庄逸的数学与技术屋

三维矩阵不变量与平行六面体体积

Vortexer99

目录

1 主要问题

1.1 混合积和体积

对于三维空间中三个不共面的向量 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 而言，其混合积 $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ 表示三个向量（经过平移使其共起点）张成的平行六面体体积。将其简记为 $[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}]$ ，其计算方法可用行列式表示。

$$[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} \tag{1}$$

1.2 矩阵的不变量

对于三维矩阵 $\boldsymbol{A}$ 而言，其三个不变量为

$$I_1(\boldsymbol{A}) = \operatorname{tr} \boldsymbol{A} \tag{2}$$

$$I_2(\boldsymbol{A}) = \frac{1}{2}(\operatorname{tr}^2 \boldsymbol{A} - \operatorname{tr} \boldsymbol{A}^2) \tag{3}$$

$$I_3(\boldsymbol{A}) = \det \boldsymbol{A} = \frac{1}{6}(\operatorname{tr}^3 \boldsymbol{A} - 3(\operatorname{tr} \boldsymbol{A})(\operatorname{tr} \boldsymbol{A}^2) + 2 \operatorname{tr} \boldsymbol{A}^3) \tag{4}$$

1.3 问题

矩阵的不变量可以如下计算。对于任意不共面矢量 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$，有

$$I_1(\boldsymbol{A}) = \frac{[\boldsymbol{Au}\ \boldsymbol{v}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{Av}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{Aw}]}{[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}]} \tag{5}$$

$$I_2(\boldsymbol{A}) = \frac{[\boldsymbol{Au}\ \boldsymbol{Av}\ \boldsymbol{w}] + [\boldsymbol{Au}\ \boldsymbol{v}\ \boldsymbol{Aw}] + [\boldsymbol{u}\ \boldsymbol{Av}\ \boldsymbol{Aw}]}{[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}]} \tag{6}$$

$$I_3(\boldsymbol{A}) = \frac{[\boldsymbol{Au}\ \boldsymbol{Av}\ \boldsymbol{Aw}]}{[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}]} \tag{7}$$

为什么？

2 证明

其实矩阵的不变量就是特征值的对称多项式。而如果 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 刚好是 $\boldsymbol{A}$ 的特征向量，结论是显然的。因此，对一般向量，可用特征向量分解，得到类似的结论。

2.1 不变量的来龙去脉

依照求特征值的方法，三维矩阵 $\boldsymbol{A}$ 的三个特征值 $\lambda_1, \lambda_2, \lambda_3$ 能使得

$$|\boldsymbol{A} - \lambda \boldsymbol{E}| = 0 \tag{8}$$

成立。其中 $\boldsymbol{E}$ 为单位矩阵。而上式左端可以看作以 λ 为变量的三次多项式。事实上，正有

$$|\boldsymbol{A}-\lambda\boldsymbol{E}| = -\lambda^3 + I_1(\boldsymbol{A})\lambda^2 - I_2(\boldsymbol{A})\lambda + I_3(\boldsymbol{A}) = 0 \tag{9}$$

证明略，根据对应的 λ 次数，算一下余子式即可。于是，根据韦达定理可知

$$I_1(\boldsymbol{A}) = \lambda_1 + \lambda_2 + \lambda_3 \tag{10}$$

$$I_2(\boldsymbol{A}) = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \tag{11}$$

$$I_3(\boldsymbol{A}) = \lambda_1\lambda_2\lambda_3 \tag{12}$$

2.2 特征分解

取矩阵的三个特征向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$，对 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 进行特征分解。利用求和约定（**省去 1 至 3 的求和号**）（且上标不表示次数）

$$\boldsymbol{u} = u^i\boldsymbol{e}_i \qquad \boldsymbol{v} = v^i\boldsymbol{e}_i \qquad \boldsymbol{w} = w^i\boldsymbol{e}_i \tag{13}$$

于是

$$[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] = u^i v^j w^k \boldsymbol{e}_i \cdot (\boldsymbol{e_j} \times \boldsymbol{e}_k) = u^i v^j w^k [\boldsymbol{e}_i\ \boldsymbol{e}_j\ \boldsymbol{e}_k] = u^i v^j w^k V(i,j,k) \tag{14}$$

其中 $V(i,j,k) = [\boldsymbol{e}_i\ \boldsymbol{e}_j\ \boldsymbol{e}_k]$。显然，由混合积的性质，只要 i,j,k 中有任意两个相等，$V(i,j,k)$ 就等于 0。

然后看第一项。根据特征向量的性质

$$\boldsymbol{A}\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i \qquad i = 1,2,3 \tag{15}$$

有

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}u^i\boldsymbol{e}_i = \lambda_i u^i \boldsymbol{e}_i \tag{16}$$

于是

$$[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] = \lambda_i u^i v^j w^k V(i,j,k) \tag{17}$$

那么，类似的有

$$[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{A}\boldsymbol{v}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{A}\boldsymbol{w}] = (\lambda_i + \lambda_j + \lambda_k) u^i v^j w^k V(i,j,k) \tag{18}$$

根据前面的讨论，求和时 i,j,k 有任意两个相等时，求和项因为 $V(i,j,k) = 0$ 而为零。因此有贡献的项是 i,j,k 互不相等时的项。而 i,j,k 互不相等时，显然 $\lambda_i + \lambda_j + \lambda_k = I_1(\boldsymbol{A})$，于是

$$\begin{aligned}[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{A}\boldsymbol{v}\ \boldsymbol{w}] + [\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{A}\boldsymbol{w}] &= I_1(\boldsymbol{A}) u^i v^j w^k V(i,j,k) && (19)\\ &= I_1(\boldsymbol{A}) [\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] && (20)\end{aligned}$$

这就证得了第一式。对于第二、第三式，做法是类似的。考虑第二式中某项，

$$[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{A}\boldsymbol{v}\ \boldsymbol{w}] = \lambda_i\lambda_j u^i v^j w^k V(i,j,k) \tag{21}$$

最后得到的系数是 $\lambda_i\lambda_j + \lambda_i\lambda_k + \lambda_j\lambda_k$，当 i,j,k 互不相等时显然就是 $I_2(\boldsymbol{A})$。

3 细枝末节的东西

3.1 如果矩阵不满秩

如果矩阵 $\boldsymbol{A}$ 不满秩会怎么样？这也就是说对应的线性映射 $\mathcal{A}$ 像空间不是全空间 V，即核空间维数不为零，会将一些非零向量映射为零向量。

此时像空间中的 r 个向量及其特征值没有影响。对于核空间，取 $3-r$ 个基向量，作为特征向量，并令其对应特征值为零即可。

这样既保证了任何向量都能被特征向量分解，又使得 $\boldsymbol{A}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$ 成立。

3.2 混合积的线性性质

将任意向量按照特征向量分解时，利用了混合积的线性性质。由于混合积是行列式运算，行列式又是多重线性函数，这一点是显然的。

一个简单的证明是，点积是线性的，即

$$(a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2)\cdot(\boldsymbol{v}\times\boldsymbol{w}) = a_1\boldsymbol{u}_1\cdot(\boldsymbol{v}\times\boldsymbol{w}) + a_2\boldsymbol{u}_2\cdot(\boldsymbol{v}\times\boldsymbol{w}) \tag{22}$$

然后利用混合积的轮换性质（可通过行列式表示证明，但是通过行列式性质亦可直接证得混合积是线性的），可知 $\boldsymbol{v},\boldsymbol{w}$ 也是线性项。

3.3 为何称作不变量

矩阵的不变量之所以被称为不变量，是因为它们在坐标变换下不变。

考虑坐标变换（正交）矩阵 $\boldsymbol{S}$，矩阵 $\boldsymbol{A}$ 经变换后得到矩阵 $\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$。计算 $\boldsymbol{B}$ 的特征多项式，利用正交矩阵和行列式的性质，

$$|\boldsymbol{B}-\lambda\boldsymbol{E}| = \left|\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}-\boldsymbol{S}^{-1}\lambda\boldsymbol{S}\right| = \left|\boldsymbol{S}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{S}\right| = \left|\boldsymbol{S}^{-1}\right||\boldsymbol{A}-\lambda\boldsymbol{E}||\boldsymbol{S}| = |\boldsymbol{A}-\lambda\boldsymbol{E}| \tag{23}$$

可见 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征多项式相同，因此由式 9和式 10的推导可知 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的不变量也相同。

3.4 一个简便证法

对于第三式，可以通过写成矩阵相乘的形式发现

$$[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{A}\boldsymbol{v}\ \boldsymbol{A}\boldsymbol{w}] = \left|\boldsymbol{A}\begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix}\right| = |\boldsymbol{A}|\,[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] \tag{24}$$

于是便很显然了。

4 * 粗暴展开硬核做法及反推结论

注：最初我是这么做的，但是卡在了一些关键步骤的证明上。现在倒过来可以证明那些关键步骤。

将点乘和叉乘全部利用求和展开，得到

$$[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] = u^i v^j w^k \varepsilon_{ijk} \tag{25}$$

此处 $\boldsymbol{u} = u^i \boldsymbol{b}_i$，$\boldsymbol{b}_i, \boldsymbol{b}_j, \boldsymbol{b}_k$ 为笛卡尔直角坐标系的三个基向量。Levi-Civita 符号 ε_{ijk} 在 ijk 中有任意两个相同时取 0，全不相同时为 ± 1，符号由 ijk 的置换奇偶性决定（例如 $\varepsilon_{123} = 1$）

由于 $\boldsymbol{A}\boldsymbol{u} = A_{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \cdot u^k \boldsymbol{e}_k = A_{ij} u^j \boldsymbol{e}_i$，因此

$$[\boldsymbol{A}\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}] = A_{li} u^i v^j w^k \varepsilon_{ljk} \tag{26}$$

同理，有

$$[\boldsymbol{u}\ \boldsymbol{A}\boldsymbol{v}\ \boldsymbol{w}] = A_{lj} u^i v^j w^k \varepsilon_{ilk} \tag{27}$$

$$[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{A}\boldsymbol{w}] = A_{lk} u^i v^j w^k \varepsilon_{ijl} \tag{28}$$

由于我们已经证明了等式，这也就是说

$$u^i v^j w^k (A_{li} \varepsilon_{ljk} + A_{lj} \varepsilon_{ilk} + A_{lk} \varepsilon_{ijl}) = \operatorname{tr} \boldsymbol{A} u^i v^j w^k \varepsilon_{ijk} \tag{29}$$

可以从中提取出一般性的抽象规律（注意对 i, j, k, l 求和)

$$f(i,j,k)[g(l,i)\varepsilon_{ljk} + g(l,j)\varepsilon_{ilk} + g(l,k)\varepsilon_{ijl}] = f(i,j,k)g(l,l)\varepsilon_{ijk} \tag{30}$$

这个还是容易证明的。首先，可以证明 i, j, k 有一组相等时左边求和项为零。例如 $i = j$ 时，中括号里为

$$[g(l,i)\varepsilon_{lik} + g(l,i)\varepsilon_{ilk} + g(l,k)\varepsilon_{iil}] \tag{31}$$

由于 $\varepsilon_{iil} = 0$，$\varepsilon_{lik} = -\varepsilon_{ilk}$，因此上式为零。

对于 i, j, k 互不相等的情况，因为 i, j, k, l 只能取 $1, 2, 3$，l 必须与其中一个相等。如果 $l = i$，则 $\varepsilon_{ilk} = \varepsilon_{ijl} = 0$。那么总共只有一项

$$f(i,j,k)g(i,i)\varepsilon_{ijk} \tag{32}$$

注意此处求和限制 $l = i$。注意到只对互不相等的 i, j, k 进行求和，对任意函数 F 有

$$\sum_{i,j,k} (F(i,j,k,i) + F(i,j,k,j) + F(i,j,k,k)) = \sum_{i,j,k} \sum_{l} F(i,j,k,l) \tag{33}$$

也就是令 $l = i, l = j, l = k$ 得到的三项加起来，结果就是

$$f(i,j,k)(g(i,i) + g(j,j) + g(k,k))\varepsilon_{ijk} = f(i,j,k)g(l,l)\varepsilon_{ijk} \tag{34}$$

对于另外两个等式，相应也可得到两个结论

$$
\begin{aligned}
f(i,j,k)[g(l,i)g(m,j)\varepsilon_{lmk} + g(l,j)g(m,k)\varepsilon_{lmi} + g(l,k)g(m,i)\varepsilon_{lmj}] \\
= f(i,j,k)\varepsilon_{ijk}\frac{1}{2}[g(l,l)g(m,m) - g(l,m)g(m,l)]
\end{aligned}
\tag{35}
$$

$$
f(i,j,k)[g(l,i)g(m,j)g(n,k)\varepsilon_{lmn}] = f(i,j,k)\varepsilon_{ijk}\,|g| \tag{36}
$$

其中 $|g|$ 表示对写成函数形式的矩阵 g 求行列式。

声明

1. 博客内容仅为经验之谈，如认为有问题请带着批判性思维自行辨别或与我讨论，本人不负责因盲目应用博客内容导致的任何损失。
2. 虽然文章的思想不一定是原创的，但是写作一定是原创的，如有雷同纯属巧合。
3.

博客信息 此文章的博客来源：https://vortexer99.github.io/

自豪地采用 LaTeX！