线性代数 (XII)

Lecturer: 杨启哲

Last modified: 2025 年 5 月 11 日

1 特征值介绍

这一章我们从斐波那契数列开始，斐波那契数列 $0,1,1,2,3,5,8,13,21,\cdots$，是一个递归数列，定义为

$$F_n = \begin{cases} 0 & n=0 \\ 1 & n=1 \\ F_{n-1}+F_{n-2} & n \geqslant 2 \end{cases}$$

如果我们将其递推过程写成矩阵的形式，即:

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}$$

我们将这个矩阵记为 $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$，则斐波那契数列可以如下计算:

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = A^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{1}$$

另一方面，由递推公式的求法（求 a, c 满足 $F_{k+2}+aF_{k+1}=c(F_{k+1}+aF_k)$）我们可以得到:

$$\begin{aligned} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k &= \frac{1+\sqrt{5}}{2}(F_k+\frac{-1+\sqrt{5}}{2}F_{k-1}) \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k &= \frac{1-\sqrt{5}}{2}(F_k+\frac{-1-\sqrt{5}}{2}F_{k-1}) \end{aligned}$$

将上述式子写成矩阵的形式有:

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_k+\frac{-1+\sqrt{5}}{2}F_{k-1} \\ F_k+\frac{-1-\sqrt{5}}{2}F_{k-1} \end{bmatrix}$$

从而通过类似的讨论可以得到:

$$\begin{aligned} \begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix} &= \left(\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}\right)^k \begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix} \begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix} \end{aligned} \tag{2}$$

注意到:

$$\begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}, \quad \begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

从而有:

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_k+\frac{-1+\sqrt{5}}{2}F_{k-1} \\ F_k+\frac{-1-\sqrt{5}}{2}F_{k-1} \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

因此我们得到了:

$$\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}$$

令: $X = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}, \Lambda = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$。则上述式子可改写为:

$$XA = \Lambda X, \quad \text{即: } A = X^{-1}\Lambda X.$$

注意到此时:

$$X^{-1} = \begin{bmatrix} \frac{5+\sqrt{5}}{10} & \frac{5-\sqrt{5}}{10} \\ \frac{\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix}$$

回到等式(1)和等式(2),注意到:

$$\begin{aligned} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\ &= \left(X^{-1} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} X \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\ &= \left(X^{-1} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix} X \right) \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \end{aligned}$$

这也说明等式等式(1)和等式(2)是一致的,同时 $A = X^{-1}\Lambda X$ 的分解也给出了一个简单的计算 A^n 的方式。

特征值和特征向量 令 X^{-1} 的列向量是 $\boldsymbol{x}_1, \boldsymbol{x}_2$。我们有:

$$AX^{-1} = \begin{bmatrix} A\boldsymbol{x}_1 & A\boldsymbol{x}_2 \end{bmatrix} = X^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \lambda \boldsymbol{x}_1 & \lambda \boldsymbol{x}_2 \end{bmatrix}$$

也就是:

$$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1, \quad A\boldsymbol{x}_2 = \lambda_1 \boldsymbol{x}_2$$

我们称 λ_1, λ_2 是 A 的特征值 (Eigenvalue),而 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 是相应的特征向量 (Eigenvector)。下面首先尝试给出如下的定义:

定义 1.1（特征值和特征向量，尝试定义）

令 A 是一个 $n \times n$ 的矩阵，λ 是一个实数，$x \in \mathbb{R}^n$ 是一个非零的 n 维向量。如果：

$$Ax = \lambda x.$$

则称 λ 是 A 的特征值 (Eigenvalue)，x 是 λ 对应的特征向量 (Eigenvector)。

一些例子 下面通过一些例子来理解上述概念。

例 1.2（斐波那契矩阵的特征值和特征向量）

先来回顾前面斐波那契上的例子：

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

由前面讨论：

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5+\sqrt{5}}{10} \\ \frac{\sqrt{5}}{5} \end{bmatrix} = \frac{1+\sqrt{5}}{2} \begin{bmatrix} \frac{5+\sqrt{5}}{10} \\ \frac{\sqrt{5}}{5} \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5-\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} \end{bmatrix} = \frac{1-\sqrt{5}}{2} \begin{bmatrix} \frac{5-\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} \end{bmatrix}$$

从而对于矩阵 A：

1. 其有两个特征值：$\lambda_1 = \frac{1+\sqrt{5}}{2}$, $\lambda_2 = \frac{1-\sqrt{5}}{2}$，对应的特征向量为：$x_1 = \begin{bmatrix} \frac{5+\sqrt{5}}{10} \\ \frac{\sqrt{5}}{5} \end{bmatrix}$，$x_2 = \begin{bmatrix} \frac{5-\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} \end{bmatrix}$.
2. 回顾之前的计算：

$$AX^{-1} = A \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = X^{-1}\Lambda$$

例 1.3（对角矩阵的特征值和特征向量）

考察对角矩阵：

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

显然我们有：

$$Ae_i = \lambda_i e_i, \quad \text{这里} e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

从而 Λ 的特征值为 $\lambda_1, \cdots, \lambda_n$，每个 λ_i 对应的特征向量为 e_i。

例 1.4 (投影矩阵的特征值和特征向量)

现在来考察投影矩阵，令 A 是一个 $m \times n$ 且 $\text{rank}(A) = n$ 的矩阵，则投影至 $\mathbf{C}(A)$ 的投影矩阵 P 可以表示为:

$$P = A(A^{T}A)^{-1}A^{T}$$

注意到 $P\mathbf{x}$ 是 $\mathbf{x}$ 到 $\mathbf{C}(A)$ 上的投影，所以 $P\mathbf{x} = \lambda\mathbf{x}$ 只有两种情况:

1. $\mathbf{x} \in \mathbf{C}(A)$，此时 $P\mathbf{x} = 1\mathbf{x} = \mathbf{x}$
2. $\mathbf{x} \in \mathbf{C}(A)^{\perp}$ 即 $\mathbf{x} \in \mathbf{N}(A^{T})$，此时 $P\mathbf{x} = 0\mathbf{x} = \mathbf{0}$

特征值为 0 的情况 假设 $n \times n$ 的矩阵 A 存在特征值 0，即存在一个非零的 $\mathbf{x} \in \mathbb{R}^{n} \setminus \{\mathbf{0}\}$ 满足:

$$A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$$

这说明 $A\mathbf{x} = \mathbf{0}$ 有非零解，即 $\dim(\mathrm{N}(A)) \geqslant 1$，进一步有:

引理 1.5

下述是等价的:

1. A 具有特征值 0。
2. $\mathbf{N}(A) \neq Z (= \{\mathbf{0}\})$。
3. $\text{rank}(A) < n$。
4. $\det(A) = 0$.

特征值和特征向量的几何解释 下面从几何的角度来理解一下特征值的意义。令 A 是一个 $n \times n$ 的矩阵，将其看作一个函数$f_A : \mathbb{R}^{n} \to \mathbb{R}^{n}$，即:

$$f_A(\mathbf{x}) = A\mathbf{x}$$

则当 $\mathbf{x}$ 是 A 的特征向量，即 $A\mathbf{x} = \lambda\mathbf{x}$ 时我们有:

$$f_A(\mathbf{x}) = \lambda\mathbf{x}$$

这意味着f_A 不改变 $\mathbf{x}$ 的方向!

另一方面，考察矩阵: $A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$，有:

$$A\begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \quad A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

从而 $1, \frac{1}{2}$ 是其两个特征值，$\begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 是相应的特征向量，注意到这两个向量是线性无关的，意味

着任何一个向量都可以表示成其线性组合，比如：

$$\begin{bmatrix}0.8\\0.2\end{bmatrix}=\begin{bmatrix}0.6\\0.4\end{bmatrix}+0.2\begin{bmatrix}1\\-1\end{bmatrix}$$

从而：

$$A^{100}\begin{bmatrix}0.8\\0.2\end{bmatrix}=(1)^{100}\begin{bmatrix}0.6\\0.4\end{bmatrix}+0.2(\frac{1}{2})^{100}\begin{bmatrix}1\\-1\end{bmatrix}\approx\begin{bmatrix}0.6\\0.4\end{bmatrix}$$

这也给出了 $A^n x$ 的一种**估计方式**。

特征值的计算　进一步回顾上述例子中特征值的计算，实际是就是：

$$Ax=\lambda x \quad \Longrightarrow \quad (\lambda E-A)x=0 \text{ 有非零解}$$

因此有下列定理：

定理 1.6

令 A 是 $n\times n$ 的矩阵，λ 是 A 的特征值当且仅当：

$$\det(\lambda E-A)=0$$

证明：

$$\begin{aligned}\lambda\text{是 } A \text{ 的特征值} &\Longleftrightarrow Ax=\lambda x \text{ 有非零解}\\ &\Longleftrightarrow (\lambda E-A)x=0 \text{ 有非零解}\\ &\Longleftrightarrow \operatorname{rank}(\lambda E-A)<n\\ &\Longleftrightarrow \det(\lambda E-A)=0\end{aligned}$$

□

将 λ 看成一个变量，则对于 $n\times n$ 的矩阵 A：

$$f(\lambda)=\det(\lambda E-A)$$

是一个**关于 λ 的 n 次多项式**，我们将其称为矩阵 A 的**特征多项式**：

定义 1.7（特征多项式 (Characteristic Polynomial)）

令 A 是 $n\times n$ 的矩阵，则称 $f(\lambda)=\det(\lambda E-A)$ 是 A 的特征多项式 (Characteristic Polynomial).

不难有如下结论：

陈述 1.8

令 A 是 $n\times n$ 的矩阵，则其特征多项式 $f(\lambda)$ 的零点即为 A 的特征值。

我们依旧通过一些例子来理解**特征多项式**：

1. 对于矩阵 $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$，其特征多项式为:

$$\det(\lambda E - A) = \begin{vmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - \lambda - 1$$

   该多项式的两个零点为:

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$$

2. 对角矩阵: $\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$ 的特征多项式为:

$$(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

   其对应的零点为:

$$\lambda_1, \cdots, \lambda_n (\text{可能重复})$$

消元法会改变特征值! 高斯-若尔当消元法是课程中一个重要的工具，但下面的例子表明，该过程会改变矩阵的特征值:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \quad \Longrightarrow U = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$$

我们有:

1. $f_A(\lambda) = (1-\lambda)(6-\lambda) - 6 = \lambda(\lambda - 7) \Longrightarrow \lambda_1 = 0,\ \lambda_2 = 7.$
2. $f_U(\lambda) = (1-\lambda)(0-\lambda) - 0 = \lambda(\lambda - 1) \Longrightarrow \lambda_1 = 0,\ \lambda_2 = 1.$

旋转矩阵的特征值和特征向量 现在来关注一个特殊的例子: 旋转矩阵。考察如下的矩阵:

$$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix}$$

其特征多项式为:

$$\det(\lambda E - A) = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1$$

其两个零点为:

$$\lambda_1 = i,\ \lambda_2 = -i$$

从而 Q 有两个复数特征值，其对应的特征向量为:

$$Q \begin{bmatrix} 1 \\ i \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$Q \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}$$

也是在$\mathbb{C}^2$上的向量!

上述例子说明了定义1.1的不足之处，即即使考虑的是实数矩阵 A，其特征值和特征向量也可能是复数的。我们需要对定义1.1进行修正:

定义 1.9（特征值和特征向量）

令 A 是一个 $n \times n$ 的矩阵，$\lambda \in \mathbb{C}$，$\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$。如果:

$$A\mathbf{x} = \lambda\mathbf{x}$$

则称 λ 是 A 的特征值 (Eigenvalue)，$\mathbf{x}$ 是 λ 对应的特征向量 (Eigenvector)。

不难验证，上述的所有讨论在定义1.9下依旧成立，因此并不需要花费篇幅修改之前的内容。现在可以总结如下计算特征值和特征向量的方法:

1. 计算 A 的特征多项式，即 $\lambda E - A$ 的行列式 $\det(\lambda E - A)$.
2. 计算 $\det(\lambda E - A)$ 的根，我们一共会得到 n 个特征值（可能重复）。这使得 $\lambda E - A$ 变成一个奇异矩阵 (rank$(A) < n$).
3. 对每个 λ，通过解方程 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ 来获取 λ 对应的特征向量 $\mathbf{x}$.

注 1.10

在一般意义下:

1. 矩阵可以有复数。
2. 向量空间是可以通过复数扩充的，即数乘允许 $c \in \mathbb{C}$ 进行运算。
3. 甚至复数域的结论会更加优美。

但在本课程中:

1. 出于课程的面向对象和内容，我们只考虑实数的矩阵。
2. 但是我们依旧会讨论到复数的特征值和特征向量。

特征值的数目　本节的最后，我们来讨论一下特征值的数目。由陈述1.8，我们知道特征值的数目和特征多项式的零点数目是相同的。注意到**代数基本定理**:

定理 1.11（Fundamental Theorem of Algebra）

对于任一的 $a_1, \cdots, a_n \in \mathbb{C}$，我们有

$$x^n + a_1x^{n-1} + \cdots a_{n-1}x + a_n = (x - b_1)(x - b_2)\cdots(x - b_n)$$

这里 $b_1, \cdots b_n \in \mathbb{C}$，即任一单变元的 n 次复系数多项式恰好有 n 个复数根（可重复）。

得益于定义1.7，不难得出：

推论 1.12

令 A 是一个 $n \times n$ 的矩阵，则在计算重根的意义下，A 恰好有 n 个 $\mathbb{C}$ 中的特征值。

例 1.13

考察矩阵

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

其特征多项式为 $f(\lambda) = \lambda^2(\lambda^2 + 1)$，注意到 $f(\lambda) = 0$ 一共有 4 个解：

$$\lambda_1 = \lambda_2 = 0,\ \lambda_3 = i,\ \lambda_4 = -i$$

这里 0 是一个重根，在计算重复次数以后可以得到其一共有 4 个特征值。

2 对角化矩阵

回顾上节斐波那契数列中提到的等式1和等式2：

$$\begin{aligned}
\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X' \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} (X')^{-1} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X' \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix} (X')^{-1} \right) \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}
\end{aligned}$$

中间的对角矩阵 Λ 极大的简化了我们对 $\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k$ 的计算，我们把这过程称为对角化 (Diagonalization)，下面给出正式的定义：

定义 2.1 (Diagonalization)

一个方阵 A 是可对角化的 (Diagonalizable)，如果其存在一个可逆矩阵 X 和对角矩阵 Λ 使得：

$$A = X\Lambda X^{-1}, \text{或者等价的 } \Lambda = X^{-1}AX$$

推论 2.2

若 $A = X\Lambda X^{-1}$，其中 $X = \text{diag}(\lambda_1, \cdots, \lambda_n)$，则对于任意的 $k \geqslant 1$ 我们有：

$$A^k = X\Lambda^k X^{-1} = X \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} X^{-1}$$

我们也称这是矩阵的**谱分解**。

例 2.3

考察矩阵 $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$，我们有：

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

其中 $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ 是可逆矩阵，并且：

$$\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$A = X\Lambda X^{-1}$ 可以理解成 $AX = X\Lambda$，从而 A 可以对角化说明的是：

1. X 的每一列都应该是 A 的特征向量。
2. X 是可逆的，所以 A 需要有 n 个线性无关的特征向量。

也即下述定理：

定理 2.4

$n \times n$ 的矩阵可对角化当且仅当 A 有 n 个线性无关的特征向量。

证明： 假设 A 有 n 个特征值 $\lambda_1, \cdots, \lambda_n$(可以重复)，对应的特征向量 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 是线性无关的：

$$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1, \cdots, A\boldsymbol{x}_n = \lambda_n \boldsymbol{x}_n$$

定义下列矩阵：

$$\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n) = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}, \quad X = \begin{bmatrix} \boldsymbol{x}_1, \cdots, \boldsymbol{x}_n \end{bmatrix}$$

则有:

1. $AX = A\begin{bmatrix}x_1, \cdots, x_n\end{bmatrix} = \begin{bmatrix}Ax_1, \cdots, Ax_n\end{bmatrix} = \begin{bmatrix}\lambda_1 x_1, \cdots, \lambda_n x_n\end{bmatrix} = \begin{bmatrix}x_1, \cdots, x_n\end{bmatrix}\Lambda = X\Lambda$.
2. $\text{rank}(X) = n$，从而 X 可逆，即：$A = X\Lambda X^{-1}$。

另一方面，定义下列矩阵:

$$\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n) = \begin{bmatrix}\lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k\end{bmatrix}, \quad X = \begin{bmatrix}x_1, \cdots, x_n\end{bmatrix}$$

容易验证:

1. 对于每个 x_i，有 $Ax_i = \lambda_i x_i$。
2. $x_1, \cdots, x_n$ 是线性无关的。

$\square$

一些例子 还是通过一些例子来进行理解。

1. 回顾之前提到的矩阵:

$$A = \begin{bmatrix}2 & -1 \\ -1 & 2\end{bmatrix} = \begin{bmatrix}1 & 1 \\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & 3\end{bmatrix}\begin{bmatrix}\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2}\end{bmatrix}$$

其特征多项式为 $f_A(\lambda) = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$，分别计算其对应的特征向量为:

(1) $\lambda = 1$ 时解方程 $(A - E)x = 0$ 得一个其特征向量为 $\begin{bmatrix}1 \\ 1\end{bmatrix}$.

(2) $\lambda = 3$ 时解方程 $(A - 3E)x = 0$ 得一个其特征向量为 $\begin{bmatrix}1 \\ -1\end{bmatrix}$.

2. 下列矩阵:

$$A = \begin{bmatrix}1 & -1 \\ 1 & -1\end{bmatrix}, \quad B = \begin{bmatrix}0 & 1 \\ 0 & 0\end{bmatrix}, \quad C = \begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}$$

不可对角化。

- A 的特征值为 $\lambda_1 = \lambda_2 = 0$，只有一个特征向量。
- B 的特征值为 $\lambda_1 = \lambda_2 = 0$，只有一个特征向量。
- C 的特征值为 $\lambda_1 = \lambda_2 = 1$，只有一个特征向量。

注 2.5

注意到矩阵 C 是可逆的，因此可逆和可对角化并没有直接的关系。

(1) 矩阵是否可逆取决于是否有零特征值 ($\lambda = 0$)。

(2) 矩阵是否可对角化取决于是否有 n 个特征向量。

下述定理给出了 n 个线性无关的特征向量的一个充分条件：

定理 2.6

令 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 是 A 的 k 个特征向量，并且其对应的特征值两两不相同。则 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 是线性无关的。

从而 n 个不同的特征值蕴含了可对角化的结论：

推论 2.7

如果 $n \times n$ 的矩阵 A 有 n 个不同的特征值，则 A 是可对角化的。

下面开始对定理2.6进行证明，其直观理解是，当 λ 互不相等的时候，下属两个等式不可能相等：

$$A\mathbf{x} = \lambda \mathbf{x} = c_1\lambda\mathbf{x}_1 + \cdots c_k\lambda\mathbf{x}_k$$
$$A(c_1\mathbf{x}_1 + \cdots c_k\mathbf{x}_k) = c_1\lambda_1\mathbf{x}_1 + \cdots c_k\lambda_k\mathbf{x}_k$$

证明： [定理2.6的证明.] 对 k 做归纳法，$k = 1$ 的时候是显然的。

假设 $k \geqslant 2$，并且：

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = \mathbf{0}$$

下面证明 $c_1 = c_2 = \cdots = c_k = 0$，反设结论不成立，即存在 $c_i \neq 0$，即：

$$\mathbf{x}_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\mathbf{x}_j$$

从而有：

$$\lambda_i\mathbf{x}_i = A\mathbf{x}_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}A\mathbf{x}_j = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\lambda_j\mathbf{x}_j$$

注意到，由归纳假设：

$$\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1} \cdots \mathbf{x}_k \text{ 是线性无关的}$$

从而：

1. 如果 $\lambda_i = 0$，则有 $\frac{c_j}{c_i}\lambda_j = 0$，从而 $\frac{c_j}{c_i} = 0$，$\mathbf{x}_i = \mathbf{0}$，矛盾。
2. 如果 $\lambda_i \neq 0$，则有：

$$\mathbf{x}_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\mathbf{x}_j \text{ 和 } \mathbf{x}_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\frac{\lambda_j}{\lambda_i}\mathbf{x}_j$$

注意到 $\lambda_j \neq \lambda_i$，所以存在不全为 0 的 $d_1, \cdots, d_{i-1}, d_{i+1}, \cdots, d_k$ 满足：

$$d_1\boldsymbol{x}_1 + \cdots + d_{i-1}\boldsymbol{x}_{i-1} + d_{i-1}\boldsymbol{x}_{i-1} + \cdots + d_k\boldsymbol{x}_k = \boldsymbol{0}$$

与 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1} \cdots \boldsymbol{x}_k$ 是线性无关的矛盾。

□

相似矩阵 再来考虑一下对角化的形式：

$$A = X\Lambda X^{-1}$$

当特征向量组成的矩阵 X 改变的时候，我们得到了无数个不同的矩阵。

- 但这些矩阵的**特征值都是相同的**。

更一般的，对于

$$A = BCB^{-1}$$

其中 B 是可逆矩阵，即使 C 不可以对角化，我们也可以得到跟 C 具有相同特征值的一大类矩阵。我们称这样一个关系叫相似 (similar)：

定义 2.8（相似矩阵（Similar Matrix）

给定矩阵 A, B，如果存在可逆矩阵 P，使得：

$$A = PBP^{-1}$$

则称矩阵 A 和 B 是相似的。

引理 2.9

相似矩阵 A 和 B 的特征值相同。

证明： 假设 $B\boldsymbol{x} = \lambda\boldsymbol{x}$，则我们有：

$$A(P\boldsymbol{x}) = PBP^{-1}(P\boldsymbol{x}) = PB\boldsymbol{x} = \lambda P\boldsymbol{x}$$

□

引理 2.10

相似关系是一个等价关系。

证明：

1. 自反性：$A = PAP^{-1}$，其中 $P = E$.
2. 对称性：$A = PBP^{-1} \Longrightarrow B = P^{-1}AP$.
3. 传递性：$A = PBP^{-1}, B = QCR^{-1} \Longrightarrow A = PQCR^{-1}P^{-1}$.

□

相似矩阵的直观理解 我们先补充一些线性变换上的概念。这一部分的定理具体证明都会在正式介绍线性变换的时候完成，现在只是为了帮助大家建立起一个概念。假设一个线性空间 V 的两组基为 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $\mathbf{f}_1,\dots,\mathbf{f}_n$，则存在 $c_{ij}\in\mathbb{R}$, $i,j\in[n]$ 满足:

$$\begin{aligned}
\mathbf{f}_1 &= c_{11}\mathbf{e}_1 + c_{21}\mathbf{e}_2 + \cdots + c_{n1}\mathbf{e}_n \\
\mathbf{f}_2 &= c_{12}\mathbf{e}_1 + c_{22}\mathbf{e}_2 + \cdots + c_{n2}\mathbf{e}_n \\
&\cdots \\
\mathbf{f}_n &= c_{1n}\mathbf{e}_1 + c_{2n}\mathbf{e}_2 + \cdots + c_{nn}\mathbf{e}_n
\end{aligned}$$

记:

$$\begin{pmatrix}\mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_n\end{pmatrix} = \begin{pmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n\end{pmatrix}\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{bmatrix} \tag{3}$$

则上式右边的矩阵被称作基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 到 $\mathbf{f}_1,\dots,\mathbf{f}_n$ 的**过渡矩阵**。当 $V=\mathbb{R}^n$ 时等式3可以转换成**矩阵的形式**:

$$\begin{bmatrix}\mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n\end{bmatrix}\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{bmatrix}$$

而对于 V 上的一个元素 $\mathbf{u}$，假设存在 $c_1,\dots c_n$ 满足:

$$\mathbf{u} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n$$

则称 $\mathbf{u}$ 在基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的**坐标为** $(c_1,\dots,c_n)$，也用列向量 $(c_1,\dots,c_n)^T$ 表示。我们有如下的定理:

定理 2.11

令一个线性空间 V 的两组基为 $\bar{\mathbf{e}}=\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ 和 $\bar{\mathbf{f}}=\{\mathbf{f}_1,\dots,\mathbf{f}_n\}$，$M$ 是 $\bar{\mathbf{e}}$ 到 $\bar{\mathbf{f}}$ 的过渡矩阵。则对于任意的 $\mathbf{u}\in V$，令其在 $\bar{\mathbf{e}}$ 下的坐标为 $(c_1,\dots,c_n)^T$，在 $\bar{\mathbf{f}}$ 下的坐标为 $(d_1,\dots,d_n)^T$，则有:

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n\end{bmatrix} = M\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n\end{bmatrix}.$$

为了叙述清楚相似矩阵的本质，我们还需要引入线性变换在不同基下的表示。设 V 是一个线性空间，$T:V\to W$ 是一个线性变换，$\bar{\mathbf{e}}=\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ 和 $\bar{\mathbf{f}}=\{\mathbf{f}_1,\dots,\mathbf{f}_m\}$ 是 V 和 W 上的一组基。则意味着

对对于 $i \in [n]$, $j \in [m]$, 存在 $a_{ij} \in \mathbb{R}$ 使得:

$$
\begin{aligned}
T(\boldsymbol{v}_1) &= a_{11}\boldsymbol{w}_1 + \cdots + a_{1m}\boldsymbol{w}_m \\
T(\boldsymbol{v}_2) &= a_{21}\boldsymbol{w}_1 + \cdots + a_{2m}\boldsymbol{w}_m \\
&\vdots \\
T(\boldsymbol{v}_n) &= a_{m1}\boldsymbol{w}_1 + \cdots + a_{nm}\boldsymbol{w}_m
\end{aligned}
$$

则称下列矩阵:

$$
A_T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}
$$

为线性变换 T 在基 $\bar{\boldsymbol{e}}$ 和 $\bar{\boldsymbol{f}}$ 下的矩阵。特别的，当 $V = W$ 时并且 $\bar{\boldsymbol{e}} = \bar{\boldsymbol{f}}$ 时，称 A_T 是**基 $\bar{\boldsymbol{e}}$ 下线性变换 T 对应的矩阵**。

下列定理给出了线性变换在不同基下的矩阵与相应坐标之间的联系:

定理 2.12

令 $\boldsymbol{v} \in \mathbb{V}$ 和 $\boldsymbol{w} \in \mathbb{W}$。若 $\boldsymbol{x} \in \mathbb{R}^n$ 是 $\boldsymbol{v}$ 在基 $\bar{\boldsymbol{v}}$ 下的坐标，$\boldsymbol{y} \in \mathbb{R}^m$ 是 $T(\boldsymbol{v})$ 在基 $\bar{\boldsymbol{w}}$ 下的坐标，则我们有:

$$
T(\boldsymbol{v}) = \boldsymbol{w} \quad \Longleftrightarrow \quad \boldsymbol{y} = A_T\boldsymbol{x}
$$

例 2.13

我们这里通过一个 $\mathbb{R}^2$ 上的例子来简单介绍一下，考虑如下的两组基:

$$
\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \ \text{和} \boldsymbol{f}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \boldsymbol{f}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}
$$

我们分别用 $\bar{\boldsymbol{e}}$ 和 $\bar{\boldsymbol{f}}$ 来表示这两组基，则 $\bar{\boldsymbol{e}}$ 到 $\bar{\boldsymbol{f}}$ 的过渡矩阵 M 满足:

$$
\begin{bmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} M \qquad \Longrightarrow \qquad \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} M
$$

从而该过渡矩阵 M 为:

$$
M = \left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix}
$$

注意到对于 $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ 有:

$$
\begin{bmatrix} 5 \\ 6 \end{bmatrix} = 5\boldsymbol{e}_1 + \boldsymbol{e}_2 = 5\boldsymbol{f}_1 + \frac{11}{3}\boldsymbol{f}_2
$$

即其在基 $\bar{\boldsymbol{e}}$ 下的坐标为 $(5,1)$，在基 $\bar{\mathbf{f}}$ 下的坐标为 $(5,\frac{11}{3})$，并且有:

$$\begin{bmatrix}5\\1\end{bmatrix}=\begin{bmatrix}1&0\\-2&3\end{bmatrix}\begin{bmatrix}5\\\frac{11}{3}\end{bmatrix}$$

再考虑如下 $\mathbb{R}^2$ 上的线性变换 T:

$$\mathsf{T}\boldsymbol{e}_1=2\boldsymbol{e}_1,\ \mathsf{T}\boldsymbol{e}_2=3\boldsymbol{e}_2$$

则 T 在 $\bar{\boldsymbol{e}}$ 上对应的矩阵为:

$$A_{\mathsf{T}}=\begin{bmatrix}2&0\\0&3\end{bmatrix}$$

另一方面，注意到:

$$\mathsf{T}\mathbf{f}_1=\mathsf{T}(\boldsymbol{e}_1-2\boldsymbol{e}_2)=\mathsf{T}\boldsymbol{e}_1-2\mathsf{T}\boldsymbol{e}_2=2\boldsymbol{e}_1-6\boldsymbol{e}_2=\begin{bmatrix}2\\-4\end{bmatrix}=2\mathbf{f}_1-\frac{2}{3}\mathbf{f}_2$$

$$\mathsf{T}\mathbf{f}_2=\mathsf{T}(3\boldsymbol{e}_2)=3\mathsf{T}\boldsymbol{e}_2=9\boldsymbol{e}_2=\begin{bmatrix}0\\9\end{bmatrix}=3\mathbf{f}_2$$

从而 T 在 $\bar{\mathbf{f}}$ 上对应的矩阵为:

$$A'_{\mathsf{T}}=\begin{bmatrix}2&0\\-\frac{2}{3}&3\end{bmatrix}$$

现在回到相似矩阵，考察如下两个矩阵:

$$Q=\begin{bmatrix}2&0\\0&3\end{bmatrix},\ Q'=\begin{bmatrix}1&1\\1&0\end{bmatrix}\begin{bmatrix}2&0\\0&3\end{bmatrix}\begin{bmatrix}0&1\\1&-1\end{bmatrix}=\begin{bmatrix}3&-1\\0&2\end{bmatrix}$$

由定义 Q 和 Q' 是相似的，考虑如下两组基:

$$\boldsymbol{e}_1=\begin{bmatrix}1\\0\end{bmatrix},\ \boldsymbol{e}_2=\begin{bmatrix}0\\1\end{bmatrix}\ 和\mathbf{f}_1=\begin{bmatrix}0\\1\end{bmatrix},\ \mathbf{f}_2=\begin{bmatrix}1\\-1\end{bmatrix}$$

给定 $\mathbf{u}=(5,4)=5\boldsymbol{e}_1+4\boldsymbol{e}_2=9\mathbf{f}_1+5\mathbf{f}_2$，则有:

$$Q\begin{bmatrix}5\\4\end{bmatrix}=\begin{bmatrix}10\\12\end{bmatrix},\quad Q'\begin{bmatrix}9\\5\end{bmatrix}=\begin{bmatrix}22\\10\end{bmatrix}$$

从而我们有:

$$\begin{bmatrix}\boldsymbol{e}_1&\boldsymbol{e}_2\end{bmatrix}\begin{bmatrix}10\\12\end{bmatrix}=10\boldsymbol{e}_1+12\boldsymbol{e}_2=\begin{bmatrix}10\\12\end{bmatrix}=22\mathbf{f}_1+10\mathbf{f}_2=\begin{bmatrix}\mathbf{f}_1&\mathbf{f}_2\end{bmatrix}\begin{bmatrix}22\\10\end{bmatrix}$$

这说明，当 Q 和 Q' 分别代表了在 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 和 $\mathbf{f}_1,\mathbf{f}_2$ 这两组基下的线性变换的时候，其实表示的是同一个线性变换，即：**两个相似矩阵本质上是在不同基下的相同线性变换!** 我们将其表述为如下的定理:

定理 2.14

令 $\mathbb{V}$ 是一个 n 维的向量空间，$\overline{v}$ 和 $\overline{v'}$ 是 $\mathbb{V}$ 的两组基。考虑 $\mathbb{V} \longrightarrow \mathbb{V}$ 的一个线性变换 T，令：

- M 是 $\overline{v}$ 到 $\overline{v'}$ 的过渡矩阵。
- A 是在基 $\overline{v}$ 下 T 对应的矩阵。
- B 是在基 $\overline{v'}$ 下 T 对应的矩阵。

则有：

$$B = M^{-1}AM.$$

即同一个线性变换在**不同基下表示的矩阵是相似的**。

例 2.15

在例2.13中，我们有：

$$\begin{bmatrix} 2 & 0 \\ -\frac{2}{3} & 3 \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix}\right)^{-1} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix},$$

即：$A'_T = M^{-1}A_TM$。