

MIT 18.06 线性代数

第二十三讲：微分方程和矩阵指数

整理自 MIT 18.06 课程及相关公开笔记

引言

本讲将特征值与特征向量应用于一阶常系数线性微分方程组的求解，通过使用特征向量矩阵对耦合的线性系统进行“解耦”，可以将复杂的微分方程化为简单的纯指数形式。同时，本讲还将引入**矩阵指数** (Matrix Exponential) e^{At} 的概念，建立微分方程解与矩阵指数之间的直接代数联系，并讨论系统稳定性的判定准则。

1 一阶微分方程组的通解

1.1 特征值的求解

考察如下的一阶常系数线性微分方程组：

$$\begin{aligned}\frac{du_1}{dt} &= -u_1 + 2u_2 \\ \frac{du_2}{dt} &= u_1 - 2u_2\end{aligned}$$

给定初始条件： $\mathbf{u}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 。将其转化为矩阵形式 $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ ，其中系数矩阵为：

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

为了追踪变量随时间的变化趋势，首先要求解矩阵 A 的特征值与特征向量，计算矩阵 A 的迹与行列式：

$$\text{tr}(A) = -1 - 2 = -3, \quad |A| = (-1)(-2) - (2)(1) = 0$$

由于行列式为 0，矩阵 A 是奇异的，因此必然有一个特征值为 $\lambda_1 = 0$ ，根据特征值之和等于矩阵的迹，可求得另一个特征值为 $\lambda_2 = -3$ 。

1.2 特征向量的计算

分别将两个特征值代入特征方程 $(A - \lambda I)\mathbf{x} = \mathbf{0}$ ，求对应的特征向量。

- 对于 $\lambda_1 = 0$ ，解方程 $A\mathbf{x} = \mathbf{0}$ ：

$$\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

- 对于 $\lambda_2 = -3$ ，解方程 $(A - (-3)I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

1.3 一般解与初始条件

一阶线性微分方程的通解由纯指数特解的线性组合构成。对于该二阶系统，其通解形式为：

$$\mathbf{u}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2$$

将特征值和特征向量代入，得到：

$$\mathbf{u}(t) = c_1 e^{0t} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

利用初始条件 $\mathbf{u}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 确定系数 c_1 与 c_2 ：

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

解得 $c_1 = \frac{1}{3}, c_2 = \frac{1}{3}$ 。因此，原方程组在给定初始条件下的解为：

$$\mathbf{u}(t) = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3} e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

解的长期行为分析：当 $t \rightarrow \infty$ 时，包含 e^{-3t} 的项迅速衰减至零，而第一项保持不变，该系统的稳态为：

$$\mathbf{u}(\infty) = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

这反映出变量 u_1 与 u_2 的分配最终达到平衡。

2 解的稳定性与系统状态

根据特征值的实数部分（Real Part，记为 $\text{Re}(\lambda)$ ），可以判断系统解 $u(t)$ 的长期性质：

- **稳定性 (Stability)：**当所有特征值的实部均满足 $\text{Re}(\lambda) < 0$ 时，解 $\mathbf{u}(t) \rightarrow \mathbf{0}$ 。
- **稳态 (Steady State)：**存在一个特征值为 0，且其余特征值的实部均小于 0。此时系统收敛到一个非零常数向量。
- **发散 (Blow Up)：**若存在任何一个特征值满足 $\text{Re}(\lambda) > 0$ ，则解将随时间发散至无穷大。

判别注意事项：对于二阶矩阵 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ ，若两个特征值实部均为负，则矩阵的迹 $a + d$ 必然为负，

但反之并不成立，例如矩阵 $\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$ 的迹为 -1 ，但它仍包含一个正特征值 1，会导致解发散；若矩阵的行列式为正、迹为负，则可以判定系统解是收敛的。

3 解耦与矩阵指数

3.1 利用特征向量矩阵解耦

在微分方程 $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ 中，矩阵 A 使得变量之间相互耦合，可以通过引入特征向量矩阵 S 来解耦此系统。设 $\mathbf{u} = S\mathbf{v}$ ，其中 S 的列是矩阵 A 的线性无关特征向量，代入原方程：

$$S \frac{d\mathbf{v}}{dt} = AS\mathbf{v} \implies \frac{d\mathbf{v}}{dt} = S^{-1}AS\mathbf{v} = \Lambda\mathbf{v}$$

由于 Λ 是对角矩阵，新系统变为完全解耦的形式：

$$\frac{dv_i}{dt} = \lambda_i v_i, \quad i = 1, 2, \dots, n$$

其解为 $v_i(t) = v_i(0)e^{\lambda_i t}$ ，即对角化向量解：

$$\mathbf{v}(t) = e^{\Lambda t}\mathbf{v}(0)$$

其中 $e^{\Lambda t}$ 为对角矩阵，对角线元素为 $e^{\lambda_i t}$ 。

3.2 矩阵指数的定义与推导

将 $\mathbf{v}(t) = e^{\Lambda t}\mathbf{v}(0)$ 代回原变量 $\mathbf{u}(t) = S\mathbf{v}(t)$ 中，可得：

$$\mathbf{u}(t) = Se^{\Lambda t}S^{-1}\mathbf{u}(0)$$

因此，定义矩阵指数为 $e^{At} = Se^{\Lambda t}S^{-1}$ ，其中：

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}, \quad e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix}$$

该公式也可以通过矩阵幂级数展开进行严格推导。根据标量指数函数的幂级数定义 $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ ，将其推广至矩阵形式：

$$e^{At} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \dots$$

利用对角化公式 $A = S\Lambda S^{-1}$ 代入上述级数：

$$\begin{aligned} e^{At} &= SS^{-1} + S\Lambda S^{-1}t + \frac{S\Lambda^2 S^{-1}}{2}t^2 + \frac{S\Lambda^3 S^{-1}}{6}t^3 + \dots \\ &= S \left(I + \Lambda t + \frac{\Lambda^2 t^2}{2} + \frac{\Lambda^3 t^3}{6} + \dots \right) S^{-1} \\ &= Se^{\Lambda t}S^{-1} \end{aligned}$$

此等式成立的前提是矩阵 A 可以被对角化（即拥有 n 个线性无关的特征向量）。使用矩阵指数的表达方式可以大大简化对微分方程系统解的结构分析。

4 高阶微分方程的降阶

对于二阶微分方程 $y'' + a_1 y' + a_0 y = 0$ ，可以通过变量代换将其转化为一阶矩阵形式。定义向量 $\mathbf{u} = \begin{bmatrix} y' \\ y \end{bmatrix}$ ，则其对时间的导数为：

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} y'' \\ y' \end{bmatrix}$$

利用原方程 $y'' = -a_1y' - a_0y$ 替换第一行，可得：

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} -a_1 & -a_0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y \end{bmatrix} = A\mathbf{u}$$

这样就成功将一个二阶微分方程转化为了一阶微分方程组，可直接套用前述的方法求解。该变换可推广至任意 k 阶微分方程，转化后的系数矩阵将是一个 $k \times k$ 的方阵，其第一行由微分方程的系数构成，次对角线上元素为 1，其余元素为 0。以 5 阶微分方程为例，设原方程形式为：

$$y^{(5)} + a_4y^{(4)} + a_3y''' + a_2y'' + a_1y' + a_0y = 0$$

将其转化为标准的一阶矩阵微分方程 $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ ：

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} y^{(5)} \\ y^{(4)} \\ y''' \\ y'' \\ y' \\ y \end{bmatrix} = \begin{bmatrix} -a_4 & -a_3 & -a_2 & -a_1 & -a_0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y^{(4)} \\ y''' \\ y'' \\ y' \\ y \end{bmatrix}$$

- 矩阵的**第一行**：由原微分方程各项的系数（取反号）构成，依次对应最高阶到最低阶的导数项。
- 矩阵的**次对角线**：元素均为 1，用于建立变量之间的连续导数关系。
- 矩阵的**其余位置**：元素均为 0。

总结

- 一阶常系数线性微分方程组的通解可以表示为特征向量与指数项 $e^{\lambda_i t}$ 的线性组合。
- 利用特征向量矩阵 S 进行**解耦**，可将原本互相耦合的线性系统简化为一组完全独立的方程。
- **矩阵指数** e^{At} 为系统解提供了统一且优雅的代数表达，其收敛性直接由特征值的实部决定。
- 任何高阶线性常微分方程都可以通过降阶转化为一阶矩阵方程，进而沿用之前的解题框架。

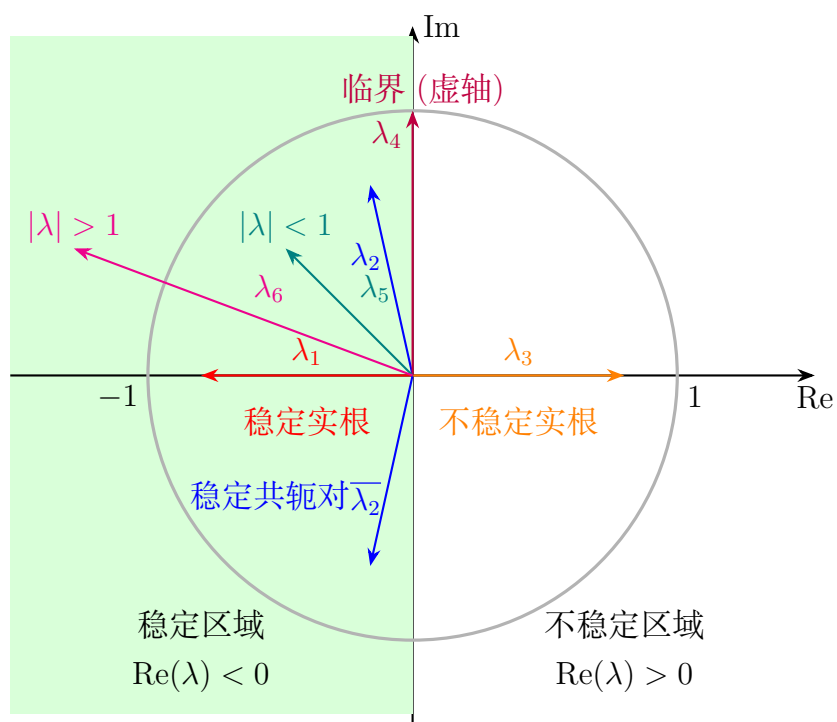


图 1: 特征值在复平面上的分布与系统稳定性的对应关系