详细解答

STEP II 1987，第 9 题

纯数学 · 正交矩阵

> **STEP II 1987，第 9 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/87-s2-q9

关键思路

把 $\mathbf{B}$ 转置（transpose）之后，条件 $\mathbf{B}^{\mathrm{T}}\mathbf{B}=\mathbf{I}$ 的最左和最右两端各有一个逆矩阵（inverse matrix）。由于这些逆矩阵本身可逆（invertible），我们可以左乘 $\mathbf{I}-\mathbf{A}^{\mathrm{T}}$、右乘 $\mathbf{I}-\mathbf{A}$，而不会损失任何信息。余下的是一条关于 $\mathbf{A}$ 和 $\mathbf{A}^{\mathrm{T}}$ 的多项式恒等式，几乎立刻就能化简。要小心的是每一步都必须可以逆推，因为我们要证明的是“若且仅若”（if and only if）。

解答

我们会用到以下标准结果：

- $(\mathbf{XY})^{\mathrm{T}}=\mathbf{Y}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}$，而且转置是线性（linear）的，所以 $(\mathbf{I}\pm\mathbf{A})^{\mathrm{T}}=\mathbf{I}\pm\mathbf{A}^{\mathrm{T}}$（因为 $\mathbf{I}^{\mathrm{T}}=\mathbf{I}$）；
- 若 $\mathbf{X}$ 是非奇异（non-singular）的，则 $\mathbf{X}^{\mathrm{T}}$ 也是，并且 $(\mathbf{X}^{-1})^{\mathrm{T}}=(\mathbf{X}^{\mathrm{T}})^{-1}$。

第 1 步：$\mathbf{B}$ 的转置。 由于 $\mathbf{I}-\mathbf{A}$ 是非奇异的，它的转置 $\mathbf{I}-\mathbf{A}^{\mathrm{T}}$ 也是非奇异的。把乘积的次序倒转，

$$\mathbf{B}^{\mathrm{T}}=\left((\mathbf{I}-\mathbf{A})^{-1}\right)^{\mathrm{T}}(\mathbf{I}+\mathbf{A})^{\mathrm{T}}=(\mathbf{I}-\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{I}+\mathbf{A}^{\mathrm{T}}).$$

因此

$$\mathbf{B}^{\mathrm{T}}\mathbf{B}=(\mathbf{I}-\mathbf{A}^{\mathrm{T}})^{-1}\,(\mathbf{I}+\mathbf{A}^{\mathrm{T}})(\mathbf{I}+\mathbf{A})\,(\mathbf{I}-\mathbf{A})^{-1}.$$

第 2 步：以可逆推的方式消去逆矩阵。 设 $\mathbf{P}=\mathbf{I}-\mathbf{A}^{\mathrm{T}}$、$\mathbf{Q}=\mathbf{I}-\mathbf{A}$ 及 $\mathbf{M}=(\mathbf{I}+\mathbf{A}^{\mathrm{T}})(\mathbf{I}+\mathbf{A})$，于是 $\mathbf{B}^{\mathrm{T}}\mathbf{B}=\mathbf{P}^{-1}\mathbf{M}\mathbf{Q}^{-1}$。我们断言

$$\mathbf{P}^{-1}\mathbf{M}\mathbf{Q}^{-1}=\mathbf{I}\quad\Longleftrightarrow\quad\mathbf{M}=\mathbf{P}\mathbf{Q}.$$

（⇒）左乘 $\mathbf{P}$、右乘 $\mathbf{Q}$：$\mathbf{M} = \mathbf{PIQ} = \mathbf{PQ}$。
（⇐）左乘 $\mathbf{P}^{-1}$、右乘 $\mathbf{Q}^{-1}$：$\mathbf{P}^{-1}\mathbf{MQ}^{-1} = \mathbf{P}^{-1}\mathbf{PQQ}^{-1} = \mathbf{I}$。

> ? **想一想** 为甚么不能只是“两边同乘”就了事?

把一条等式乘以一个矩阵，得到的必定仍是正确的等式，但这一步一般不能还原：由 $\mathbf{XY} = \mathbf{XZ}$ 不能推断 $\mathbf{Y} = \mathbf{Z}$，除非 $\mathbf{X}$ 可逆。这里反方向之所以成立，全因 $\mathbf{P}$ 和 $\mathbf{Q}$ 都有逆矩阵，而这正是用上“$\mathbf{I} - \mathbf{A}$ 是非奇异的”（因而 $\mathbf{I} - \mathbf{A}^{\mathrm{T}}$ 也是）这个假设的地方。还要留意，$\mathbf{P}$ 必须保持在左边，$\mathbf{Q}$ 必须保持在右边，因为矩阵乘法不符合交换律（commutative）。

第 3 步：展开。 把 $\mathbf{M} = \mathbf{PQ}$ 的两边展开，并小心保持各因子的次序，

$$(\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A}) = \mathbf{I} + \mathbf{A} + \mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A},$$
$$(\mathbf{I} - \mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}) = \mathbf{I} - \mathbf{A} - \mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A}.$$

$\mathbf{I}$ 和 $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 这两项在两边都出现，所以相减后得

$$\mathbf{M} = \mathbf{PQ} \iff 2(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) = \mathbf{O} \iff \mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{O}.$$

以上每一步都可以逆推：两边同时加上或减去同一个矩阵，以及乘以非零纯量（scalar）2 或 $\frac{1}{2}$。

结论。 把第 1 至第 3 步串连起来，每一环都是等价关系，

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I} \iff (\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A}) = (\mathbf{I} - \mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}) \iff \mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{O},$$

这正是我们要证明的结果。用文字表达：$\mathbf{B}$ 是正交矩阵（orthogonal matrix），当且仅当 $\mathbf{A}$ 是斜对称矩阵（skew-symmetric matrix）。

验算

- 取 $\mathbf{A} = \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix}$，它是斜对称的。这时 $\det(\mathbf{I} - \mathbf{A}) = 1 + t^2 \neq 0$，而
$$\mathbf{B} = \frac{1}{1+t^2}\begin{pmatrix} 1 - t^2 & 2t \\ -2t & 1 - t^2 \end{pmatrix},$$
这是一个旋转矩阵（rotation matrix）（代入 $t = \tan\frac{\theta}{2}$），所以一如所料 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}$。
- 1×1 的情况：$\mathbf{A} = (a)$，其中 $a \neq 1$，得 $\mathbf{B} = \left(\frac{1+a}{1-a}\right)$，而 $\mathbf{B}^2 = 1$ 迫使 $1 + a = \pm(1 - a)$，即 $a = 0$（另一个符号不可能成立）。这与 $2a = 0$ 吻合。
- 我们以精确的有理数运算测试了约 2000 个大小由 1 至 5 的随机矩阵：每一个斜对称的 $\mathbf{A}$ 都准确地给出 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}$，其余的则全部不成立。

常见错误

- 没有倒转次序就写 $\mathbf{B}^{\mathrm{T}} = (\mathbf{I}+\mathbf{A}^{\mathrm{T}})(\mathbf{I}-\mathbf{A}^{\mathrm{T}})^{-1}$。（这条式子碰巧是对的，因为 $\mathbf{I}+\mathbf{A}^{\mathrm{T}}$ 与 $(\mathbf{I}-\mathbf{A}^{\mathrm{T}})^{-1}$ 可交换，但这一点需要另行说明理由。）
- 展开 $(\mathbf{I}+\mathbf{A}^{\mathrm{T}})(\mathbf{I}+\mathbf{A})$ 时写成 $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ 而不是 $\mathbf{A}^{\mathrm{T}}\mathbf{A}$。在这一题里两种写法都会抵销，但前提是两边的展开方式必须一致。
- 只证明了一个方向。题目要求的是“若且仅若”，所以要说明每一步为何可以逆推。

STEP 之外

映射 $\mathbf{A} \mapsto (\mathbf{I}+\mathbf{A})(\mathbf{I}-\mathbf{A})^{-1}$ 称为凯莱变换（Cayley transform）。它把构成一个向量空间（vector space）的斜对称矩阵，变成构成一个群（group）的正交矩阵，从而给出旋转的有理参数化（rational parametrisation）。（可以证明，当 $\mathbf{A}$ 是实斜对称矩阵时，$\mathbf{I}-\mathbf{A}$ 必定可逆，而且所得到的正交矩阵恰好是那些没有特征值（eigenvalue）-1 的正交矩阵。）这让我们初次窥见李群（Lie group）与它的李代数（Lie algebra）之间的关系。想进一步了解正交矩阵和斜对称矩阵，Strang 的 *Introduction to Linear Algebra*（线性代数入门）是很好的读物。

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。