，可得到$P_2:(p_{1B},p_{1A},p_{1C})$

将σ_5作用于$P_1:(p_{1A},p_{1B},p_{1C})$，可得到$P_3:(p_{1B},p_{1C},p_{1A})$

将σ_6作用于$P_1:(p_{1A},p_{1B},p_{1C})$，可得到$P_5:(p_{1C},p_{1A},p_{1B})$

矩阵李群

相对于正方形而言，圆具有更强的连续对称性。描述这种连续对称的图形，仅仅用群或者矩阵群的概念往往是不够的，很多时候需要引入李群或矩阵李群的概念



定义6（**矩阵李群**）任何一个一般线性群 $\mathrm{GL}(n;\mathbb{R})$ 的子群 H，如果满足如下性质：对于 H 中任意一个收敛于矩阵 $\mathbf{A}$ 的矩阵序列 $\{\mathbf{A}_n\}$ ，要么 $\mathbf{A}\in H$，要么 $\mathbf{A}$ 不可逆。则称 H 为一个矩阵李群。

注：上述定义相当于说，矩阵李群为一般线性群的封闭子群。

矩阵李群

常见的矩阵李群

1. 一般线性群$\mathrm{GL}(n;\mathbb{R})$是矩阵李群。

对于该群中的任意一个收敛于矩阵$\mathbf{A}$的矩阵序列$\{\mathbf{A}_n\}$，由于$\mathbf{A}_n \in \mathrm{GL}(n;\mathbb{R})$均为实矩阵，所以$\mathbf{A}$必为实矩阵。矩阵$\mathbf{A}$要么可逆，要么不可逆，无论哪种情况发生，$\mathrm{GL}(n;\mathbb{R})$均满足矩阵李群的定义。

2. 特殊线性群$\mathrm{SL}(n;\mathbb{R})$是矩阵李群。

假设$\mathrm{SL}(n;\mathbb{R})$中的矩阵序列$\{\mathbf{A}_n\}$收敛于矩阵$\mathbf{A}$，即$\lim\limits_{n\to\infty}\mathbf{A}_n=\mathbf{A}$。由于行列式运算的连续性，则有$\lim\limits_{n\to\infty}|\mathbf{A}_n|=\left|\lim\limits_{n\to\infty}\mathbf{A}_n\right|=|\mathbf{A}|$，又因为$|\mathbf{A}_n|$恒为1，所以$|\mathbf{A}|=1$。

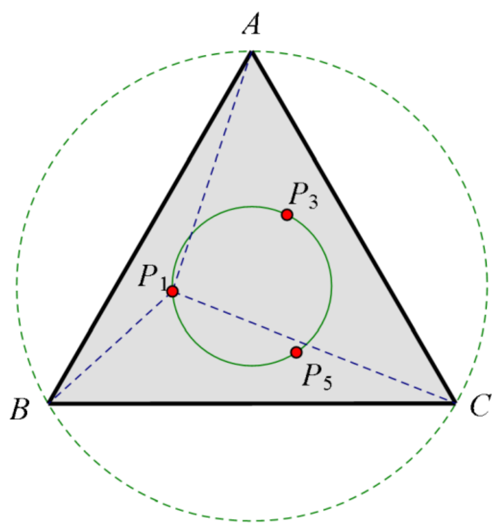
3. 正交群$\mathrm{O}(n)$和特殊正交群$\mathrm{SO}(n)$是矩阵李群。

假设$\{\mathbf{A}_n\}$为$\mathrm{O}(n)$中收敛于矩阵$\mathbf{A}$的矩阵序列，即$\lim\limits_{n\to\infty}\mathbf{A}_n=\mathbf{A}$。由于下式成立：$\lim\limits_{n\to\infty}\mathbf{A}_n^{\mathrm{T}}\mathbf{A}_n=\lim\limits_{n\to\infty}\mathbf{A}_n^{\mathrm{T}}\lim\limits_{n\to\infty}\mathbf{A}_n=\mathbf{A}^{\mathrm{T}}\mathbf{A}$，且$\mathbf{A}_n^{\mathrm{T}}\mathbf{A}_n=\mathbf{I}$恒成立，则必有$\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}$。即有$\mathbf{A}\in\mathrm{O}(n)$。因此，$\mathrm{O}(n)$为矩阵李群。同样地，由于极限操作不改变矩阵的正交性和行列式为1的性质，$\mathrm{SO}(n)$也是矩阵李群。

矩阵李群

4. 循环移位群 $CS(n)=\left\{\mathbf{Q}_n^x \middle| x\in\mathbb{R}, n\in 2\mathbb{Z}+1\right\}$是矩阵李群。

n 阶循环移位群可以体现$(n-1)$ 维单形体内各个顶点的重心坐标的连续对称性。下图以二维单形体（三角形）为例。



$$\mathbf{Q}^0\tilde{\mathbf{p}}_1=\tilde{\mathbf{p}}_1$$

$$\mathbf{Q}^1\tilde{\mathbf{p}}_1=\tilde{\mathbf{p}}_3$$

$$\mathbf{Q}^2\tilde{\mathbf{p}}_1=\tilde{\mathbf{p}}_5$$

$$\mathbf{Q}^x\tilde{\mathbf{p}}_1=?$$

$$\tilde{\mathbf{p}}_1=\begin{bmatrix}p_{1A} & p_{1B} & p_{1C}\end{bmatrix}^{\mathrm{T}} \quad \tilde{\mathbf{p}}_3=\begin{bmatrix}p_{1B} & p_{1C} & p_{1A}\end{bmatrix}^{\mathrm{T}} \quad \tilde{\mathbf{p}}_5=\begin{bmatrix}p_{1C} & p_{1A} & p_{1B}\end{bmatrix}^{\mathrm{T}}$$

绿色（椭）圆与循环移位群$CS(3)$对应。并且，这个圆也可以取到三角形之外。

李群

李群理论自诞生以来，对数学和物理的发展产生了重要的影响。尤其是，某些特殊的李群，竟直接对应了微观粒子世界的基本物理结构。

麦克斯韦方程对应了纤维丛（底空间是闵可夫斯基四维时空）上主丛（U(1)李群）之间的联络

杨米尔斯规范场对应了纤维丛（底空间仍是四维时空）上规范群（李群）之间的联络，规范群取决于具体的杨米尔斯理论，常见的如SU(3)李群，用于量子色动力学描述强相互作用；SU(2)*SU(1)李群，用于描述电磁相互作用和弱相互作用的统一理论。

李群

接下来我们讨论矩阵李群和李群的关系，这里首先给出李群的定义。

定义7（**李群**）李群 G 是微分流形同时也是一个群，且群乘积运算以及群的逆运算从流形映射角度是可微的。（其中群乘积运算可以看作从$G\times G$ 到流形 G 的映射，逆运算则可以看作从G 到 G的映射）

从上述定义可以看出，李群不仅仅是一个群，还是一个微分流形。一般来说，微分流形是嵌入在$\mathbb{R}^n$ 中的一个光滑结构，因此李群可以认为是嵌入在 $\mathbb{R}^n$中的某种具有对称性的光滑结构。此外，李群中的群乘积运算不限于矩阵乘积，而是一般的、抽象的从流形 $G\times G$到流形 G 的映射。而矩阵李群是一般线性群的封闭子群，是一堆 $n\times n$ 矩阵的集合。由于所有$n\times n$实矩阵的集合可以等同为 $\mathbb{R}^{n^2}$，因此，一般线性群继而每一个矩阵李群都是嵌入在 $\mathbb{R}^{n^2}$中的一个对称结构。那么矩阵李群和李群之间有什么关联么？

李群

比较矩阵李群和李群的定义，无论从所嵌入的空间的维度，还是从群运算，亦或从集合的结构而言，似乎都看不出二者有任何联系。但是下面的两个定理揭示了二者之间的内在关联。

定理1 每一个矩阵李群必然是李群。

定理2 每一个紧李群都同构于一个矩阵李群。

定理1相当于告诉我们，矩阵李群是李群的特例。但定理2却告诉我们，这个特例几乎就是全部，这是一个令人震惊的结论。也就是说，对于绝大多数抽象的李群，我们都可以用矩阵李群为之具象化，并且李群中的群运算都可以用矩阵乘积代替。毋庸置疑，这对很多问题的理解和处理带来了极大的方便。

矩阵指数

李代数是研究矩阵李群的必不可少的工具。一方面，矩阵李群是结构复杂的微分流形，在上面只有矩阵乘积而没有矩阵加法运算。而李代数是更容易理解的线性空间，线性代数中包含的各种工具包都可以在上面展开应用。另一方面，李代数包含了矩阵李群的大部分信息，尤其是李代数几乎可以提供矩阵李群的所有局部信息。矩阵指数是联系矩阵李群和其李代数的关键所在，因此，接下来首先对其进行简要介绍。

定义8（**矩阵指数**） 对于任意一个$n\times n$实矩阵$\mathbf{X}$，它的指数定义为

$$e^{\mathbf{X}}=\sum_{m=0}^{\infty}\frac{\mathbf{X}^m}{m!}$$

可以证明，对于任意的实方阵，上述定义的级数必然收敛。因此，矩阵指数对于任意的实方阵都是有意义的。

矩阵指数

关于矩阵指数，有如下重要结论：

定理3 假设$\mathbf{X}, \mathbf{Y}$是任意两个方阵，那么

（1）$e^{\mathbf{0}} = \mathbf{I}$

（2）$e^{\mathbf{X}}$ 可逆，且 $\left(e^{\mathbf{X}}\right)^{-1} = e^{-\mathbf{X}}$

（3）对于任意的实数α, β，都有 $e^{(\alpha+\beta)\mathbf{X}} = e^{\alpha\mathbf{X}} e^{\beta\mathbf{X}}$

（4）如果$\mathbf{XY} = \mathbf{YX}$，则 $e^{\mathbf{X}+\mathbf{Y}} = e^{\mathbf{X}} e^{\mathbf{Y}} = e^{\mathbf{Y}} e^{\mathbf{X}}$

（5）如果矩阵 $\mathbf{P}$ 可逆，则 $e^{\mathbf{PXP}^{-1}} = \mathbf{P} e^{\mathbf{X}} \mathbf{P}^{-1}$

定理4 对于任意一个实方阵$\mathbf{X}$，都有

$$\lim_{n\to\infty}\left(\mathbf{I} + \frac{\mathbf{X}}{n}\right)^n = e^{\mathbf{X}}$$

定理5 对于任意一个实方阵$\mathbf{X}$，都有（请尝试证明）

$$\det\left(e^{\mathbf{X}}\right) = e^{trace(\mathbf{X})}$$

矩阵指数

矩阵指数的计算，可以分为如下三种情形

（1）当方阵 $\mathbf{X}$ 可对角化时。

首先对其特征分解，$\mathbf{X}=\mathbf{PDP}^{-1}$，根据定理3的性质5，有

$$e^{\mathbf{PDP}^{-1}}=\mathbf{P}e^{\mathbf{D}}\mathbf{P}^{-1}$$

其中

$$\mathbf{D}=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix} \qquad e^{\mathbf{D}}=\begin{bmatrix}e^{\lambda_1} & & 0\\ & \ddots & \\ 0 & & e^{\lambda_n}\end{bmatrix}$$

（2）当方阵 $\mathbf{X}$ 为幂零矩阵时。

因为 $\mathbf{X}$ 是幂零矩阵，所以存在正整数m，使得对于任意一个等于或者大于m的正整数 l，都有 $\mathbf{X}^l=\mathbf{0}$，因此可以直接根据矩阵指数的定义计算得

$$e^{\mathbf{X}}=\sum_{k=0}^{m-1}\frac{\mathbf{X}^k}{k!}$$

矩阵指数

矩阵指数指数的计算，可以分为如下三种情形

（3）当方阵$\mathbf{X}$即不能对角化，也不是幂零矩阵时。

首先，根据若当标准型理论，可以对该矩阵分解为若当标准型，即

$$\mathbf{X}=\mathbf{PJP}^{-1}$$

其中$\mathbf{J}$为若当矩阵，显然可以将其分解为对角矩阵和次对角矩阵之和，即有$\mathbf{J}=\mathbf{D}+\widehat{\mathbf{D}}$，记$\mathbf{S}=\mathbf{PDP}^{-1}$，$\widehat{\mathbf{S}}=\mathbf{P}\widehat{\mathbf{D}}\mathbf{P}^{-1}$，显然他们分别为可对角矩阵和幂零矩阵。这样，我们就可以把任意矩阵分解为对角化矩阵和幂零矩阵之和，即

$$\mathbf{X}=\mathbf{S}+\widehat{\mathbf{S}}$$

注意到,$\mathbf{S},\widehat{\mathbf{S}}$是可交换的，即$\mathbf{S}\widehat{\mathbf{S}}=\widehat{\mathbf{S}}\mathbf{S}$，根据定理3的性质4，有

$$e^{\mathbf{X}}=e^{\mathbf{S}+\widehat{\mathbf{S}}}=e^{\mathbf{S}}e^{\widehat{\mathbf{S}}}$$

其中,$e^{\mathbf{S}}$和$e^{\widehat{\mathbf{S}}}$的计算可以归结为上面两种情形。

矩阵指数

例1 试计算如下矩阵的矩阵指数

$$\mathbf{X}=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$$

首先对矩阵进行特征分解，得到$\mathbf{X}=\mathbf{UDU}^{-1}$，其中

$$\mathbf{U}=\begin{bmatrix}\sqrt{3}/3&\sqrt{3}/3&\sqrt{3}/3\\\sqrt{3}/3&-\sqrt{3}/6+i/2&-\sqrt{3}/6-i/2\\\sqrt{3}/3&-\sqrt{3}/6-i/2&-\sqrt{3}/6+i/2\end{bmatrix}\qquad\mathbf{D}=\begin{bmatrix}1&0&0\\0&-1/2+i\sqrt{3}/2&0\\0&0&-1/2-i\sqrt{3}/2\end{bmatrix}$$

则

$$e^{\mathbf{D}}=\begin{bmatrix}e^{1}&0&0\\0&e^{-1/2+i\sqrt{3}/2}&0\\0&0&e^{-1/2-i\sqrt{3}/2}\end{bmatrix}=\begin{bmatrix}2.7183&0&0\\0&0.3929+0.4620i&0\\0&0&0.3929-0.4620i\end{bmatrix}$$

$$e^{\mathbf{X}}=\mathbf{U}\,\mathrm{e}^{\mathbf{D}}\,\mathbf{U}^{-1}=\begin{bmatrix}1.1681&1.0419&0.5084\\0.5084&1.1681&1.0419\\1.0419&0.5084&1.1681\end{bmatrix}\qquad e^{\mathbf{X}}\approx\left(1+\frac{\mathbf{X}}{10000}\right)^{10000}=\begin{bmatrix}1.1680&1.0418&0.5083\\0.5083&1.1680&1.0418\\1.0418&0.5083&1.1680\end{bmatrix}$$

矩阵指数

例2 试计算如下矩阵的矩阵指数

$$\mathbf{X}=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$$

该矩阵可以分解为对角矩阵和幂零矩阵的和，即

$$\mathbf{X}=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}+\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}=\mathbf{I}+\mathbf{X}_1$$

因为

$$e^{\mathbf{X}_1}=\sum_{k=0}^{2}\frac{\mathbf{X}_1^k}{k!}=\mathbf{I}+\mathbf{X}_1+\frac{\mathbf{X}_1^2}{2}=\begin{bmatrix}1&1&1/2\\0&1&1\\0&0&1\end{bmatrix}$$

所以

$$e^{\mathbf{X}}=e^{\mathbf{I}+\mathbf{X}_1}=e^{\mathbf{I}}e^{\mathbf{X}_1}=\begin{bmatrix}e&0&0\\0&e&0\\0&0&e\end{bmatrix}\begin{bmatrix}1&1&1/2\\0&1&1\\0&0&1\end{bmatrix}=\begin{bmatrix}e&e&e/2\\0&e&e\\0&0&e\end{bmatrix}$$

矩阵李代数

定义9（**矩阵李代数**）假设G是一个矩阵李群，则G的李代数（记为$\mathbf{g}$）是所有满足下列条件的矩阵$\mathbf{X}$的集合：对于任意的实数t，总有$e^{t\mathbf{X}} \in G$。

几何上，矩阵李群的李代数是该矩阵李群在单位元处的切空间。

定义10（**括号运算**）给定两个方阵$\mathbf{A},\mathbf{B}$，他们的括号运算定义为

$$[\mathbf{A},\mathbf{B}] = \mathbf{AB} - \mathbf{BA}$$

定理6 假设G是一个矩阵李群，$\mathbf{g}$是G的李代数，并且$\mathbf{X},\mathbf{Y}$均为$\mathbf{g}$中的元素，则

（1）对于任意实数s，都有$s\mathbf{X} \in \mathbf{g}$

（2）$\mathbf{X}+\mathbf{Y} \in \mathbf{g}$

（3）$\mathbf{XY}-\mathbf{YX} \in \mathbf{g}$

$$\boxed{e^{t(\mathbf{X}+\mathbf{Y})} = \lim_{m\to\infty}\left(e^{t\mathbf{X}/m}e^{t\mathbf{Y}/m}\right)^m}$$

$$\boxed{\left.\frac{d}{dt}\right|_{t=0}\left(e^{t\mathbf{X}}\mathbf{Y}e^{-t\mathbf{X}}\right) = \mathbf{XY}-\mathbf{YX}}$$

由定理6的性质1,2可知，矩阵李群的李代数必然是线性空间。由性质3可知，矩阵李群的李代数在括号运算下是封闭的。

矩阵李代数

常见的矩阵李群的李代数

1. 一般线性群 $\mathrm{GL}(n;\mathbb{R})$。

对于任意一个 $n\times n$ 实矩阵 $\mathbf{X}$，根据定理3，对于所有的实数 t，都有 $e^{t\mathbf{X}}$ 为实矩阵且可逆，即 $\mathrm{e}^{t\mathbf{X}}\in\mathrm{GL}(n;\mathbb{R})$。此外，如果对于任意的实数 t，都有 $\mathrm{e}^{t\mathbf{X}}\in\mathrm{GL}(n;\mathbb{R})$，那么 $\mathbf{X}=\left.\frac{\mathrm{d}e^{t\mathbf{X}}}{dt}\right|_{t=0}$ 也必然是实矩阵。因此，根据定义9，$\mathrm{GL}(n;\mathbb{R})$ 的李代数就是所有 $n\times n$ 实矩阵的集合，记为 $\mathrm{gl}(n;\mathbb{R})$。

注：$e^{t\mathbf{X}}=\sum_{m=0}^{\infty}\frac{(t\mathbf{X})^m}{m!}\rightarrow\frac{\mathrm{d}e^{t\mathbf{X}}}{dt}=e^{t\mathbf{X}}\mathbf{X}\rightarrow\mathbf{X}=\left.\frac{\mathrm{d}e^{t\mathbf{X}}}{dt}\right|_{t=0}$

2. 特殊线性群 $\mathrm{SL}(n;\mathbb{R})$。

根据定理5公式，$\det\left(e^{\mathbf{X}}\right)=e^{trace(\mathbf{X})}$。当 $trace(\mathbf{X})=0$，对于所有的实数 t，都有下式成立：$\det\left(e^{t\mathbf{X}}\right)=e^{0}=1$。此外，如果 $\mathbf{X}$ 是一个 $n\times n$ 的矩阵，使得 $\det\left(e^{t\mathbf{X}}\right)=1$ 对于任意实数 t 都成立。因此，对于所有的 t，$(t)(trace(\mathbf{X}))$ 必然是 $2\pi i$ 的整数倍，而这只有在 $trace(\mathbf{X})=0$ 才成立。又因为，对于所有的实数 t，$e^{t\mathbf{X}}$ 都为实矩阵，$\mathbf{X}=\left.\frac{\mathrm{d}e^{t\mathbf{X}}}{dt}\right|_{t=0}$ 必然也是实矩阵。因此 $\mathrm{SL}(n;\mathbb{R})$ 的李代数是所有迹为0的 $n\times n$ 实矩阵的集合，记为 $\mathrm{sl}(n;\mathbb{R})$。

矩阵李代数

常见的矩阵李群的李代数

3. 正交群 $O(n)$。

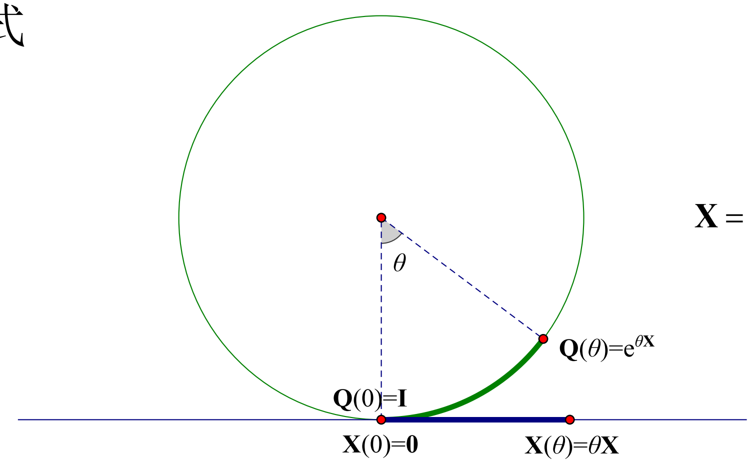
注意到，一个 $n\times n$ 矩阵 $\mathbf{U}$ 是正交的，当且仅当 $\mathbf{U}^{\mathrm{T}}=\mathbf{U}^{-1}$。所以，给定一个 $n\times n$ 实矩阵，$e^{t\mathbf{X}}$ 是正交阵，当且仅当 $e^{t\mathbf{X}^{\mathrm{T}}}=\left(e^{t\mathbf{X}}\right)^{-1}=e^{-t\mathbf{X}}$。显然，如果 $\mathbf{X}^{\mathrm{T}}=-\mathbf{X}$，上式必然成立。反之，如果该式子成立，公式两边同时在 $t=0$ 处对 t 求导，得 $\mathbf{X}^{\mathrm{T}}=-\mathbf{X}$。因此，正交群 $O(n)$ 的李代数是所有满足 $\mathbf{X}^{\mathrm{T}}=-\mathbf{X}$ 的 $n\times n$ 实矩阵的集合。特殊正交群和正交群具有相同的李代数，将他们的李代数统一记为 $\mathbf{so}(n)$。

$n=2$ 时，$SO(2)$ 中的元素都具有如下形式

$$\mathbf{Q}(\theta)=\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}$$

$\mathbf{so}(2)$ 中的元素都具有如下形式

$$\mathbf{X}(\theta)=\begin{bmatrix}0 & -\theta\\ \theta & 0\end{bmatrix}$$



$$\mathbf{X}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$$

矩阵李代数

常见的矩阵李群的李代数

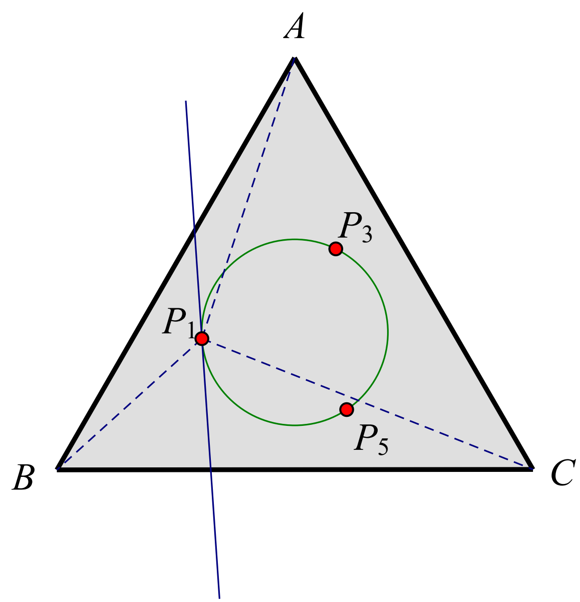
4. 循环移位群 $\mathrm{CS}(n)=\left\{\mathbf{Q}_n^x \middle| x\in\mathbb{R}, n\in 2\mathbb{Z}+1\right\}$。

对于循环移位矩阵$\mathbf{Q}_n$，可以通过矩阵对数得到 $\mathbf{X}_n=\log(\mathbf{Q}_n)$，满足 $e^{\mathbf{X}_n}=\mathbf{Q}_n$ 。因此，循环移位群的任意元素都可以表示为 $\mathbf{Q}_n^x=e^{x\mathbf{X}_n}$。这意味着$\mathrm{CS}(n)$的李代数$\mathrm{cs}(n)$为所有形如$x\mathbf{X}_n$的矩阵的几何，即循环移位群的李代数为矩阵$\mathbf{X}_n$这个向量张成的一维子空间。

$n=3$时，三阶循环移位矩阵的矩阵对数为

$$\mathbf{X}_3=\log(\mathbf{Q}_3)=\begin{bmatrix} 0 & 1.2092 & -1.2092 \\ -1.2092 & 0 & 1.2092 \\ 1.2092 & -1.2092 & 0 \end{bmatrix}$$

事实上，$\mathrm{CS}(3)$与$\mathrm{SO}(2)$ 同构。如右图，过单位元的切线也对应着 $\mathrm{CS}(3)$ 的李代数，$\mathrm{cs}(3)$



李代数

定义11（**李代数**）假设 $\mathfrak{g}$ 是域 F 上的有限维线性空间，如果 $\mathfrak{g}$ 上有一个满足如下三个条件的从 $\mathfrak{g}\times\mathfrak{g}$ 到 $\mathfrak{g}$ 的映射 $[\ ]$，则称 $\mathfrak{g}$ 为一个有限维李代数：

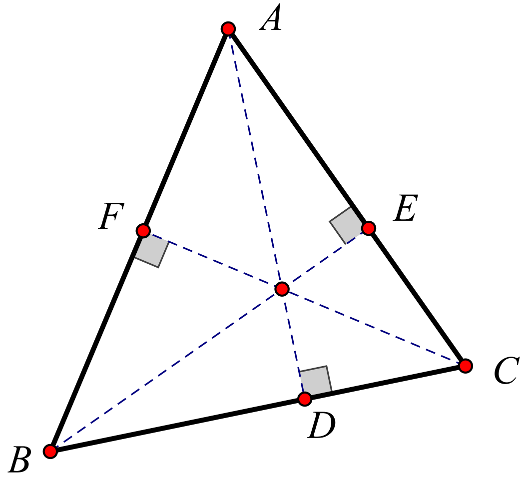
（1）双线性性：$[\ ]$ 是一个双线性映射；

（2）反对称性：对于任意的 $\mathbf{X},\mathbf{Y}\in\mathfrak{g}$，都有 $[\mathbf{X},\mathbf{Y}]=-[\mathbf{Y},\mathbf{X}]$；

（3）雅可比恒等式：$\left[\mathbf{X},[\mathbf{Y},\mathbf{Z}]\right]+\left[\mathbf{Y},[\mathbf{Z},\mathbf{X}]\right]+\left[\mathbf{Z},[\mathbf{X},\mathbf{Y}]\right]=\mathbf{0}$，其中 $\mathbf{X},\mathbf{Y},\mathbf{Z}\in\mathfrak{g}$。

例3 验证 $\mathrm{gl}(n;\mathbb{R})$ 在矩阵的括号运算下为李代数。（请自行验证）

例4 三维欧氏空间在向量叉积运算下为李代数。（请自行验证）



三维欧氏空间的雅可比恒等式的几何解释：三角形的三条垂线必然交于一点。

矩阵李群同态定理

在前面，定理1和定理2给出了矩阵李群和李群之间的关系，即每一个矩阵李群都是李群，几乎每一个李群都是矩阵李群。那么矩阵李代数和李代数又有什么关系呢？可以验证所有的矩阵李群的李代数都是李代数。那么每一个李代数是否都对应矩阵李代数呢？既然李代数的定义并不要求元素为矩阵，所以李代数是一个比矩阵李代数更抽象的概念，它的元素似乎也理所应当具有更广的范畴。但是，下面的定理给出了一个令人吃惊的事实。

定理7 每一个李代数都同构于一个矩阵李代数。

定理7表明，尽管李代数是一个更抽象的概念，但是矩阵李代数已经是它的全部了。这是一个令人震惊的结论。

矩阵李群同态定理

定义12 假设G,H是两个群。对于任意的$g_1,g_2\in G$，一个从G到H的映射 ϕ如果满足$\phi(g_1g_2)=\phi(g_1)\phi(g_2)$，则称该映射为同态映射。此外，如果该映射还是一个一一到上的映射，则称其为同构映射。

比如，对于一般线性群 $\mathrm{GL}(n;\mathbb{R})$ 上的任意矩阵 $\mathbf{A}$，其行列式$|\mathbf{A}|$ 就是一个从一般线性群 $\mathrm{GL}(n;\mathbb{R})$ 到一般线性群$\mathrm{GL}(1;\mathbb{R})$ 的映射。其中，$\mathrm{GL}(1;\mathbb{R})$ 又称非零实数群。又由于对于任意的$\mathbf{A},\mathbf{B}\in\mathrm{GL}(n;\mathbb{R})$，都有 $|\mathbf{AB}|=|\mathbf{A}||\mathbf{B}|$，因此行列式是从一般线性群到非零实数群的同态映射。

矩阵李群同态定理

定理8 假设G, H是两个矩阵李群，$\boldsymbol{g}$，$\boldsymbol{h}$ 分别是他们的李代数。假设$\phi: G \to H$是一个李群同态。那么存在唯一的实线性映射 $\tilde{\phi}: \boldsymbol{g} \to \boldsymbol{h}$，使得 $\phi\left(e^{\mathbf{X}}\right) = e^{\tilde{\phi}(\mathbf{X})}$对于所有的$\mathbf{X} \in \boldsymbol{g}$都成立。并且映射 $\tilde{\phi}$ 具有如下性质：

（1）对于任意的$\mathbf{X} \in \boldsymbol{g}$，$\mathbf{A} \in G$，都有 $\tilde{\phi}\left(\mathbf{A}\mathbf{X}\mathbf{A}^{-1}\right) = \phi(\mathbf{A})\tilde{\phi}(\mathbf{X})\phi(\mathbf{A})^{-1}$。

（2）对于任意的$\mathbf{X}, \mathbf{Y} \in \boldsymbol{g}$，都有 $\tilde{\phi}([\mathbf{X}, \mathbf{Y}]) = \left[\tilde{\phi}(\mathbf{X}), \tilde{\phi}(\mathbf{Y})\right]$。

（3）对于任意的$\mathbf{X} \in \boldsymbol{g}$，$\tilde{\phi}(\mathbf{X}) = \left.\dfrac{d\phi\left(e^{t\mathbf{X}}\right)}{dt}\right|_{t=0}$。

定理8表明，任意两个李群间的李群同态映射，必然诱导出它们的李代数间的线性映射。特别的，该定理告诉我们，两个同构的李群，有相同的李代数。

如上所述，行列式是从$\mathrm{GL}(n;\mathbb{R})$到 $\mathrm{GL}(1;\mathbb{R})$ 的映射，根据定理8，必然诱导出从$\mathrm{GL}(n;\mathbb{R})$的李代数$\mathrm{gl}(n;\mathbb{R})$ 到 $\mathrm{GL}(1;\mathbb{R})$ 的李代数 $\mathrm{gl}(1;\mathbb{R})$ 的线性映射，这个映射正是矩阵的迹。

$$\mathrm{trace}(\mathbf{X}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \det\left(\phi\left(e^{t\mathbf{X}}\right)\right)$$

矩阵李群的应用

在图像匹配中，我们曾经建立了平移匹配优化模型，其中的一项为

$$\psi_{x,y}(\mathbf{B}) = \mathbf{Q}_1^x \mathbf{B} \left(\mathbf{Q}_2^y\right)^{\mathrm{T}}$$

利用公式

$$\boxed{vec(\mathbf{ABC}) = \left(\mathbf{C}^{\mathrm{T}} \otimes \mathbf{A}\right) vec(\mathbf{B})}$$

可以将上述函数转化为

$$vec\left(\psi_{x,y}(\mathbf{B})\right) = \left(\mathbf{Q}_2^y \otimes \mathbf{Q}_1^x\right) vec(\mathbf{B})$$

因此，我们可以把平移匹配问题，转化为下列集合的代数结构问题

$$G = \left\{\mathbf{Q}_2^y \otimes \mathbf{Q}_1^x \mid x, y \in \mathbb{R}\right\}$$

矩阵李群的应用

定理9　集合 $G=\left\{\mathbf{Q}_2^y\otimes\mathbf{Q}_1^x \mid x,y\in\mathbb{R}\right\}$构成双参数矩阵李群，称之为双向循环移位群，或简称循环移位群。

上述矩阵李群，有两个特殊的一维子李群，分别为

$$G_x=\left\{\mathbf{I}_2\otimes\mathbf{Q}_1^x \mid x\in\mathbb{R}\right\}\qquad G_y=\left\{\mathbf{Q}_2^y\otimes\mathbf{I}_1 \mid y\in\mathbb{R}\right\}$$

这两个子群分别对应了图像的单向循环移位。

这两个子群的矩阵李代数的基底分别为

$$\left.\frac{\partial\left(\mathbf{I}_2\otimes\mathbf{Q}_1^x\right)}{\partial x}\right|_{x=0}=\mathbf{I}_2\otimes\ln\mathbf{Q}_1\qquad\left.\frac{\partial\left(\mathbf{Q}_2^y\otimes\mathbf{I}_1\right)}{\partial y}\right|_{y=0}=\ln\mathbf{Q}_2\otimes\mathbf{I}_1$$

矩阵李群的应用

而这两个基底，正好构成了循环移位群 G 的矩阵李代数的一组标准正交基。那么，平移参数的确定，就可以从循环移位群上转化到其李代数上。相应的，可以得到公式

$$\ln\left(\mathbf{Q}_2^y \otimes \mathbf{Q}_1^x\right) = x\mathbf{I}_2 \otimes \ln \mathbf{Q}_1 + y \ln \mathbf{Q}_2 \otimes \mathbf{I}_1$$

讨论如何利用上述公式进行图像匹配?

矩阵李群的应用

记 $$\mathbf{Q}^{x,y}=\mathbf{Q}_2^y\otimes\mathbf{Q}_1^x$$

则 $$vec(\mathbf{A})=\mathbf{Q}^{x,y}vec(\mathbf{B})$$

其，对于 $$i=0,1,\cdots,M,\ j=0,1,\cdots,N$$

都有 $$\mathbf{Q}^{i,j}vec(\mathbf{A})=\mathbf{Q}^{x,y}\left(\mathbf{Q}^{i,j}vec(\mathbf{B})\right)$$

记 $$\mathbf{T}_1=\left[\mathbf{Q}^{0,0}vec(\mathbf{B}),\mathbf{Q}^{1,0}vec(\mathbf{B}),\cdots,\mathbf{Q}^{M-1,N-1}vec(\mathbf{B})\right]$$

$$\mathbf{T}_2=\left[\mathbf{Q}^{0,0}vec(\mathbf{A}),\mathbf{Q}^{1,0}vec(\mathbf{A}),\cdots,\mathbf{Q}^{M-1,N-1}vec(\mathbf{A})\right]$$

则 $$\mathbf{T}_2=\mathbf{Q}^{x,y}\mathbf{T}_1\Rightarrow\mathbf{Q}^{x,y}=\mathbf{T}_2\mathbf{T}_1^{-1}$$

然后基于最小二乘法，即可得到平移参数 x,y

矩阵李群的应用

循环移位矩阵总结：

1. 循环移位矩阵的特征值（基本线性动作）
2. 循环移位矩阵的开方条件（矩阵开方定理，要求奇数阶）
3. 循环移位矩阵的特征向量（离散傅里叶变换）
4. 循环移位矩阵的任意次幂（sinc插值）
5. 循环移位矩阵的任意次幂的集合（构成循环移位群）
6. 循环移位群的李代数（图像匹配矩阵公式）
7. 旋转和缩放也可以转化为平移，一定程度可以说，循环移位群及李代数揭示了图像匹配的基本代数结构

小结

1. 群的概念源于自然界中的对称性
2. 李群既是群又是微分流形，既有对称性，又有光滑性
3. 一般线性群是最大的矩阵李群，所有的矩阵李群都是一般线性群的子群。
4. 每一个矩阵李群都是李群。
5. 几乎每一个李群都是矩阵李群。
6. 李代数是李群在单位元处的切空间。
7. 李代数同构于矩阵李代数。
8. 根据矩阵李群同态定理，矩阵的迹可以由矩阵的行列式定义。

谢　谢

耿修瑞

中国科学院空天信息创新研究院

gengxr@sina.com.cn