详细解答

STEP II 2020，第 6 题

纯数学 · 矩阵的幂与迹

STEP II 2020，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2020 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696589-2020-past-papers.zip

本题在 *Beyond STEP* 第 7.4 节回到起点的矩阵中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch07/finite-order.html

关键思路

每个实 2×2 矩阵（matrix）都满足它自己的特征方程（characteristic equation）：

$$\mathbf{M}^2 - t\,\mathbf{M} + \delta\,\mathbf{I} = \mathbf{0}, \qquad t = \operatorname{tr}(\mathbf{M}),\ \delta = \det(\mathbf{M}),$$

这就是凯莱–哈密顿定理（Cayley–Hamilton theorem）（把乘积展开，两行便可验证）。一旦把 $\mathbf{M}^2$ 写成 $\mathbf{M}$ 与 $\mathbf{I}$ 的组合，任何像 $\mathbf{M}^2 = \mathbf{I}$ 这样的方程都会变成“$t\mathbf{M}$ 是 $\mathbf{I}$ 的倍数”，这就迫使 $t = 0$，或者 $\mathbf{M}$ 本身是 $\mathbf{I}$ 的倍数。之后，第 (iii) 及 (iv) 部分只需把第 (ii) 部分应用到 $\mathbf{M}^2$ 和 $\mathbf{M}^4$（而非 $\mathbf{M}$）上便可得出。

全文记 $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$、$t = a + d$ 及 $\delta = ad - bc$。

解答

第 (i) 部分

把乘积展开，

$$\mathbf{M}^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & bc + d^2 \end{pmatrix} = (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} - (ad - bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

因为例如 $(a+d)a-(ad-bc)=a^2+bc$ 及 $(a+d)b=ab+bd$。所以

$$\mathbf{M}^2=t\mathbf{M}-\delta\mathbf{I}. \qquad (*)$$

对 $(*)$ 取迹（trace），并利用 $\operatorname{tr}(\mathbf{I})=2$，得

$$\operatorname{tr}(\mathbf{M}^2)=t\cdot t-2\delta=\operatorname{tr}(\mathbf{M})^2-2\det(\mathbf{M}),$$

这正是所求。（直接计算亦可：$a^2+2bc+d^2=(a+d)^2-2(ad-bc)$。）

第 (ii) 部分

第一个等价（equivalence），$(\Rightarrow)$。设 $\mathbf{M}^2=\mathbf{I}$ 且 $\mathbf{M}\neq\pm\mathbf{I}$。代入 $(*)$ 得

$$t\,\mathbf{M}=(1+\delta)\,\mathbf{I}.$$

若 $t\neq 0$，便可两边除以 t，得 $\mathbf{M}=k\mathbf{I}$，其中 $k=(1+\delta)/t$ 为实数；于是 $\mathbf{M}^2=k^2\mathbf{I}=\mathbf{I}$ 给出 $k=\pm1$，即 $\mathbf{M}=\pm\mathbf{I}$，但这已被排除。因此 $t=0$，从而 $(1+\delta)\mathbf{I}=\mathbf{0}$，所以 $\delta=-1$。

$(\Leftarrow)$。若 $t=0$ 且 $\delta=-1$，则 $(*)$ 给出 $\mathbf{M}^2=0\cdot\mathbf{M}+\mathbf{I}=\mathbf{I}$。此外 $\operatorname{tr}(\pm\mathbf{I})=\pm2\neq0$，所以 $\mathbf{M}\neq\pm\mathbf{I}$。

第二个等价，$(\Rightarrow)$。若 $\mathbf{M}^2=-\mathbf{I}$，则 $(*)$ 给出 $t\mathbf{M}=(\delta-1)\mathbf{I}$。若 $t\neq0$，则 $\mathbf{M}=k\mathbf{I}$，其中 k 为实数，而 $\mathbf{M}^2=k^2\mathbf{I}=-\mathbf{I}$ 便需要 $k^2=-1$，这对实数 k 是不可能的。所以 $t=0$，从而 $\delta=1$。

$(\Leftarrow)$。若 $t=0$ 且 $\delta=1$，则 $(*)$ 立即给出 $\mathbf{M}^2=-\mathbf{I}$。

> ? **想一想** 我们究竟在哪里用到 $\mathbf{M}$ 是实矩阵这一点?

只在第二个等价中用到：$k^2=-1$ 没有实数解。对于复矩阵，$\mathbf{M}=\mathrm{i}\mathbf{I}$ 满足 $\mathbf{M}^2=-\mathbf{I}$，但其迹为 $2\mathrm{i}\neq0$，所以结论便不成立。

第 (iii) 部分

$(\Leftarrow)$ 是显然的：若 $\mathbf{M}^2=\pm\mathbf{I}$，则 $\mathbf{M}^4=(\pm\mathbf{I})^2=\mathbf{I}$。

$(\Rightarrow)$。设 $\mathbf{N}=\mathbf{M}^2$，这是一个实 2×2 矩阵，且 $\mathbf{N}^2=\mathbf{I}$。要么 $\mathbf{N}=\pm\mathbf{I}$，这正是我们想要的；要么 $\mathbf{N}\neq\pm\mathbf{I}$，此时第 (ii) 部分告诉我们 $\det(\mathbf{N})=-1$。但

$$\det(\mathbf{N})=\det(\mathbf{M})^2\geq0,$$

所以第二种情况不可能出现。因此 $\mathbf{M}^2=\pm\mathbf{I}$。

$\mathbf{M}^4=-\mathbf{I}$ 的条件。同样设 $\mathbf{N}=\mathbf{M}^2$。把第 (ii) 部分的第二个等价应用到 $\mathbf{N}$ 上，

$$\mathbf{M}^4=-\mathbf{I}\iff\operatorname{tr}(\mathbf{N})=0\text{ 且 }\det(\mathbf{N})=1.$$

利用第 (i) 部分及 $\det(\mathbf{N}) = \delta^2$，这等于说

$$t^2 - 2\delta = 0 \quad 及 \quad \delta^2 = 1.$$

由 $\delta^2 = 1$ 得 $\delta = \pm 1$，而 $t^2 = 2\delta \geq 0$ 排除了 $\delta = -1$。所以 $\delta = 1$ 且 $t^2 = 2$。每一步都是等价的，因此

$$\boxed{\mathbf{M}^4 = -\mathbf{I} \iff \det(\mathbf{M}) = 1 \text{ 且 } \operatorname{tr}(\mathbf{M}) = \pm\sqrt{2}}.$$

第 (iv) 部分

把第 (iii) 部分应用到 $\mathbf{M}^2$ 上：$\mathbf{M}^8 = \mathbf{I} \iff \mathbf{M}^4 = \pm\mathbf{I}$，所以满足 $\mathbf{M}^8 = \mathbf{I}$ 的矩阵恰好是满足以下条件的矩阵：

$$\mathbf{M}^2 = \pm\mathbf{I} \qquad 或 \qquad \det(\mathbf{M}) = 1,\ \operatorname{tr}(\mathbf{M}) = \pm\sqrt{2}.$$

旋转（rotation）和反射（reflection）就是正交（orthogonal）矩阵：它们的列是互相垂直的单位向量（unit vector）（旋转的行列式（determinant）为 1，反射的行列式为 -1）。所以我们只需在上述矩阵中找一个列不是单位向量的矩阵。例如

$$\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

的迹为 0、行列式为 -1，所以由第 (ii) 部分得 $\mathbf{M}^2 = \mathbf{I}$（验证：$\mathbf{M}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$），因而 $\mathbf{M}^8 = \mathbf{I}$。它的第二列 $(1, -1)$ 的长度为 $\sqrt{2} \neq 1$，所以它既不是旋转，也不是反射。（它是一个“斜反射”：它保持沿 x 轴的每个向量不变，并把沿 $(1, -2)$ 方向的每个向量反向，而这两个方向并不互相垂直。）

较有趣的那一类中的一个例子是 $\mathbf{M} = \begin{pmatrix} \sqrt{2} & 1 \\ -1 & 0 \end{pmatrix}$：行列式为 1、迹为 $\sqrt{2}$，所以 $\mathbf{M}^4 = -\mathbf{I}$ 及 $\mathbf{M}^8 = \mathbf{I}$，但它的第一列长度为 $\sqrt{3}$。它是一个“变形了的”45° 旋转：它的阶（order）恰好是 8，与该旋转相同，但不保持长度。

验算

- 转角为 45° 的旋转 $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ 的行列式为 1、迹为 $\sqrt{2}$，而转四次 45° 便是转半圈，即 $-\mathbf{I}$：与第 (iii) 部分一致。
- 把 $\mathbf{M} = \begin{pmatrix} \sqrt{2} & 1 \\ -1 & 0 \end{pmatrix}$ 平方，得 $\begin{pmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & -1 \end{pmatrix}$，其迹为 0、行列式为 1，所以由第 (ii) 部分，它的平方是 $-\mathbf{I}$。

- 以电脑搜索所有元素属于 $\{-3, -\frac{5}{2}, \ldots, 3\}$ 的矩阵，确切验证了第 (ii) 部分的两个等价，以及第 (iii) 部分的第一个等价。

常见错误

- 未先处理 $t = 0$ 的情况便除以 t；第 (ii) 部分的整个论证，正是按 t 是否为零来分情况讨论。
- 在第 (iii) 部分，忘记 $\det(\mathbf{M}^2) = \det(\mathbf{M})^2$ 不可能是负数；正是这一点排除了 $\mathbf{I}$ 的“其他”平方根（square root）。
- 在第 (iv) 部分以 45° 旋转作为例子：它虽然满足 $\mathbf{M}^8 = \mathbf{I}$，却正正是题目所排除的情况。

STEP 之外

恒等式 $(*)$ 是凯莱–哈密顿定理在 2×2 情况下的特例：每个方阵都满足它的特征多项式（characteristic polynomial）。像“哪些 $\mathbf{M}$ 满足 $\mathbf{M}^n = \mathbf{I}$?”这类问题，实质上是关于特征值（eigenvalue）的问题：在这里，特征值必须是八次单位根（roots of unity），而一个阶为 8 的实矩阵与某个旋转相似（similar），但本身不一定是旋转。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）正是本着这种精神，阐述特征值、相似及特征多项式。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。