MATH10003.01 高等线性代数 Fall 2025

Lecture 1 复数与复变函数 — 2025.09.09

教授：邵美悦

Scribe: 路人甲

目录

1 记号约定

在本讲义中，我们将使用以下记号约定：

1. 实数集记为 $\mathbb{R}$，复数集记为 $\mathbb{C}$。
2. 设 A 是一个矩阵，则 $A^\top$ 表示矩阵 A 的转置矩阵，A^* 表示矩阵 A 的共轭转置矩阵，亦即先转置后共轭或先共轭后转置。

此外，酉矩阵的定义如下：

定义 1.1 (酉矩阵)**.** *设* $U \in \mathbb{C}^{n\times n}$*，若*

$$U^*U = UU^* = I,$$

则称矩阵 U *为一个酉矩阵。*

酉矩阵是正交矩阵在复数域上的推广。在本讲中，我们将会频繁地使用以上概念。

2 复数与复矩阵

2.1 复数的矩阵定义

在高中时，我们已经学习过复数的基本知识。我们将复数定义为形如 $x + y\mathrm{i}\,(x, y \in \mathbb{R})$ 的数，其中 i 是虚数单位，满足 $\mathrm{i}^2 = -1$。然后，我们定义了复数的加法和乘法运算法则：与实数的加法和乘法运算法则相同。此外，我们还定义了复数的共轭、模和辐角：对于复数 $z = x + y\mathrm{i}$，其共轭定义为 $\bar{z} = x - y\mathrm{i}$，表示复数在复平面上的对称点；其模定义为 $|z| = \sqrt{x^2 + y^2}$，表示复数在复平面上与原点的距离；其辐角定义为复数所表示的向量 $\overrightarrow{OP} = (x, y)$ 与正实轴的夹角。

但是，这种定义方式并没有指出复数的本质。我们的定义只指出了复数的形，但还有一些关键的问题并没有回答清楚。例如，

1. 为什么我们要定义一个虚数单位 i，满足 $\mathrm{i}^2 = -1$？不定义这样一个数行不行？
2. 为什么复数的加法和乘法运算法则与实数相同？运算规律不是我们想怎么定义就怎么定义的。历史上就曾有数学家尝试过定义其他的数系，如四元数、八元数等，它们就并不满足乘法交换律。

我们先来回顾一下历史，解释一下历史上数学家们为什么要引入复数。复数的引入是为了求解某些在实数域 $\mathbb{R}$ 上只能求出一个解的三次方程。例如，方程 $x^3 - 15x - 4 = 0$ 在实数域上只有一个根，但通过引入复数，我们可以找到另外两个解。具体地，我们可以将方程重写为

$$(x-4)(x^2+4x+1) = 0.$$

通过解这个方程，我们可以得到一个实数解 $x = 4$，以及两个复数解 $x = -2+\sqrt{3}\mathrm{i}$ 和 $x = -2-\sqrt{3}\mathrm{i}$。这表明复数的引入使得我们能够在更广泛的范围内求解方程。但是，这个理由并不充分。我们为什么一定要让方程能够求解呢？我们明明可以不去理会那些在实数域上无解的方程——方程 $\frac{1}{x} = 0$ 在实数域上就没有解，我们也没有因此引入新的数。我们又凭什么为了让二次方程、三次方程有解就引入复数呢？

我们先不回答上面的问题。下面我们给出复数的新的定义方式，它或许能够更好地揭示复数的本质。

定义 2.1 (复数的矩阵定义)**.** *我们定义复数集 $\mathbb{C}$ 为*

$$\mathbb{C} = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \right\}$$

首先，我们来证明一下复数对加法和乘法是封闭的。对任意两个复数 $\begin{bmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{bmatrix}$ 和 $\begin{bmatrix} \alpha_2 & -\beta_2 \\ \beta_2 & \alpha_2 \end{bmatrix}$，

它们的和为

$$\begin{bmatrix}\alpha_1 & -\beta_1\\ \beta_1 & \alpha_1\end{bmatrix}+\begin{bmatrix}\alpha_2 & -\beta_2\\ \beta_2 & \alpha_2\end{bmatrix}=\begin{bmatrix}\alpha_1+\alpha_2 & -(\beta_1+\beta_2)\\ \beta_1+\beta_2 & \alpha_1+\alpha_2\end{bmatrix}\in\mathbb{C},$$

它们的乘积为

$$\begin{bmatrix}\alpha_1 & -\beta_1\\ \beta_1 & \alpha_1\end{bmatrix}\begin{bmatrix}\alpha_2 & -\beta_2\\ \beta_2 & \alpha_2\end{bmatrix}=\begin{bmatrix}\alpha_1\alpha_2-\beta_1\beta_2 & -(\alpha_1\beta_2+\beta_1\alpha_2)\\ \alpha_1\beta_2+\beta_1\alpha_2 & \alpha_1\alpha_2-\beta_1\beta_2\end{bmatrix}\in\mathbb{C}.$$

按这种定义方式，复数就自然地继承了矩阵的运算性质：复数像矩阵一样，满足加法交换律、加法结合律、乘法结合律、分配律。此外，由于上面的矩阵的特殊结构，我们还可以验证复数满足乘法交换律：

$$\begin{bmatrix}\alpha_1 & -\beta_1\\ \beta_1 & \alpha_1\end{bmatrix}\begin{bmatrix}\alpha_2 & -\beta_2\\ \beta_2 & \alpha_2\end{bmatrix}=\begin{bmatrix}\alpha_1\alpha_2-\beta_1\beta_2 & -(\alpha_1\beta_2+\beta_1\alpha_2)\\ \alpha_1\beta_2+\beta_1\alpha_2 & \alpha_1\alpha_2-\beta_1\beta_2\end{bmatrix}=\begin{bmatrix}\alpha_2 & -\beta_2\\ \beta_2 & \alpha_2\end{bmatrix}\begin{bmatrix}\alpha_1 & -\beta_1\\ \beta_1 & \alpha_1\end{bmatrix}$$

这就自然地解释了为什么复数继承了实数的运算规律。

在中学阶段，我们将复数 $z=\alpha+\mathrm{i}\beta$ 的共轭定义为 $\bar{z}=\alpha-\mathrm{i}\beta$. 当时我们并没有解释为什么要这样定义。现在，按照矩阵的定义方式，我们可以将复数 z 的共轭表示为

$$\bar{z}\leftrightarrow\begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}.$$

这正好是矩阵 z 的转置矩阵。因此，复数的共轭实际上就是对应矩阵的转置。从矩阵的角度来看，复数的共轭定义是非常自然的。

这种定义方式也使得复数的几何意义变得非常清晰。事实上，矩阵

$$\begin{bmatrix}\alpha & -\beta\\ \beta & \alpha\end{bmatrix}$$

是一种特殊的线性变换：旋转变换和伸缩变换的复合。如果我们提取出系数 $\rho=\sqrt{\alpha^2+\beta^2}$，并令

$$\cos\theta=\frac{\alpha}{\rho},\quad \sin\theta=\frac{\beta}{\rho}\quad(\theta\in[0,2\pi)),$$

那么上面的矩阵可以写成

$$\begin{bmatrix}\alpha & -\beta\\ \beta & \alpha\end{bmatrix}=\rho\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

其中，矩阵 $\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$ 表示将一个向量逆时针旋转 θ 角度的线性变换，而系数 ρ 则表示将一个向量按原点放大 ρ 倍的线性变换。因此，复数可以看作是对二维实数空间 $\mathbb{R}^2$ 上的旋转和伸缩变换的抽象表示。在这个角度下，复数的模 $|z|$ 就是缩放因子，而辐角 $\arg z$ 就是旋转角。

此外，我们还可以引入复数的指数形式。对于复数 z，我们可以证明

$$z = \rho e^{\mathrm{i}\theta}, \quad \rho = |z|, \theta = \arg(z).$$

这个结论可以直接从 Taylor 级数展开式中得到。

现在，我们就可以解释复数乘法的几何意义了。对于两个复数的乘法

$$\begin{bmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{bmatrix} \begin{bmatrix} \alpha_2 & -\beta_2 \\ \beta_2 & \alpha_2 \end{bmatrix},$$

这个乘法对应着两个线性变换的复合。由于复数对应的线性变换是伸缩变换和旋转变换的复合，因此复数的乘法实际上对应着两个伸缩变换和两个旋转变换的复合。两个伸缩变换的复合仍然是一个伸缩变换，两个旋转变换的复合仍然是一个旋转变换。于是，复数乘法的结果仍然对应着一个伸缩变换和旋转变换。这也从另一个角度解释了复数对乘法的封闭性。

现在我们来看看虚数单位 i 在这种定义方式下的表示。我们可以将虚数单位 i 表示为矩阵

$$\mathrm{i} \leftrightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

而这个矩阵正好表示将一个向量逆时针旋转 90° 的线性变换。我们可以验证这个矩阵的平方为

$$\mathrm{i}^2 \leftrightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I,$$

其中 I 是 2×2 的单位矩阵。这说明，i^2 实际上对应着将一个向量旋转 180° 的线性变换——这正好是将一个向量取反。这也解释了为什么 $\mathrm{i}^2 = -1$.

在线性代数中，我们常常使用分块矩阵的技术。现在，我们来研究一种特殊结构的分块矩阵：

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix},$$

其中 A 和 B 是任意的 $n \times n$ 实矩阵。我们可以将这个分块矩阵看作是复数矩阵的推广——它可以看作是复矩阵 $A + \mathrm{i}B$ 的实嵌入。

此外，我们可以发现，矩阵中只需要确定两个元素 α 和 β，就可以唯一确定一个复数。这表明，复数的矩阵定义其实是有着信息的冗余的。为避免信息冗余，我们还可以将矩阵定义为一个有序对 (α, β)，定义其加法满足 $(\alpha_1, \beta_1) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2)$，定义其乘法满足 $(\alpha_1, \beta_1)(\alpha_2, \beta_2) = (\alpha_1\alpha_2 - \beta_1\beta_2, \alpha_1\beta_2 + \beta_1\alpha_2)$。我们可以验证这样的有序对对加法和乘法是封闭的，并且满足加法交换律、加法结合律、乘法结合律、分配律和乘法交换律。

这种定义方式与我们在高中时学习的复数定义是等价的。事实上，我们可以分别建立从复数 $x + y\mathrm{i}$ 到形如 $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$ 的矩阵和到有序对 (x, y) 的双射映射。因此这三种定义方式是等价的。

从有序对定义出发，我们很容易能看到复数加法的几何意义：复数的加法等价于二维实数空间 $\mathbb{R}^2$ 上的向量加法。

2.2 复数、复矩阵与多项式方程

虽然矩阵定义存在着信息的冗余，但是我们可以发现这种冗余实际上是有意义的，它使得我们可以将复数的运算与线性代数中的矩阵运算联系起来，从而利用线性代数的工具来研究复数及其相关的数学对象。

我们之前还剩下一个问题没有解决：复数和多项式方程有什么关系？为什么我们要引入复数来求解多项式方程呢？代数基本定理回答了这个问题。

定理 2.2 (代数基本定理)**.** *设 $p(z)$ 是一个非常数的单变量复系数多项式，即*

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0,$$

其中 $a_n, a_{n-1}, \ldots, a_0 \in \mathbb{C}$ 且 $a_n \neq 0$. 则存在一个复数 $z_0 \in \mathbb{C}$ 使得 $p(z_0) = 0$.

定理 2.2 表明，复数域 $\mathbb{C}$ 是一个代数闭域，即所有非常数的单变量多项式在复数域上都有根。定理的证明比较复杂，通常需要使用拓扑学或复分析的方法。在这里，我们不打算给出这个定理的证明，而将它看作一个基本事实。

这个定理有一个等价形式：

定理 2.3 (代数基本定理的等价形式)**.** *一元 n 次复多项式恰好有 n 个复根（计重数）。*

证明. 充分性显然。

必要性：设 $p(z)$ 是一个一元 n 次复多项式。根据代数基本定理，$p(z)$ 至少有一个复根 z_1.

因此，我们可以将 $p(z)$ 分解为

$$p(z) = (z - z_1) q(z),$$

其中 $q(z)$ 是一个一元 $n-1$ 次复多项式。我们可以对 $q(z)$ 重复上述过程，直到将 $p(z)$ 分解为 n 个一次因子的乘积：

$$p(z) = (z - z_1)(z - z_2) \cdots (z - z_n),$$

其中 $z_1, z_2, \ldots, z_n$ 是 $p(z)$ 的复根（计重数）。这就证明了一元 n 次复多项式恰好有 n 个复根（计重数）。 □

定理 2.3 解释了为什么我们要引入复数来求解多项式方程：在复数域上，多项式方程总是有解的。

现在我们从矩阵的角度来看这个定理。特征值问题是线性代数的一个重要问题。我们知道，一个 n 阶复方阵 A 对应着一个线性变换 $T : \mathbb{C}^n \to \mathbb{C}^n$. 在这种线性变换下，存在一个非零向量 $v \in \mathbb{C}^n$，使得 $T(v) = \lambda v, \lambda \in \mathbb{C}$，亦即 $Av = \lambda v$. 这也就是说，在线性变换下，存在一个非零向量 v 只被伸缩而不被旋转。这个 λ 就被称为线性变换 T 和矩阵 A 的一个**特征值**，而 v 就是对应的**特征向量**。

那么，为什么一定存在这样的 λ 和 v 呢？我们可以将方程 $Av = \lambda v$ 重写为

$$(\lambda I - A)v = 0,$$

其中 I 是 $n \times n$ 的单位矩阵。由于 $v \neq 0$，因此矩阵 $\lambda I - A$ 必须是奇异矩阵，即 $\det(\lambda I - A) = 0$. 而特征多项式 $p(\lambda) = \det(\lambda I - A)$ 是 λ 的一个 n 次多项式，因此根据代数基本定理，$p(\lambda) = 0$ 至少存在一个复数解 λ. 根据代数基本定理的等价形式，$p(\lambda) = 0$ 恰好有 n 个复数解（计重数），因此线性变换 T 和矩阵 A 恰好有 n 个特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$（计重数）。

那么，一个 n 阶多项式方程 $p(\lambda) = 0$ 到底怎样对应于一个 n 阶复方阵 A 呢？从矩阵 A 出发，我们只需求其特征多项式 $p(\lambda) = \det(\lambda I - A)$ 即可将其与多项式方程联系起来。那么，从多项式方程出发，我们能不能找到一个对应的复方阵 A 呢？我们可以通过 Frobenius 友阵来实现这一点。

定义 2.4 (Frobenius 友阵). *设 $p(\lambda) = \lambda^n + \cdots + a_1\lambda + a_0$ 是一个首系数为 1 的一元 n 次复系数多项式。则与 $p(\lambda)$ 对应的* ***Frobenius 友阵****定义为*

$$\begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}$$

上面的定义中，我们将多项式的首系数规定为 1，这是为了简化表示。如果多项式的首系数不是 1，我们可以先将多项式除以首系数，使其变为首系数为 1 的多项式，然后再应用定义 2.4。

或许读者会对 Frobenius 友阵的定义感到困惑：凭什么这个矩阵就对应着给定的多项式方程呢？常规的教材或许会直接验证 Frobenius 友阵的特征多项式正是 $p(\lambda)$，但这显然无法令我们挑剔的读者满意。下面，我们将完全抛开定义 2.4，从方程 $p(\lambda) = 0$ 出发，推导出其对应的复方阵。

证明. 假设矩阵 A 对应的特征多项式是 $p(\lambda)$，这意味着

$$\det(\lambda I - A) = p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

仅凭这一个条件，想要构造出符合条件的 A 是非常困难的。邵老师教导我们，题目做不出来就加一些条件。我们需要引入一些额外的假设来简化问题。

1. 设 $p(\lambda) = 0$ 的 n 个根分别为 $\omega_1, \ldots, \omega_n \in \mathbb{C}$. 根据定理 2.3，这个假设是合理的。

2. 我们如果有特征向量，反解出矩阵就简单多了。因此，我们假设矩阵 A 对应于 ω_k 的特征向量为

$$v_k = \begin{bmatrix} \omega_k^{n-1} \\ \omega_k^{n-2} \\ \vdots \\ 1 \end{bmatrix}$$

现在，我们来推导矩阵 A 的形式。在矩阵 $\omega I - A$ 的作用下，v 会被变换为零向量。那么，我们该如何构造出这样的 $\omega I - A$ 呢？我们考虑从下到上地构造这个矩阵。我们不妨从矩阵 ωI 出发，逐步将其变换为矩阵 $\omega I - A$. 首先，ωI 对 v 的最后一个分量的作用是伸缩 ω 倍。这样，原本的 1 就被变成了 ω. 那么，我们就需要让 Av 的最后一个分量也是 ω，从而使得 $(\omega I - A)v$ 的最后一个分量变为零。

在下面的推导中，我们将广泛运用一种观点：向量 By 的第 i 个分量等于矩阵 B 的第 i 个行向量与向量 y 的内积。

根据这一观点，Av 的最后一个分量等于矩阵 A 的最后一个行向量与 v 的内积。我们希望这个内积是 ω. 注意到：v 的倒数第二个分量是 ω，那么 A 的最后一个行向量可以自然地构造为

$$[0, 0, \ldots, 1, 0].$$

这样，Av 的最后一个分量就等于 $1 \cdot \omega = \omega$.

接下来，我们来考虑 $(\omega I - A)v$ 的倒数第二个分量。类似地，ωI 对 v 的倒数第二个分量的作用是伸缩 ω 倍，使得原本的 ω 变成了 ω^2. 因此，我们需要让 Av 的倒数第二个分量也是 ω^2，从而使得 $(\omega I - A)v$ 的倒数第二个分量变为零。类似于上面的推导，我们可以将矩阵 A 的倒数第二个行向量构造为

$$[0, 0, \ldots, 1, 0, 0].$$

想来读者已经能猜到了，我们可以将矩阵 A 的第 i 个行向量构造为

$$e_{i-1}^\top = [0, 0, \ldots, 1, 0, \ldots, 0],$$

这种构造在 $i \geq 2$ 时是行之有效的，读者可自行验证。但对于 $i = 1$，e_{i-1} 就没有定义了，Python 会提示你 `IndexError!` 因此，我们需要单独考虑矩阵 A 的第一个行向量。类似地，ωI 对 v 的第一个分量的作用是伸缩 ω 倍，使得原本的 ω^{n-1} 变成了 ω^n. 因此，我们需要让 Av 的第一个分量也是 ω^n，从而使得 $(\omega I - A)v$ 的第一个分量变为零。别忘了，我们还有一个条件没用：多项式方程 $p(\lambda) = 0$ 的根 ω 满足

$$\omega^n + a_{n-1}\omega^{n-1} + \cdots + a_1\omega + a_0 = 0.$$

因此，我们可以将矩阵 A 的第一个行向量构造为

$$[-a_{n-1}, -a_{n-2}, \ldots, -a_1, -a_0].$$

这样，矩阵 A 的第一个行向量与 v 的内积就等于

$$-a_{n-1}\omega^{n-1}-a_{n-2}\omega^{n-2}-\cdots-a_1\omega-a_0=\omega^n.$$

综上所述，我们得到了矩阵 A 的完整形式：

$$A=\begin{bmatrix}-a_{n-1} & -a_{n-2} & \cdots & -a_0\\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0\end{bmatrix}.$$

它的特征多项式正是 $p(\lambda)$. 如果 ω 是 $p(\lambda)=0$ 的根，那么 ω 就是矩阵 A 的一个特征值，而对应的特征向量就是

$$v=\begin{bmatrix}\omega^{n-1}\\ \omega^{n-2}\\ \vdots\\ 1\end{bmatrix}.$$

请读者注意，这个推导是充分性的，并不能保证必要性——事实上，对应于多项式方程 $p(\lambda)$ 的复方阵并不唯一。如果我们对 Frobenius 友阵进行相似变换，得到的矩阵仍然对应着同一个多项式方程。这个结论可以利用 Jordan 标准型来证明，我们在此不再赘述。 □

现在我们来看看方程 $\lambda^2+(0\cdot\lambda)+1=0$ 的 Frobenius 友阵：

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

可以发现，它正是复数 i 的矩阵表示！而 i 也正是这个矩阵的特征方程一个解。这也解释了为什么我们要引入复数来求解多项式方程，以及为什么要引入矩阵来表示复数。在复数域上，多项式方程总是有解的，而它的解又可以看作是对应的复方阵的特征值，我们便可以自然地将复数和矩阵联系起来。这种代数上的精妙结构使得复数和矩阵在数学中有着广泛的应用。

下面我们补充复方阵的一个性质，它与代数基本定理和特征值密切相关：

命题 2.5 (Schur 分解)**.** *设 A 是一个 n 阶复方阵，则 A 酉相似于一个上三角阵，且对角元的次序可任意指定。*

证明. 我们使用数学归纳法和分析法来证明这个命题。

当 $n=1$ 时，命题显然成立。

假设当 $n=k$ 时，命题成立。现在我们来证明当 $n=k+1$ 时，命题也成立。

我们想要证明：$\exists P \in \mathbb{C}^{(k+1)\times(k+1)}$ 是酉矩阵，使得 P^*AP 是一个上三角阵 T，即证 $AP = PT$. 我们将 P 和 T 分块表示为

$$P = \begin{bmatrix} p_1, p_2, \ldots, p_{k+1} \end{bmatrix}, \quad T = \begin{bmatrix} t_1 & * \\ 0 & T_1 \end{bmatrix},$$

其中 $p_i \in \mathbb{C}^{(k+1)}, t_1 \in \mathbb{C}, T_2 \in \mathbb{C}^{k\times k}$.

于是，$AP = PT$ 可写为

$$A \begin{bmatrix} p_1, p_2, \ldots, p_{k+1} \end{bmatrix} = \begin{bmatrix} Ap_1, Ap_2, \ldots, Ap_{k+1} \end{bmatrix} \begin{bmatrix} t_1 & * \\ 0 & T_1 \end{bmatrix}.$$

提取出第一项：

$$Ap_1 = t_1 p_1.$$

这就说明 t_1 需要是 A 的一个特征值，p_1 需要是对应的特征向量。

根据代数基本定理，A 至少有一个特征值 t_1 和对应的特征向量 p_1. 因此，我们总是可以找出一个特征对，用以填充 P 的第一列 p_1 和 T 的第 $(1,1)$ 元素 t_1.

接下来，我们需要构造出 P 的剩余列 $p_2, \ldots, p_{k+1}$ 和 T 的剩余部分 $*$ 和 T_1.

将 p_1 正规化为单位向量，并扩张为 $\mathbb{C}^n$ 上的一组标准正交基 $\{q_1, \ldots, q_n\}$. 记 $Q = \begin{bmatrix} q_1, q_2, \ldots, q_n \end{bmatrix}$. 因此 Q 是一个酉矩阵。

计算

$$\begin{aligned} Q^*AQe_1 &= Q^*Aq_1 \\ &= Q^*t_1q_1 \\ &= t_1Q^*q_1 \\ &= t_1e_1. \end{aligned}$$

最后一个等号是因为我们有如下恒等式

$$Q^*Q = I.$$

因此，Q^*AQ 具有如下分块形式

$$Q^*AQ = \begin{bmatrix} t_1 & * \\ 0 & A_1 \end{bmatrix},$$

其中 $A_1 \in \mathbb{C}^{k\times k}$.

由归纳假设，存在酉矩阵 $U \in \mathbb{C}^{k\times k}$，使得 $U^*A_1U = T_1$ 是一个上三角阵，且对角元的次序可任意指定。记

$$P = Q \begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix}.$$

则 P 是一个酉矩阵，且

$$P^*AP = \begin{bmatrix}1 & 0\\0 & U^*\end{bmatrix} Q^*AQ \begin{bmatrix}1 & 0\\0 & U\end{bmatrix} = \begin{bmatrix}t_1 & *\\0 & T_1\end{bmatrix},$$

显然，由于 q_1 是可以从特征向量中任意选取的，而 T_1 的对角元次序也可以任意指定，因此整个上三角阵的对角元次序也可以任意指定。这就完成了证明。

事实上，T 是上三角阵，其对角元正是其特征值。由于 Schur 分解是酉相似变换，相似变换不改变特征值，因此 T 的对角元也是矩阵 A 的特征值。 □

3 复变函数

在中学阶段，我们接触的主要是实变量实值函数，即函数的定义域和值域都是实数集 $\mathbb{R}$. 但是，在大学阶段，我们将接触到更为广泛的函数类型，其中之一就是复变函数，即函数的定义域和值域都是复数集 $\mathbb{C}$. 在本节中，我们补充一些关于复变函数的基本概念，为后续有关矩阵函数的内容做铺垫。

定义 3.1 (复变函数)**.** *设 $D \subseteq \mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个子集，$f: D \to \mathbb{C}$ 是一个映射，则称 f 为一个***复变函数***，其定义域为 D.*

我们可以将复变函数 $f: D \to \mathbb{C}$ 看作是一个二维向量值函数 $f: D_{xy} \to \mathbb{R}^2$，即

$$f(x, y) = \begin{bmatrix}u(x, y)\\v(x, y)\end{bmatrix}$$

其中 $D_{xy} = \{(x, y) \mid x + \mathrm{i}y \in D\}$，$u(x, y)$ 和 $v(x, y)$ 分别是 $f(x + \mathrm{i}y)$ 的实部和虚部。

对于一个函数，我们通常关心它的连续性和可微性。在本节中，我们不会严格地建立复变函数上的分析理论，而只是给出一些基本的定义和结论。如果对这些内容感兴趣，可以参考相关的复分析教材。

定义 3.2 (复变函数的极限)**.** *设 $f: D \to \mathbb{C}$ 是一个复变函数，$z_0 \in D$. 如有 $\forall \varepsilon > 0$，$\exists \delta_\varepsilon > 0$, s.t. $|f(z) - A| < \varepsilon$，$\forall 0 < |z - z_0| < \delta_\varepsilon$，则称 A 是 $f(z)$ 在 z_0 处的极限，记为 $\lim\limits_{z \to z_0} f(z) = A$.*

定义 3.3 (复变函数的连续性)**.** *设 $f: D \to \mathbb{C}$ 是一个复变函数，$z_0 \in D$. 若 $\lim\limits_{z \to z_0} f(z) = f(z_0)$，则称 $f(z)$ 在 z_0 处连续。*

定义 3.4 (复变函数的可微性)**.** *设 $f: D \to \mathbb{C}$ 是一个复变函数，$z_0 \in D$. 若极限*

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

存在，则称 $f(z)$ 在 z_0 处可微，称该极限为 $f(z)$ 在 z_0 处的导数，记为 $f'(z_0)$.

需要注意，复变函数的可微性比实变函数更加严格。事实上，一个复变函数在某点处可微，意味着它在该点处的导数与取值路径无关。下面我们以一道题为例，来说明这里所说的“取值路径无关”是什么意思。

例 3.5. 设

$$f(z)=\begin{cases}\exp(-z^{-4}) & z\neq 0\\ 0 & z=0\end{cases}.$$

证明 $f(z)$ 在 $z=0$ 处不可微。

证明. 首先判断 $f(z)$ 在 $z=0$ 处是否连续。

$$\lim_{z\to 0} f(z)=\lim_{z\to 0}\exp(-z^{-4})=0=f(0).$$

有些同学可能在这里就认为 $f(z)$ 在 $z=0$ 处连续了（笔者也是），但事实上这只说明了 $f(z)$ 在 $z=0$ 处沿实轴方向连续。我们还需要检查 $f(z)$ 在 $z=0$ 处沿其他路径的连续性。

$$\begin{aligned}\lim_{z\to 0} f(z) &\overset{z=\rho e^{\mathrm{i}\theta}}{=} \lim_{\rho\to 0} f(\rho e^{\mathrm{i}\theta})\\ &= \lim_{\rho\to 0}\exp(-\rho^{-4}e^{-4\mathrm{i}\theta})\end{aligned}$$

我们取 $\theta=\frac{\pi}{4}$，有 $e^{-4\mathrm{i}\theta}=-1$，因此有

$$\lim_{\rho\to 0} f(\rho e^{\mathrm{i}\theta})=\lim_{\rho\to 0}\exp(\rho^{-4})=+\infty.$$

这就说明 $f(z)$ 在 $z=0$ 处极限与路径相关，因此不连续，从而也不可微。 □

定义 3.6 (全纯函数)**.** 设 $D\subseteq\mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个子集，$f:D\to\mathbb{C}$ 是一个复变函数。若 $f(z)$ 在 D 上的每一点都可微，则称 $f(z)$ 是 D 上的一个**全纯函数**。

定义 3.7 (解析函数)**.** 设 $D\subseteq\mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个子集，$f:D\to\mathbb{C}$ 是一个复变函数。若 $f(z)$ 在 D 上的每一点都可以展开成幂级数，则称 $f(z)$ 是 D 上的一个**解析函数**。

定理 3.8. 全纯函数和解析函数是等价的。

定理 3.8 的证明比较复杂，通常需要使用复分析的方法。在这里，我们不打算给出这个定理的证明，而将它看作一个基本事实。定理 3.8 表明，复变函数在区域 D 上可导，就意味着它在区域 D 上可以展开成幂级数。

前面提到，我们可以用二维向量值映照的观点来看待复变函数。基于这种观点，我们可以得到一个重要的方程组。

定理 3.9 (Cauchy–Riemann 方程组). 设 $f : D \to \mathbb{C}$ 是一个复变函数，$D \subseteq \mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个子集，$f(z) = u(x, y) + \mathrm{i}v(x, y), z = x + \mathrm{i}y$. 若 $f(z)$ 在 D 上可微，则 $u(x, y)$ 和 $v(x, y)$ 在 $D_{xy} = \{(x, y) \mid x + \mathrm{i}y \in D\}$ 上可微，且满足 *Cauchy–Riemann* 方程组

$$\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \end{cases}.$$

证明. 我们先给一个极其不严谨的证明。

$f(z)$ 在 D 上可微，这意味着它对 x 和 y 都可微。因此，$u(x, y)$ 和 $v(x, y)$ 在 D_{xy} 上可微。而向量 $\begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$ 在 (x, y) 处的 Jacobi 矩阵为

$$D \begin{bmatrix} u \\ v \end{bmatrix} (x, y) = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}.$$

显然它应该具有复数的结构，即有

$$\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \end{cases}.$$

然后我们再给一个比较严谨的证明。

由于 $f(z)$ 在 D 上可微，它要求 Δz 无论从哪个方向趋近 0，极限值都要相同。

于是，我们可以选取两个特殊的方向，建立相等关系。

1. 当 Δz 沿实轴方向趋近 0 时，有

$$\begin{aligned} f'(z) &= \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \\ &= \lim_{\Delta x \to 0} \frac{f((x + \Delta x) + \mathrm{i}y) - f(x + \mathrm{i}y)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{u(x + \Delta x, y) + \mathrm{i}v(x + \Delta x, y) - u(x, y) - \mathrm{i}v(x, y)}{\Delta x} \\ &= \frac{\partial u}{\partial x} + \mathrm{i}\frac{\partial v}{\partial x}. \end{aligned}$$

2. 当 Δz 沿虚轴方向趋近 0 时，有

$$\begin{aligned}
f'(z) &= \lim_{\Delta z\to 0}\frac{f(z+\Delta z)-f(z)}{\Delta z}\\
&= \lim_{\Delta y\to 0}\frac{f(x+\mathrm{i}(y+\Delta y))-f(x+\mathrm{i}y)}{\mathrm{i}\Delta y}\\
&= \lim_{\Delta y\to 0}\frac{u(x,y+\Delta y)+\mathrm{i}v(x,y+\Delta y)-u(x,y)-\mathrm{i}v(x,y)}{\mathrm{i}\Delta y}\\
&= -\mathrm{i}\lim_{\Delta y\to 0}\frac{u(x,y+\Delta y)-u(x,y)}{\Delta y}-\mathrm{i}^2\lim_{\Delta y\to 0}\frac{v(x,y+\Delta y)-v(x,y)}{\Delta y}\\
&= -\mathrm{i}\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}.
\end{aligned}$$

由于 $f'(z)$ 与 Δz 的取值路径无关，因此有

$$\frac{\partial u}{\partial x}+\mathrm{i}\frac{\partial v}{\partial x}=-\mathrm{i}\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}.$$

由实部虚部对应相等，我们有

$$\begin{cases}\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}\\ \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}\end{cases}.$$

这就得到了 Cauchy–Riemann 方程组。 □

定理 3.8 还可以解释我们之前在级数中遇到的一些现象。例如，$\frac{1}{1+x^2}$ 展开成幂级数

$$\frac{1}{1+x^2}=\sum_{n=0}^{+\infty}(-1)^n x^{2n}, |x|<1,$$

它的收敛半径为 1，在实数域上，我们只能通过比值判别法对这一现象进行解释。但在复数域上，我们还有另外一种更接近问题本质的解释。这是因为 $\frac{1}{1+z^2}$ 在复数域上有两个奇点 $z=\mathrm{i}$ 和 $z=-\mathrm{i}$，也就是说它在这两点不可导，也就不是全纯函数，进而不是解析函数，因此在整个复数域上没有幂级数展开。这两个点到原点的距离都是 1. 因此，$\frac{1}{1+z^2}$ 在复数域上展开成幂级数的收敛半径为 1.

我们自然会想到：该如何计算复变函数的导数呢？以下我们分别以两类函数为例，来说明如何计算复变函数的导数。

第一类函数是只含有 z，不含有 $\bar{z}$ 的函数。

例 3.10. 设 $f(z)=z^2$. 求 $f(z)$ 的导数.

解.

$$
\begin{aligned}
f'(z) &= \lim_{h\to 0}\frac{f(z+h)-f(z)}{h}\\
&= \lim_{h\to 0}\frac{(z+h)^2-z^2}{h}\\
&= \lim_{h\to 0}\frac{2zh+h^2}{h}\\
&= \lim_{h\to 0}(2z+h)\\
&= 2z.
\end{aligned}
$$

显然，对于只含有 z 的函数，我们可以直接使用实变函数的导数公式——它们在形式上是完全相同的。

第二类函数是同时含有 z 和 $\bar{z}$ 的函数。

例 3.11. 设 $f(z)=|z|^2=z\bar{z}$. 求 $f(z)$ 的导数。

解.

$$
\begin{aligned}
f'(z) &= \lim_{h\to 0}\frac{f(z+h)-f(z)}{h}\\
&= \lim_{h\to 0}\frac{(z+h)(\bar{z}+\bar{h})-z\bar{z}}{h}\\
&= \lim_{h\to 0}\frac{z\bar{h}+\bar{z}h+h\bar{h}}{h}\\
&= \lim_{h\to 0}(z\frac{\bar{h}}{h}+\bar{z}+\bar{h})\\
&= \lim_{h\to 0}z\frac{\bar{h}^2}{|h|^2}+\bar{z}+\bar{h}\\
&\overset{h=|h|e^{\mathrm{i}\theta}}{=} \lim_{h\to 0}ze^{-2\mathrm{i}\theta}+\bar{z}+\bar{h}.
\end{aligned}
$$

由于 $\lim_{h\to 0}ze^{-2\mathrm{i}\theta}$ 与 θ 路径相关，因此 $\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}$ 不存在。这就说明 $f(z)$ 在任何点 z 处都不可微。

对于非数学专业的学生来说，我们只需要记住含有 $\bar{z}$ 的函数不可导即可。这一点可以利用上面的 Cauchy–Riemann 方程来证明。

定理 3.12 (Cauchy–Riemann 方程组的等价形式)**.** 设 $f: D\to\mathbb{C}$ 是一个复变函数，$D\subseteq\mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个子集，$f(z)=u(x,y)+\mathrm{i}v(x,y), z=x+\mathrm{i}y$. 若 $f(z)$ 在 D 上可微，则

$$
\frac{\partial f}{\partial\bar{z}}=0.
$$

证明. 注意到我们可以用 x 和 y 来表示 z 和 $\bar{z}$，也可以用 z 和 $\bar{z}$ 来表示 x 和 y，因此从 $\begin{bmatrix}x\\y\end{bmatrix}$ 到

$\begin{bmatrix} z \\ \bar{z} \end{bmatrix}$ 的映射是微分同胚。

因此，我们有

$$\begin{cases} u(x,y) = u \circ \begin{bmatrix} z \\ \bar{z} \end{bmatrix}(x,y) \\ v(x,y) = v \circ \begin{bmatrix} z \\ \bar{z} \end{bmatrix}(x,y) \end{cases}$$

我们先看第一个式子。由链式求导法则，我们有

$$Du(x,y) = Du(z,\bar{z})D\begin{bmatrix} z \\ \bar{z} \end{bmatrix}(x,y).$$

亦即

$$\begin{aligned} \left[\tfrac{\partial u}{\partial x}, \tfrac{\partial u}{\partial y}\right] &= \left[\tfrac{\partial u}{\partial z}, \tfrac{\partial u}{\partial \bar{z}}\right]\begin{bmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{bmatrix} \\ &= \left[\tfrac{\partial u}{\partial z} + \tfrac{\partial u}{\partial \bar{z}}, \mathrm{i}(\tfrac{\partial u}{\partial z} - \tfrac{\partial u}{\partial \bar{z}})\right]. \end{aligned}$$

同理可有

$$\left[\tfrac{\partial v}{\partial x}, \tfrac{\partial v}{\partial y}\right] = \left[\tfrac{\partial v}{\partial z} + \tfrac{\partial v}{\partial \bar{z}}, \mathrm{i}(\tfrac{\partial v}{\partial z} - \tfrac{\partial v}{\partial \bar{z}})\right].$$

由 Cauchy–Riemann 方程组，我们有

$$\begin{cases} \tfrac{\partial u}{\partial z} + \tfrac{\partial u}{\partial \bar{z}} = \mathrm{i}(\tfrac{\partial v}{\partial z} - \tfrac{\partial v}{\partial \bar{z}}) \\ \mathrm{i}(\tfrac{\partial u}{\partial z} - \tfrac{\partial u}{\partial \bar{z}}) = -(\tfrac{\partial v}{\partial z} + \tfrac{\partial v}{\partial \bar{z}}) \end{cases}.$$

我们先搁置这两个方程，来看看 $\frac{\partial f}{\partial \bar{z}}$ 的表达式。

由于 $f(z) = u(z,\bar{z}) + \mathrm{i}v(z,\bar{z})$，我们有

$$\frac{\partial f}{\partial \bar{z}} = \frac{\partial u}{\partial \bar{z}} + \mathrm{i}\frac{\partial v}{\partial \bar{z}}.$$

因此，对上面的两个方程，我们需要消去所有含 z 的项，因此将一式乘上 i，再将两式相减，立刻有

$$\frac{\partial u}{\partial \bar{z}} + \mathrm{i}\frac{\partial v}{\partial \bar{z}} = 0.$$

亦即

$$\frac{\partial f}{\partial \bar{z}} = 0.$$

容易验证反向命题也成立，因此定理得证。

这也就说明了，一个解析函数是不可能含有 $\bar{z}$ 的。

□

定义 3.13 (复积分). 设 $f : D \to \mathbb{C}$ 是一个复变函数，C 是 D 上的一条光滑曲线，C 的参数方程为 $z = z(t), t \in [a, b]$. 则称

$$\int_C f(z)\,\mathrm{d}z = \int_a^b f \circ z(t) z'(t)\,\mathrm{d}t$$

为 $f(z)$ 沿曲线 C 的**复积分**。

复积分的计算与实积分类似，只不过被积函数是复值的。我们可以将复积分看作是二维向量值映照的线积分。

定理 3.14 (Cauchy–Goursat 定理). 设 $D \subseteq \mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个单连通开子集，$f : D \to \mathbb{C}$ 是 D 上的一个全纯函数，C 是 D 上的一条分段光滑闭曲线。则

$$\int_C f(z)\,\mathrm{d}z = 0.$$

Cauchy–Goursat 定理表明，在单连通区域上，解析函数的积分与路径无关，这与实变函数的情形有很大的不同。这使得复变函数在积分计算上有很大的便利性。

定理 3.15 (Cauchy 积分公式). 设 $D \subseteq \mathbb{C}$ 是复数集 $\mathbb{C}$ 的一个单连通开子集，$f : D \to \mathbb{C}$ 是 D 上的一个全纯函数，C 是 D 上的一条分段光滑闭曲线，且 z_0 是 C 围成的区域内的一点。则

$$f(z_0) = \frac{1}{2\pi\mathrm{i}} \int_C \frac{f(z)}{z - z_0}\,\mathrm{d}z.$$

Cauchy 积分公式表明，复变函数在某点处的值可以通过该点周围的闭曲线上的积分来表示。

最后，我们给出矩阵函数的定义，结束本节的讨论。

定义 3.16 (矩阵函数). 若 Ω 是复平面上的区域，且包含 $A \in \mathbb{C}^{n\times n}$ 的所有特征值。那么对于 Ω 上的解析函数 $f(z)$，定义矩阵函数

$$f(A) = \frac{1}{2\pi\mathrm{i}} \int_{\partial\Omega} f(z)(zI - A)^{-1}\,\mathrm{d}z.$$

这个定义表明，如果两个解析函数满足 $f(z) = g(z)$，那么它们对应的矩阵函数也相等，即 $f(A) = g(A)$.

因此，在复数域成立的方程，对应地在矩阵赋范线性空间上也成立。

例如，$(\cos x)^2 + (\sin x)^2 = 1$ 对 $x \in \mathbb{C}$ 成立，根据定义，我们有

$$(\cos A)^2 + (\sin A)^2 = I$$

对任意 $A \in \mathbb{C}^{n\times n}$ 成立。