Chapter21微分方程与矩阵的指数函数

Section1.微分方程的解法

Subsection1.基础知识介绍

我们这里主要介绍一阶微分方程系统，即一阶常系数微分方程组，考虑两个变量，他们之间相互耦合，导数之间也会有联系，即:

$$\begin{cases}\frac{dx}{dt}=ax+by\\ \frac{dy}{dt}=cx+dy\end{cases}$$

或者我们写为矩阵形式:

$$\begin{bmatrix}x'\\ y'\end{bmatrix}=\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

除了$\begin{bmatrix}x\\ y\end{bmatrix}$的写法，我们也可以写成$u(t)=\begin{bmatrix}u_1\\ u_2\end{bmatrix}$的形式，二者是一个意思。

Subsection2.此种方法的解法

有方程组$\begin{cases}\frac{\mathrm{d}u_1}{\mathrm{d}t} & =-u_1+2u_2\\ \frac{\mathrm{d}u_2}{\mathrm{d}t} & =u_1-2u_2\end{cases}$，则系数矩阵是$A=\begin{bmatrix}-1 & 2\\ 1 & -2\end{bmatrix}$，设初始条件为在0时刻$u(0)=\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}$。

对于性质的分析

由于初始时刻$u(0)=\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}$，$u_1>0,u_2=0$，且关注$\frac{\mathrm{d}u_1}{\mathrm{d}t}$和$\frac{\mathrm{d}u_2}{\mathrm{d}t}$的表达式，可以发现$u_1$下降，$u_2$上升，$u_1$中的事物会流向$u_2$。

这里我们把A的特征值和特征向量依次解出来：

$$|A-\lambda I|=\begin{vmatrix}-1-\lambda & 2\\ 1 & -2-\lambda\end{vmatrix}=\lambda^2+3\lambda=0$$

求特征向量，$\lambda_1=0$时,$x_1=\begin{bmatrix}2\\ 1\end{bmatrix}$；$\lambda_2=-3$时,$x_2=\begin{bmatrix}1\\ -1\end{bmatrix}$。

得到方程组的通解为：$u(t)=c_1e^{\lambda_1t}x_1+c_2e^{\lambda_2t}x_2$，通解的前后两部分都是该方程组的纯解，即方程组的通解就是两个与特征值、特征向量相关的纯解的线性组合。我们来验证一下，比如取$u=e^{\lambda_1t}x_1$带入$\frac{\mathrm{d}u}{\mathrm{d}t}=Au$，对时间求导得到$\lambda_1e^{\lambda_1t}x_1=Ae^{\lambda_1t}x_1$，化简得$\lambda_1x_1=Ax_1$。

继续求c_1,c_2，$u(t)=c_1\cdot 1\cdot\begin{bmatrix}2\\1\end{bmatrix}+c_2\cdot e^{-3t}\cdot\begin{bmatrix}1\\-1\end{bmatrix}$，已知$t=0$时，$\begin{bmatrix}1\\0\end{bmatrix}=c_1\begin{bmatrix}2\\1\end{bmatrix}+c_2\begin{bmatrix}1\\-1\end{bmatrix}$（$Sc=u(0)$），所以$c_1=\frac{1}{3},c_2=\frac{1}{3}$。

解法总结

1.求解系数矩阵的特征值与特征向量
2.得到$u(t)=c_1e^{\lambda_1t}x_1+c_2e^{\lambda_2t}x_2$
3.利用初始条件得到c_1,c_2

Subsection3.性质与要求

- 稳态：
  经过无限的时间最终达到稳态$u(\infty)=\begin{bmatrix}\frac{2}{3}\\\frac{1}{3}\end{bmatrix}$。所以，要使得$u(t)\to 0$，则需要负的特征值，并且需要一个特征值实部为0。
- 收敛态：
  需要其中一个特征值实部为0，而其他特征值的实部皆小于0。
- 发散态：
  如果某个特征值实部大于0。上面的例子中，如果将A变为$-A$，特征值也会变号，结果发散。

如果特征值为复数，何时收敛？

如$\lambda=-3+6i$，我们来计算$\left|e^{(-3+6i)t}\right|$，其中的$\left|e^{6it}\right|$部分为$|\cos 6t+i\sin 6t|=1$，因为这部分的模为$\cos^2\alpha+\sin^2\alpha=1$，这个虚部就在单位圆上转悠。所以只有实数部分才是重要的。所以我们可以把前面的结论改为**需要实部为负数的特征值**。实部会决定最终结果趋近于0或∞，虚部不过是一些小杂音。

再进一步，我们想知道如何从直接判断任意二阶矩阵的特征值是否均小于零。对于二阶矩阵$A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$，如果矩阵稳定，则所有$\lambda<0$，矩阵的迹为$a+d=\lambda_1+\lambda_2$，迹应为负数，$\det A=\lambda_1\cdot\lambda_2$，行列式为正数。

Section2.指数矩阵e^{At}

Subsection1.引入

我们在Section1.微分方程的解法中已经解出了答案，但是我们一开始就说到，这种一阶微分方程系统的变量是互相耦合的，反应到矩阵就是A不是对角线矩阵。那么特征值和特征向量的作则就是解耦，也就是对角化（diagonalize）。

回到原方程组$\frac{\mathrm{d}u}{\mathrm{d}t}=Au$，将$u$表示为特征向量的线性组合$u=Sv$，代入原方程有$S\frac{\mathrm{d}v}{\mathrm{d}t}=ASv$，两边同乘以$S^{-1}$得$\frac{\mathrm{d}v}{\mathrm{d}t}=S^{-1}ASv=\Lambda v$。以特征向量为基，将$u$表示为$Sv$，得到关于$v$的对角化方程组，新方程组不存在耦合，此时$\begin{cases}\frac{\mathrm{d}v_1}{\mathrm{d}t} & =\lambda_1v_1\\ \frac{\mathrm{d}v_2}{\mathrm{d}t} & =\lambda_2v_2\\ \vdots & \vdots\\ \frac{\mathrm{d}v_n}{\mathrm{d}t} & =\lambda_nv_n\end{cases}$，这是一个各未知函数间没有联系的方程组，它们的解的一般形式为$v(t)=e^{\Lambda t}v(0)$，则原方程组的解的一般形式为$u(t)=e^{At}u(0)=Se^{\Lambda t}S^{-1}u(0)$，我想$e^{\Lambda t}$还是好理解一点的，但是$e^{At}$是啥?

Subsection2.定义

像$e^x=1+\frac{x^2}{2}+\frac{x^3}{6}+\cdots$一样，将$e^{At}$展开成幂级数的形式为：

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

其实指数矩阵都可以类似于实数的函数来做幂级数的展开，类似几何级数：

$$(I-At)^{-1}=I+At+(At)^2+(At)^3+\cdots$$

但是就如同实数的幂函数求和一样，e^x的级数总是收敛，而$\frac{1}{1-x}$需要$|x|<1$才收敛。第一个级数对我们而言比第二个级数好，因为第一个级数总会收敛于某个值，所以e^x总会有意义，而第二个级数需要A特征值的绝对值小于1（因为涉及矩阵的幂运算）。

那么我们注意e^{At}长什么样子:

$$e^A=\sum_{n=0}^{\infty}\frac{A^n}{n!}=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{bmatrix}+\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix}+\frac{1}{2!}\begin{bmatrix}a_{11}^2+a_{12}a_{21}&a_{11}a_{12}+a_{12}a_{22}\\a_{21}a_{11}+a_{22}a_{21}&a_{21}a_{12}+a_{22}^2\\\vdots&\vdots\\a_{n1}a_{11}+a_{nn}a_{21}&a_{n1}a_{12}+a_{nn}a_{22}\end{bmatrix}$$

这个矩阵形式相当复杂，而再看$e^{\lambda t}$的样子:

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

有了$u(t)=Se^{\Lambda t}S^{-1}u(0)$，我们将$e^{At}$变为对角矩阵就是因为对角矩阵简单、没有耦合。

> **Note**
>
> 这里有一个要求，就是如果A是可以相似对角化的，那么就会有一个等式:
>
> $$e^{At}=Se^{\lambda t}S^{-1}$$

再来看矩阵的稳定性可知，所有特征值的实部均为负数时矩阵收敛，此时对角线上的指数收敛为0。如果我们画出复平面，则要使微分方程存在稳定解，则特征值存在于复平面的左侧（即实部为负）；要使矩阵的幂收敛于0，则特征值存在于单位圆内部（即模小于1），这是幂稳定区域。（上一讲的差分方程需要计算矩阵的幂。）

Section3.应用

我们注意Subsection1.基础知识介绍中提到的一阶微分方程系统的基本形式:

> ### Subsection1.基础知识介绍
>
> 我们这里主要介绍一阶微分方程系统，即一阶常系数微分方程组，考虑两个变量，他们之间相互耦合，导数之间也会有联系，即:
>
> $$\begin{cases}\frac{dx}{dt}=ax+by\\\frac{dy}{dt}=cx+dy\end{cases}$$
>
> 或者我们写为矩阵形式:

> $$\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$
>
> 除了$\begin{bmatrix}x\\y\end{bmatrix}$的写法，我们也可以写成$u(t)=\begin{bmatrix}u_1\\u_2\end{bmatrix}$的形式，二者是一个意思。

我们来看二阶情况如何计算，有$y''+by'+ky=0$。我们也模仿差分方程的情形，构造方程组$\begin{cases}y'' & =-by'-ky\\y' & =y'\end{cases}$，写成矩阵形式有$\begin{bmatrix}y''\\y'\end{bmatrix}=\begin{bmatrix}-b & -k\\1 & 0\end{bmatrix}\begin{bmatrix}y'\\y\end{bmatrix}$，令$u'=\begin{bmatrix}y''\\y'\end{bmatrix}, u=\begin{bmatrix}y'\\y\end{bmatrix}$。

那么我们同样的求特征值，特征向量就可以得到$u=\begin{bmatrix}y'\\y\end{bmatrix}$。
例如我们求解$y''+5y'+4y=0$，可以得到$\lambda_1=-1,\lambda_2=-4$，以及$x_1=\begin{bmatrix}-1\\1\end{bmatrix}x_2=\begin{bmatrix}-4\\1\end{bmatrix}$，那么根据通解的公式:$u=c_1\lambda_1x_1+c_2\lambda_2x_2=\begin{bmatrix}-c_1e^{-x}-4c_2e^{-4x}\\c_1e^{-x}+c_2e^{-4x}\end{bmatrix}=\begin{bmatrix}y'\\y\end{bmatrix}$，因此很容易得到解。

继续推广，对于5阶微分方程$y'''''+by''''+cy'''+dy''+ey'+f=0$，则可以写作$\begin{bmatrix}y'''''\\y''''\\y'''\\y''\\y'\end{bmatrix}=\begin{bmatrix}-b & -c & -d & -e & -f\\1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}y''''\\y'''\\y''\\y'\\y\end{bmatrix}$，这样我们就把一个五阶微分方程化为$5\times5$一阶方程组了，然后就是求特征值、特征向量了步骤了。