

## 2.10 微分方程和矩阵指数

本讲将脱离一般的线性方程组，介绍一阶线性常微分方程的矩阵解法，也就是将微分方程用矩阵表示，通过“解耦”，计算出对应系数，最终得到解。这里会牵涉到 $e^{Ax}$ 计算问题，(A是矩阵)，所以也会引出幂指数是矩阵时算式的计算问题。最后扩展介绍了高阶微分方程的降阶求解方法。这讲的内容较多，计算难度大，大家要有心理准备！

### 微分方程的求解

再说求解之前，我们来说什么是微分方程！他和导函数的区别是什么！

导函数就是描述了一个函数每一个时刻变换律的函数。而微分方程就是一个把一个未知函数及其导数（或高阶导数）联系起来，或者说描述一个量或者多个量与其导数（或高阶导数）的关系的等式(当然上一个描述更准确)！你可以这样想：导数是微分方程这个“语言”中的“词汇”。微分方程用这些“词汇”(导数、函数、变量)来“造句”(建立方程)，描述世界的运行规则。

我们以一道例题来贯穿微分方程的讲解：有两个函数： $u_1(t), u_2(t)$

由他们的导函数我们得到一个**二维常微分方程组**：

$$\begin{cases} \frac{du_1}{dt} = -u_1(t) + 2u_2(t) & (\text{也可以写成 } u_1'(t) = -u_1(t) + 2u_2(t)) \\ \frac{du_2}{dt} = u_1(t) - 2u_2(t) \end{cases}$$
。然后有一个函数 $u(t)$ ，函数 $u$ 是

根据时间改变作用于 $u_1$ 和 $u_2$ 的。那么把 $u_1$ 和 $u_2$ 放在一个矩阵中，就是说 $u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$ ，那么他们的导函数总和是 $\frac{du}{dt} = Au$ （这个就是微分方程的矩阵形式）。然后有条件： $u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 。现在我们要求出函数 $u$ 的具体方程，也就是 $u_1$ 和 $u_2$ 的具体的方程！

tips：其实这就是矩阵的核心：规模化的去处理方程组，之前是一般线性方程，这里我们换成了微分方程而已！当我们把开始的微分方程组改为一般线性方程组就很好理解了！

我们先对题目信息进行分析，当时间 $t=0$ 时， $u_1(0) = 1, u_2(0) = 0$ ，那么在0时刻 $u_1$ 的导函数在 $t=0$ 时小于0，在 $t=0$ 时 $u_2$ 的导函数大于0，所以 $u_1$ 的值会上涨， $u_2$ 会下降。

然后我们由方程组可以得到： $A = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$ 。正如我们用线性代数求解差分方程一样，我们也可以用它来求解这个微分方程。

首先我们给出矩阵微分方程的通解形式： $u(t) = \sum_{i=1}^n C_i e^{\lambda_i t} x_i$ 。（在上讲求差分函数的时候我们推导了差分函数通解形式的由来，但是这里这个通解形式我们就不做研究，但其实过程是差不多一样，只不过细节上的证明会比较麻烦，感兴趣的可以去了解！）

然后我们来求矩阵的特征向量与特征值！首先我们可以看出矩阵A是不可逆的，那么他的行列式就是0！而且一定有一个特征值是0。而矩阵A的迹是-3，所以还有一个特征值是-3。然后代入矩

阵 $A-\lambda I$ 中求得两个特征向量是： $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ,  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 。由于只有两个特征值，那么通解是有两个项的，

$u(t) = C_1 e^{\lambda_1 t} x_1 + C_2 e^{\lambda_2 t} x_2 = C_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + C_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 。然后代入初始值： $u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ，得到

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

分析这个通解，我们发现随着时间  $t$  的增加，后一项  $\frac{1}{3} e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$  逐渐衰减，最后趋于 0，而前一项

$\frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  不随时间改变，这也符合我们一开始分析微分方程意义的时候  $u$  的走势。

那么  $u_1 = 2C_1 + C_2 e^{-3t} = \frac{1}{3} e^{-3t} + \frac{2}{3}$ ,  $u_2 = -\frac{1}{3} e^{-3t} + \frac{1}{3}$

到此这个题目就求解完成！

通过这道题，我们可以得到解决微分方程过程中遇到的某些特点：

(1) 当特征值都为负数的时候， $u(t)$  趋于 0。这个很好理解！

(2) 当稳态存在时，一个特征向量或者有多个=0，其余的特征向量全部<0。那么就会趋近于一个常数！

(3) 如果有任何特征值实数部分>0，则解无法收敛。

(4) 但是要注意一种特殊情况，就是特征值为复数时， $a+bi$  怎么去判断  $u(t)$  的趋势呢？答案是只有实数部分决定  $u(t)$  趋势。可知： $C_i e^{(1+2i)t} x_i = C_i e^{1t} x_i + C_i e^{2it} x_i$ 。而后者虚部只产生振荡；对“衰减还是增长”起决定作用的永远是实部。证明涉及欧几里得范数，我们略过！

## 解耦与指数矩阵 $e^{AT}$

首先先说明什么是解耦，什么是耦合！

耦合：变量之间互相牵扯，方程必须一起求解。像上面我们给的微分方程组：

$$\begin{cases} \frac{du_1}{dt} = -u_1(t) + 2u_2(t) \\ \frac{du_2}{dt} = u_1(t) - 2u_2(t) \end{cases}$$

右边  $u_2$  出现在第一个方程里， $u_1$  又出现在第二个方程里。想求  $u_1(t)$  必须先知道  $u_2(t)$ ，反之亦然——这就叫**耦合**。

解耦：**变量独立，分开解**，原本互相纠缠的方程组被转换成彼此独立的方程。

回到我们的题目，我们有  $\frac{du}{dt} = Au$ ，这个是耦合的， $u_1$  和  $u_2$  相互牵制！那么我们怎么解耦呢？还

记得我们的对角化矩阵吗，如果能够把  $A$  换成是一个对角矩阵，那么  $\begin{bmatrix} a & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & c \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ ，那么

是不是说结果矩阵  $\frac{d}{dt} \begin{bmatrix} u_1 \\ \dots \\ u_3 \end{bmatrix}$  的第一个结构就是有  $a$  和  $u_1$  决定，这样就解耦成功了！那么怎么换算出

对角矩阵呢，接下来有一个很妙的换元！

令  $u = Sv$ ， $S$  是矩阵  $A$  的特征向量矩阵。代入原方程得到

$\frac{du}{dt} = Au \rightarrow \frac{dSv}{dt} = ASv \rightarrow S \frac{dv}{dt} = ASv \rightarrow \frac{dv}{dt} = S^{-1}ASv$ 。而在上一讲我们推导过：

$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ 。所以得到：

$$\frac{d}{dt} \begin{bmatrix} u_1 \\ \vdots \\ u_3 \end{bmatrix} \rightarrow \frac{d}{dt} \begin{bmatrix} v_1 \\ \vdots \\ v_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 \\ \lambda_2 v_2 \end{bmatrix} \rightarrow \frac{dv_1}{dt} = \lambda_1 v_1 \text{ 以及 } \frac{dv_2}{dt} = \lambda_2 v_2 \text{。解耦成功了！}$$

把  $\frac{dv_1}{dt} = \lambda_1 v_1$  以及  $\frac{dv_2}{dt} = \lambda_2 v_2$  他们积分得到  $v_1(t) = e^{\lambda_1 t + C} = C' e^{\lambda_1 t}$ 。（其中常数  $C'$  由初值  $v_1(0)$  决定，所以直接写成  $v_1(0)$  即可，因为  $v_1(0) = C' e^0 = C'$ ），所以  $v_1(t) = e^{\lambda_1 t} v_1(0)$ 。同理

$$v_2(t) = e^{\lambda_2 t} v_2(0)。$$

那么把：  $\begin{cases} v_1(t) = e^{\lambda_1 t} v_1(0) \\ v_2(t) = e^{\lambda_2 t} v_2(0) \end{cases}$  矩阵化得到：  $v(t) = e^{\Lambda t} v(0)$ ，其中  $e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix}$

因为  $u = Sv$ ，所以  $u(t) = Sv(t) = Se^{\Lambda t} v(0)$ ，而  $u(0) = Sv(0) \Rightarrow v(0) = S^{-1}u(0)$ ，那么代入上式得到  $u(t) = Se^{\Lambda t} S^{-1}u(0)$ 。

而又一个定理是说： $e^{At} = Se^{\Lambda t} S^{-1}$ 。所以我们得到了： $u(t) = e^{At} u(0)$ 。

那这个定理怎么来的呢？证明如下：

首先我们先看看幂级公式（也就是大名鼎鼎的泰勒展开）：

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!}。(\text{收敛域是全体实数，不用考虑特征值问题})$$

扩展到矩阵的计算中，同样，1 代替 1，矩阵代替 x，得到：

$$e^{At} = I + At + \frac{At^2}{2} + \frac{At^3}{6} + \cdots + \frac{At^n}{n!}$$

接下来我们对角化形式化简 A，得到：

$$e^{At} = SS^{-1} + S\Lambda tS^{-1} + \frac{S(\Lambda t)^2 S^{-1}}{2} + \frac{S(\Lambda t)^3 S^{-1}}{6} + \dots$$

我们提取出  $S$  和  $S^{-1}$ ：

$$e^{At} = S(I + \Lambda t + \frac{(\Lambda t)^2}{2} + \frac{(\Lambda t)^3}{6} + \dots)S^{-1}$$

然后根据幂级公式得到：

$$e^{At} = S(e^{\Lambda t})S^{-1}$$

而可以知道：

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

那么根据幂级公式，展开后相加得到  $e^{\Lambda t}$  是：

$$\begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & e^{\lambda_n t} \end{bmatrix}$$

同样，这里的判断是否收敛与微分方程中的判别差不多，即比较  $\lambda$  的实部的绝对值与 1 的大小关系。

上面这一系列妙不可言的证明，难度很大，涉及到了很深的微积分知识！我们在上这门课前说到高中数学基础即可，尽管我在尽力的把教授讲课的内容联系为我们高中阶段说到的导数的内容，但是在这里似乎仍然不够，而且我个人对这讲内容也是云里雾里，在网上也看到许多人说这讲内容在现代工程控制里面有涉及。反正很难，很抽象，多多理解，有条件有时间的可以在上完微分

方程的课程后（mit的18.02和18.01）来复习这讲内容！重点在把握运用特征向量与特征值求解一阶微分方程即可！

## 高阶微分方程的求解

上面我们都在讲一阶微分方程组，那我们高阶的微分方程怎么求解呢，还记得我们怎么求解斐波那契数列的吗，方法类似！

有  $y'' + by' + ky = 0$

那么我们设：  $U = \begin{bmatrix} y' \\ y \end{bmatrix}$ ，得到  $U' = \begin{bmatrix} y' \\ y \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y \end{bmatrix}$ 。这样我们就将二阶微分方程化简为一阶微分方程乘上一个矩阵。同样，如果是求解一个五阶微分方程的话，我们只需要像上面那样化简，但是矩阵会变成一个  $5 \times 5$  的矩阵，接下来只要使用一阶微分方程正常求解就可以了。这是求解高阶微分方程很聪明的解法，把  $y$  的不同阶导数看出是不同的函数，这样就可以类比到一阶的微分方程组了！

## 讨论课

接下来我们来看看一个完整的高阶微分方程的求解。对于一阶（常数）微分方程组的求解我们在第一部分说到了就不再重复！把握好着两个求解过程，其实比理解上面那一大段的证明更加的重要！

我们有一个高阶微分方程：  $y'' + 2y' - y' - 2y = 0 (y = y(t))$ ，求出他的一般解，然后找出矩阵  $A$  和  $y$ ，以及指数矩阵  $\exp(At)$ 。

我们设：  $U(t) = \begin{bmatrix} y \\ y' \\ y \end{bmatrix}$ ，那么  $U'(t) = \begin{bmatrix} y' \\ y' \\ y' \end{bmatrix}$

那么  $\begin{bmatrix} y \\ y' \\ y' \end{bmatrix} = \begin{bmatrix} -2 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \\ y \end{bmatrix}$

找到矩阵  $A$  的特征值与特征向量。

$\det(A - \lambda I) = \det\left(\begin{bmatrix} -2 - \lambda & 1 & 2 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{bmatrix}\right)$ ，解得特征值是 1, -1, -2。对应得到的特征向量是：

$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$ 。

那么  $U(t) = C_1 e^t x_1 + C_2 e^{-t} x_2 + C_3 e^{-2t} x_3$ 。而看前面的矩阵可以看出  $y$  是  $U(t)$  的最后一个坐标，那么  $y = C_1 e^t + C_2 e^{-t} + C_3 e^{-2t}$ 。

那么现在我们要求解指数矩阵  $\exp(At)$ ，也就是  $e^{At}$ 。上面推导可知：  $S(e^{\Lambda t})S^{-1}$

而  $e^{At} = \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-2t} \end{bmatrix}$ 。而  $S = \begin{bmatrix} 1 & 1 & 4 \\ 1 & -1 & -2 \\ 1 & 1 & 1 \end{bmatrix}$ ，然后求出  $S$  的逆矩阵。

$$S^{-1} = \frac{1}{\det s} C^T = \frac{1}{6} \begin{bmatrix} 1 & 3 & 2 \\ -3 & -3 & 6 \\ 2 & 0 & -2 \end{bmatrix}$$

然后代入求解即可，我们这里就不细说了！

## 习题课

### 问题一

本题中的矩阵是反对称矩阵（即  $A^T = -A$ ）：

$$\frac{d}{dt} = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix} \quad \text{或写为} \quad \begin{cases} u_1' = cu_2 - bu_3 \\ u_2' = au_3 - cu_1 \\ u_3' = bu_1 - au_2 \end{cases}$$

利用定义求  $\| (t) \|^2$  的导数：

$$\| (t) \|^2 = u_1^2 + u_2^2 + u_3^2$$

这说明了向量  $(t)$  长度的变化率如何？又能说明  $(t)$  的取值范围如何？

解答：

$$\begin{aligned} \frac{d\| (t) \|^2}{dt} &= \frac{d(u_1^2 + u_2^2 + u_3^2)}{dt} \\ &= 2u_1u_1' + 2u_2u_2' + 2u_3u_3' \\ &= 2u_1(cu_2 - bu_3) + 2u_2(au_3 - cu_1) + 2u_3(bu_1 - au_2) \\ &= 0 \end{aligned}$$

这意味着  $\| (t) \|^2$  保持不变，始终等于  $\| (0) \|^2$ 。因此， $(t)$  的长度始终不变，始终位于以原点为中心、半径为  $\| (0) \|^2$  的球面上。

### 问题二

给定矩阵

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

将  $A$  写成  $S\Lambda S^{-1}$  的形式，计算  $Se^{\Lambda t}S^{-1}$  以求得矩阵指数  $e^{At}$ 。

在  $t = 0$  时，验证  $e^{At}$  及其导数的值。

解答：

矩阵  $A$  的特征值为  $\lambda_1 = 1$  和  $\lambda_2 = 3$ ，对应的特征向量为  $\mathbf{x}_1 = (1, 0)$  和  $\mathbf{x}_2 = (1, 2)$ 。由此可得：

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

利用以上矩阵计算  $e^{At}$ :

$$\begin{aligned} e^{At} &= S e^{\Lambda t} S^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{3t} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} e^t & \frac{1}{2}e^{3t} - \frac{1}{2}e^t \\ 0 & e^{3t} \end{bmatrix} \end{aligned}$$

**验证:**

当  $t = 0$  时:

$$e^{A \cdot 0} = \begin{bmatrix} e^0 & \frac{1}{2}e^0 - \frac{1}{2}e^0 \\ 0 & e^0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

计算导数:

$$\frac{d}{dt} e^{At} = \begin{bmatrix} e^t & \frac{3}{2}e^{3t} - \frac{1}{2}e^t \\ 0 & 3e^{3t} \end{bmatrix}$$

当  $t = 0$  时:

$$\left. \frac{d}{dt} e^{At} \right|_{t=0} = \begin{bmatrix} 1 & \frac{3}{2} - \frac{1}{2} \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = A$$

验证通过。

这节涉及的求导的内容比较多!