线性代数

-23 课 微分方程和矩阵指数

一、知识概要

本节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抽象，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。

二．解微分方程

解决微分方程问题重点在于其流程，我们通过一道例题来介绍本部分内容。

【例 1】

$$\frac{du_1}{dt} = -u_1 + 2u_2 \;;\quad \frac{du_2}{dt} = u_1 - 2u_2 \;;\; u(0) = \begin{bmatrix}1\\0\end{bmatrix}，求 u(t)。$$

解题流程：

解题之前首先要搞清楚这个微分方程的意义，我们看到，$u(0) = \begin{bmatrix}1\\0\end{bmatrix}$，就是说明在最初 0 时刻，$u_1 = 1$，所有的值都在$u_1$中；而此时$u_2 = 0$，但是随着时间流逝，t 增加时，我们可以看到 $\frac{du_2}{dt} > 0$。说明u_2的导数大于 0，u_2会慢慢增加，u_1慢慢减少；（或者理解为u_1中的值流向u_2）。最终达到某一状态，这需要我们计算来得到。

列出方程 $\frac{du}{dt} = Au$,其中系数矩阵 A 应综合 $\frac{du_1}{dt}$ 与 $\frac{du_2}{dt}$,写做$\begin{bmatrix}-1 & 2\\1 & -2\end{bmatrix}$ 。我们先给出通解形式：

$$\sum_{i=1}^{n} C_i e^{\lambda_i t} x_i$$

通解形式是如何得到的我们不做研究。具体验证可以将通解看做几个纯指数解的组合，随便挑一个代入验证一下即可。

解 A 矩阵的特征值与特征向量。不难得到这个矩阵有两个特征值：λ_1=0，λ_2=-3。特征向量为：$x_1 = \begin{bmatrix}2\\1\end{bmatrix}$，$x_2 = \begin{bmatrix}1\\-1\end{bmatrix}$，代入通解得到其形式如下：

$$u(t) = C_1 e^{0t}\begin{bmatrix}2\\1\end{bmatrix} + C_2 e^{-3t}\begin{bmatrix}1\\-1\end{bmatrix}$$

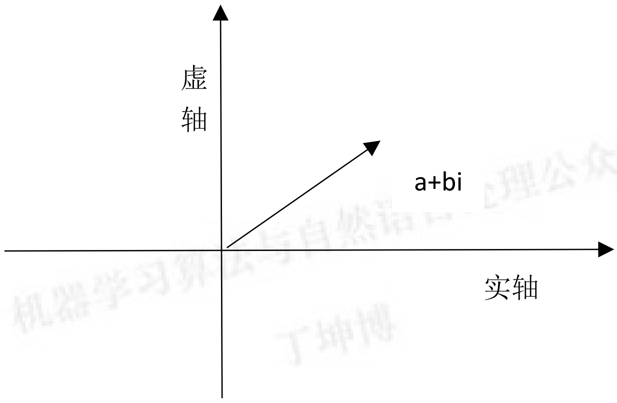
再代入初值：$u(0) = \begin{bmatrix}1\\0\end{bmatrix}$，确定$C_1$与$C_2$，最终解为：

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

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。

⊿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）特征值是负数时，u(t)趋于 0。这个特点很简单，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。我们可以画一个复数坐标系



不难看出，投影到实数轴，只有实数部分 a 决定正负性，而虚部 b 的作用是在另一条轴上指明方向，所以不影响我们的判断。

（2）稳态存在时（如【例 1】中最后 t 趋于无穷时，u 趋于一个确数），一个特征向量=0，其余的特征向量全部<0。

（3）如果有任何特征值实数部分>0，则解无法收敛。

三．解耦与e^{At}

3.1 解耦

回到【例 1】，$\frac{du}{dt} = Au$，矩阵 A 中，有u_1，u_2耦合，我们的处理，就是计算出特征值与特征向量将 A 解耦。我们设 S 是特征向量矩阵。

令 u = SV

$$\frac{du}{dt} = S\frac{dv}{dt} = Au = ASV，提取出 S\frac{dv}{dt} = ASV。$$

两边同时左乘上S^{-1}，得到：$\frac{\mathrm{dv}}{\mathrm{dt}} = S^{-1}\mathrm{ASV} = \Lambda \mathrm{V}$，得到关于 V 的对角化方程组。新方程不耦合，$\frac{\mathrm{d}v_1}{\mathrm{dt}} = \lambda_1 v_1$，$\frac{\mathrm{d}v_2}{\mathrm{dt}} = \lambda_2 v_2$以此类推。最终可得到：

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

同类型地，也有：$$\mathrm{u(t)} = e^{\mathrm{At}}\mathrm{u}(0) = \mathrm{S}e^{\Lambda \mathrm{t}}S^{-1}\mathrm{u}(0)$$

由$\frac{\mathrm{du}}{\mathrm{dt}}$ = Au 得到

?

这里就牵扯到了一个新问题，$e^{\mathrm{A}t}$和$e^{\Lambda \mathrm{t}}$是什么？

表面来看，e^{At}就是 u(t)的解，那么e^{At}为什么与 $\mathrm{S}e^{\Lambda \mathrm{t}}S^{-1}$相等呢？它们表示的是什么呢？我们接下的来重点就在介绍这些式子上。

3.2 e^{At}与 $\mathrm{S}(e^{\Lambda \mathrm{t}})S^{-1}$

首先熟悉一下幂级数公式：

$$e^{\mathrm{x}} = 1 + \mathrm{x} + \frac{x^2}{2} + \frac{x^3}{6} + \cdots\cdots + \frac{x^n}{\mathrm{n}!} \quad \cdots\cdots$$

（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，I 代替 1，矩阵代替 x，得到：

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

接下来我们对角化形式化简 A，得到

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

提出 S 和S^{-1}，得到：

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

综合幂级数公式，得到：$e^{At} = \mathrm{S}(e^{\Lambda \mathrm{t}})S^{-1}$

注意：这步的化简是有条件的，即 A 必须可以对角化，即有 n 个独立的特征向量，S^{-1}存在。

3.3 矩阵指数

上面的等式将对角矩阵与一般矩阵联系了起来，那么其中的$e^{\Lambda t}$代表着什么呢？

我们知道 $\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$,

那么$e^{\Lambda t}$就可以如下表示：

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

其主对角线上皆为$e^{\lambda_i t}$，其余位置为 0。

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λ的实部的绝对值与 1 的大小关系。

四．二阶微分方程的解

解二阶微分方程时，我们可以将它降阶处理，步骤如下：

二阶微分方程 $y'' + by' + ky = 0$

设 $u = \begin{bmatrix} y' \\ y \end{bmatrix}$，可利用 u 将上面的方程化简为：

$$u' = \begin{bmatrix} y'' \\ y' \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y \end{bmatrix}$$

这样我们就将二阶微分方程化简为了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，只不过其中的$\begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix}$会变成一个 5*5 的矩阵，类似于$\begin{bmatrix} a & b & c & d & e \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$的矩阵，其中 a，b，c，d，e 都是方程中的系数，而且主对角线上的元素下的元素都是 0。这样的矩阵将五阶微分方程转化为一阶向量方程。接下来只要使用一阶微分方程正常求解就可以了。

五、学习感悟

本节内容较多，主要目的是在实际情况下使用矩阵对角化，特征值等方法求解微分方程，给出了一种使用矩阵求解微分方程的通用规律，即高阶降阶，一阶用特征值和特征向量将原系数矩阵 A 解耦，最后得到结果。并介绍了在我们解耦 A 时使用矩阵对角化将其与特征向量联系起来运算的方法。另外介绍了判断收敛性的方法，即看特征值实部绝对值与 1 的大小关系。这些内容都是特征值与特征向量的实际应用，较为重要。