[微积分学] 无痛的 微分方程入门 I

引言

这个文章标题来自我看过的很多文章 *xxx Without the Agonizing Pain*，我本来很想硬套或者干脆直接『微分方程 没有剧烈的疼痛』之类的，但我克制住了（

基于往期博客『微积分基础入门 V2』的内容，本文我以随笔的形式整理了微分方程的基础概念、解法和简单应用，本文假设读者有往期微积分基础文章+线性代数 P1 的知识门槛。
复分析建议看，因为如果要继续学基本上逃不掉，但是不作为本文门槛。

> 感谢曾经在我自学微分学时提供过帮助的 田中飒太 先生 和 Mr. Jürgen，还有我亲爱的 W女士（七海）

内容范围

由于就算是基础概念内容也比较多，所以我决定微分方程会作为微积分博客的子系列分集写，本文作为 I 篇的涉及范围有：

- 微分方程的基础概念
- 常见的一阶常微分方程解法
- 可拓展到高阶的常系数线性微分方程解法
- 初步的齐次线性常微分方程组的认识
- 主要应用于计算机等领域的数值方法

其中部分我觉得比较需要解释的结论会包括推论和证明，或者对于繁琐于需要挨个列举的东西（例如单边拉普拉斯变换的性质）给出有完整/正规的列举的链接。如果你对某个我直接忽略掉证明的结论有疑惑可以以任何我提供的方式联系我。

基础概念

微分方程及其意义

微分方程就是包含未知函数及其导数的方程，它描述了一个函数及其导数之间的关系，我们解微分方程本质关注的是能够找到满足给定微分方程的函数以进行进一步分析。

我们用阶数来表示一个微分方程中最高阶导数的阶数。
微分方程分为常微分方程（Ordinary Differential Equations，ODE）和偏微分方程（Partial Differential Equations，PDE）。其中常微分方程是指微分方程的未知数是自变量只有一个的函

数，对于一个未知函数 y，自变量为 x，则一个 n 阶的常微分方程形式如下，为了避免歧义这里使用莱布尼兹记法。

$$f(x, \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}}, \frac{\mathrm{d}y}{\mathrm{d}x}, y) = 0$$

对应的，偏微分方程是指微分方程的未知数是有多个变量的函数，我们定义 $\mathbf{x} = [x_1, x_2, \cdots, x_m]$，该函数表示为 $u(\mathbf{x})$，则一个 n 阶的偏微分方程形式如下

$$f(\mathbf{x}, u, \cdots, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \cdots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \cdots, \frac{\partial^n u}{\partial x_1, \partial x_2, \cdots, x_n}) = 0$$

乍一看这坨东西真烦人，实际上后面那些就是表示方程中所有涉及到的函数及其偏导数的组合，例如 $\frac{\partial u^n}{\partial x_i}$ 表示的是 u 对 x_i 的 n 阶偏导数，$\frac{\partial u}{\partial x_i x_j}$ 是 u 对 x_i 和 x_j 的混合偏导数，综合一下，$\frac{\partial u^n}{\partial x_1, x_2, \cdots, x_m}$ 就是 u 对 m 个自变量的 n 阶混合偏导数。

好，我知道即便如此它还是很不美丽，但是先别着急，形式而已形式而已（

> 在我时隔两年三刷《黑客与画家》，我注意到了其中提到的用『美丽』这个词来形容数学相关的东西。

初值问题和边值问题

为了在微分方程中确定唯一的解，我们需要一个初始条件提供函数在某一特定点上的值，这个道理初中生都知道，就类似初中的给你特定的点然后求函数解析式。这种求解类型的问题叫做初值问题。
一个一阶常微分方程的初值问题一般形式可以表示为：

$$y'(x) = f(x, y(x)),\ y(x_0) = y_0$$

其中 $f(x, y(x))$ 是描述因变量 $y(x)$ 关于独立变量 x 的变化率的函数，也就是表示 $y(x)$ 在每个时刻 x 处的导数的函数，我们的目标就算找到一个合适的 $y(x)$ 使得方程两边相等。
不同于初值问题，边值问题的条件是自变量的定义域，一个一阶常微分方程的边值问题一般形式可以表示为：

$$y'(x) = f(x, y(x)), a \le x \le b, y(a) = \alpha, y(b) = \beta$$

微分方程的解

在微分方程中，我们会用通解和特解来讨论解的类别。通解是一个包含了微分方程所有解的解集，简单的说它是通用的，一般不包含任何特定的初始条件，而是表示微分方程所有可能解的形式，形如 $y(x) = \varphi(x, C)$，其中 C 是待定常数，通过对 C 的具体取值，我们可以得到特解。

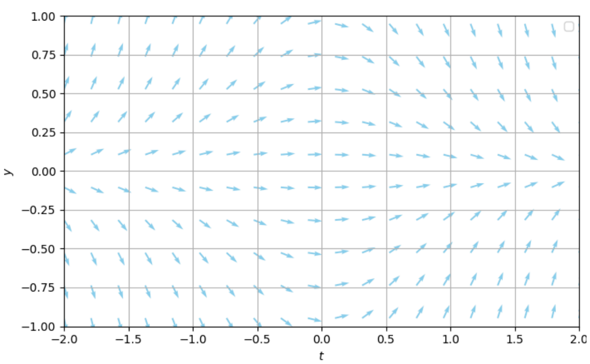
对应的，特解就是微分方程满足某个特定初值条件或边值条件的解，如果我们有一对特定的值 x_0 和 y_0，使得 $y(x_0)=y_0$，那么 $y(x)$ 就是微分方程在 x_0 处的特解。

如果 $y_1(x), y_2(x), \cdots, y_n(x)$ 是微分方程的解，且它们线性无关，则我们称其为线性无关解。我们可以通过 Wronskian 行列式来确定解组是否线性无关，定义为：

$$W(y_1, y_2, \ldots, y_k) = \begin{vmatrix} y_1 & y_2 & \cdots & y_k \\ y_1' & y_2' & \cdots & y_k' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(k-1)} & y_2^{(k-1)} & \cdots & y_k^{(k-1)} \end{vmatrix}$$

如果朗斯基行列式在某个 $x_0 \neq 0$，那么解在该点是线性无关的。如果对于所有的 x 都有 $W(y_1, y_2, \ldots, y_k) \neq 0$，那么这组解在整个定义域都是线性无关的。

方向场



在此我们需要引入一个概念叫方向场，一种将微分方程图形化表示的方法，它显示了在平面上每一点的切线方向，这些切线方向由微分方程确定。

图例是一个简单的一阶常微分方程 $\frac{\mathrm{d}y}{\mathrm{d}x} = -2ty$ 在平面上的解的整体行为，其中的箭头就是解在每个点上的变化方向，其长度表示解在该点上的变化速率。

通过绘制方向场，我们可以在整个平面上获得微分方程解的一些定性特征，这种图形化的方法为理解微分方程的行为提供了直观的手段。此处你应对此概念有基本的认识。

完全微分

根据之前所述，我们知道一个一阶微分方程的标准形式为（但是注意不是所有都行）：

$$y' = f(x, y)$$

如果存在一个函数 $\phi(x,y)$ 满足 $\frac{\partial \phi}{\partial x} = M(x,y), \frac{\partial \phi}{\partial y} = N(x,y)$ （M, N 在定义域内连续）使得

$$d\phi = M(x,y)\mathrm{d}x + N(x,y)\mathrm{d}x$$

则该微分形式是完全微分的。在这里，ϕ 被称为势函数，它的存在意味着该方程是完全微分方程，也

就是恰当方程，并且 ϕ 是微分形式的积分因子。

上式意味着其混合偏导数的次序可以交换。回顾微积分基础V2的补充节，施瓦茨定理，若一个函数 f 某一点的混合偏导数存在且连续，那么这些混合偏导数的次序可以交换，这就是Schwarz定理，例如一个二元函数 $f(x_1,x_2)$，则可以表示为：

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

当我们关注的是这个势函数 ϕ 是否存在时，我们就能反过来使用施瓦茨定理的条件来确保 ϕ 的存在性。

线性微分方程

显而易见地，线性微分方程是指一阶或高阶导数以及未知函数本身的线性组合等形式的微分方程。一个 n 阶的线性微分方程形如：

$$a_n(x)\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1}(x)\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a_0(x)y = f(x)$$

其中 $a_0(x), a_1(x), \cdots, a_n(x)$ 是系数函数，若系数函数都不依赖于 x，也就是每个系数函数都是常数，则这个方程是常系数线性微分方程，若至少有一个系数函数依赖于 x，则这是变系数线性微分方程。不同的系数函数导致不同的微分方程，因此通过调整这些系数，我们可以调整方程中各个导数阶次的权重，从而改变方程的性质。

在后面的章节我们将涉及一些数学技巧来求解这一类微分方程。

一阶常微分方程解法

在基础概念章节我们已经对微分方程有了一个基本的概念，在推广到更广泛的微分方程前我们不如先来点前菜，先来针对一阶ODE了解一些解法找找感觉。

变量分离方程

一类可以通过变量分离方法来求解的微分方程（注意这并不限于常微分方程和偏微分方程），我们称其为变量分离方程，一般形式为：

$$
\frac{\mathrm{d}y}{\mathrm{d}x} = g(x)h(y)
$$

这意味着它可以被携程两个变量的乘积的形式，也就是把上面那坨东西变成：

$$
\begin{aligned}
\frac{1}{h(y)} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = g(x) \\
\int \frac{1}{h(y)}\,\mathrm{d}y = \int g(x)\,\mathrm{d}x + C
\end{aligned}
$$

其中 C 是积分常数，想起之前我们提过的积分是对导数的逆运算，所以我们只需要这样使用积分反过来求解这个函数即可。

方法简单，我们还是来点例子，考虑一个简单的微分方程：

$$
\frac{\mathrm{d}y}{\mathrm{d}x} = x \cdot y^2
$$

分离并进行积分操作，再进行代数运算就可解出我们需要的 y

$$
\begin{aligned}
\int \frac{1}{y^2}\,\mathrm{d}y = \int x\,\mathrm{d}x + C \\
-\frac{1}{y} = \frac{x^2}{2} + C \\
y = -\frac{1}{\frac{x^2}{2} + C}
\end{aligned}
$$

伯努利方程

我们遇到的方程不一定要线性，但是这会使得它比较难处理，而之后我们章节的方法经常要求方程是线性的，所以我们先考虑如何将非线性的微分方程化为线性的，在一阶线性微分方程的标准形式 $\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)$ 的基础上上再加点什么东西，把它变成：

$$
\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)y^\alpha
$$

这个看起来不是很友好的 y^α 就是方程的非线性项，这个形式的方程叫伯努利方程。微软我们可以通过适当的变量替换使其变成线性的，我们先对方程进行一个简单的变形：

$$
\begin{aligned}
y^{-\alpha} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} + P(x)z = Q(x) \\
\frac{1}{1-\alpha} \cdot \frac{\mathrm{d}z}{\mathrm{d}x} + P(x)z = Q(x)
\end{aligned}
$$

然后引入一个新的变量 $z=y^{1-\alpha}$，则我们能由此推出：

$$\frac{\mathrm{d}z}{\mathrm{d}x}=(1-\alpha)y^{-\alpha}\frac{\mathrm{d}y}{\mathrm{d}x}$$

代入得到：

$$(1-\alpha)y^{-\alpha}\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)(1-\alpha)z=Q(x)(1-\alpha)$$

这样简单的操作，它就变成了一个一阶线性常微分方程，之后我们就可以使用标准的积分因子或其他方法来求解。

积分因子法

积分因子法也是我们常用于求解一阶常微分方程的方法，首先我们先考虑线性的方程（这种情况往往更简单），还是强调如下形式：

$$\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y=Q(x)$$

我们需要找到那个积分因子 $\mu(x)$ 将其转化为完全微分方程，以便我们更容易对其进行积分，也就是使得方程可以写成形式：

$$\mu(x)\left(\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y\right)=\mu(x)Q(x)$$

关于这个积分因子如何选择的问题，我们可以发现 μ 需要满足以下方程，所以我们直接求解得到：

$$\begin{aligned}\mu(x)P(x)&=\frac{\mathrm{d}\mu(x)}{\mathrm{d}x}\\ \mu(x)&=e^{\int P(x)\mathrm{d}x}\end{aligned}$$

这就是我们在积分因子法中经常选择 $\mu(x)=e^{\int P(x)\mathrm{d}x}$ 作为积分因子的原因，它能确保我们的方程在与积分因子相乘后可以写成一个完全微分形式。

所以我们只需要把这坨东西带进去，应用链式法则（第二行）再同时积分（第三行），再运用一些微分积分的基本性质直接操作（四到五行）得到最后的结果。

$$
\begin{aligned}
e^{\int P(x)\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y\right) &= e^{\int P(x)\mathrm{d}x}Q(x) \\
\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\int P(x)\mathrm{d}x}P(x)y\right) &= e^{\int P(x)\mathrm{d}x}Q(x) \\
\int \frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\int P(x)\mathrm{d}x}P(x)y\right) &= \int e^{\int P(x)\mathrm{d}x}Q(x) \\
e^{\int P(x)\mathrm{d}x}y &= \int e^{\int P(x)\mathrm{d}x}Q(x)\mathrm{d}x+C \\
y = e^{-\int P(x)\mathrm{d}x}\left(\int e^{\int P(x)\mathrm{d}x}Q(x)\mathrm{d}x+C\right)
\end{aligned}
$$

好，现在我们推出来是同样适用于非其次的一阶线性微分方程的解，但是如果说它是齐次的话，也就是 $Q(x)=0$ 的话，我们带进去就得到了只适用于齐次一阶线性微分方程的解。

$$
y = Ce^{-\int P(x)\mathrm{d}x}
$$

更一般地，对于一阶齐次非线性微分方程，我们也经常可以考虑使用积分因子法，根据上一章我们对线性微分方程的形式定义，我们可以将其改写为可能让大部分人更为熟悉的一阶微分方程形式：

$$
M(x,y)\mathrm{d}x+N(x,y)\mathrm{d}y=0
$$

我们首先需要检查它是否完全微分，如果不是，则我们需要找到那个积分因子 $\mu(x,y)$ 将其转化为完全微分方程，以便我们更容易对其进行积分，也就是使得方程可以写成形式：

$$
\mu M(x,y)\mathrm{d}x+\mu N(x,y)\mathrm{d}y=0
$$

对于一个不完全微分的一阶微分方程，我们可以通过这个公式找到使其完全微分的积分因子 μ：

$$
\mu = e^{\int \frac{M_y-N_x}{M}\mathrm{d}y}
$$

或者

$$
\mu = e^{\int \frac{N_y-M_x}{M}\mathrm{d}x}
$$

常数变易法

基于本节对非齐次和齐次一阶线性微分方程的解推导，我们可以再引入一个求解的方法，即在齐次解 $y=Ce^{-\int P(x)\mathrm{d}x}$ 的基础上引入一个待定函数 $u(x)$ 来替换掉 积分常数 C，再回代到一阶线性微分方程并解出 u：

$$
\begin{aligned}
u'e^{-\int P(x)\mathrm{d}x} - uP(x)e^{-\int P(x)\mathrm{d}x} + P(x)ue^{-\int P(x)\mathrm{d}x} &= Q(x) \\
u'e^{-\int P(x)\mathrm{d}x} &= Q(x) \\
ue^{-\int P(x)\mathrm{d}x} &= \int Q(x)\mathrm{d}x + C_1 \\
u &= e^{-\int P(x)\mathrm{d}x}\left(\int Q(x)\mathrm{d}x + C_1\right)
\end{aligned}
$$

然后我们把这个结果代回 $y = Ce^{-\int P(x)\mathrm{d}x}$ 就得到了一阶非齐次线性方程的通解：

$$
y = e^{-\int P(x)\mathrm{d}x}\left(\int Q(x)\mathrm{d}x + C_1\right)e^{-\int P(x)\mathrm{d}x}
$$

这种方法的优势在于，通过引入待定函数，我们可以更灵活地处理常数项（因为 C_1 既包含了原始齐次解中的积分常数 C，也包含了非齐次方程求解过程中引入的积分常数），从而更方便地求解非齐次线性微分方程。

例题

自己编了点例题，读者自行练习，学数学做点题还是很必要的。

答案不分中英版，发布后见文末链接，正好可以让某些懒惰的家伙不要第一眼就想着看答案（

(1) $\frac{\mathrm{d}y}{\mathrm{d}x} = 2x$

- Tip: 分离变量并积分两边以求解微分方程。

(2) $\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = e^{-x}$

- Tip: 将方程重写为标准形式并找到合适的积分因子，将左侧表示为微分，然后对方程两边进行积分。

(3) $\frac{\mathrm{d}y}{\mathrm{d}x} + (x - 2y) = 0$

- Tip: 检查它是否是一个完全微分方程，如果不是，找到积分因子 μ 并将其转换成完全微分形式。

常系数线性微分方程

常系数线性微分方程泛指方程中各项系数都为常数的方程，和上一章的定义范围不同。一个 n 阶的

常系数线性微分方程形式为：

$$a_n\frac{\mathrm{d}^n y}{dx^n}+a_{n-1}\frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}}+\cdots+a_1\frac{\mathrm{d}y}{\mathrm{d}x}+a_0y=f(x)$$

本章的内容包括但不限于高阶的常系数线性微分方程的解法，因为我意识到了在『一阶』的title下会使得我写的内容非常受限，而单纯写『高阶』又好像显得很狭隘。了解我的读者可能会知道我一般会很想直接上来就是更通用的、可以推广到高阶的方法，所以理论上我的正文现在才开始（被打）

特征方程法

首先我们要考虑这个很常见的解常系数线性齐次微分方程的特征方程法，显而易见就是将微分方程转化为一个被称为特征方程的代数方程，一个齐次的常系数线性微分方程形如：

$$a_n\frac{\mathrm{d}^n y}{dx^n}+a_{n-1}\frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}}+\cdots+a_1\frac{\mathrm{d}y}{\mathrm{d}x}+a_0y=0$$

对于 r，指数函数 e^{ex} 的导数都是它自身的常数倍，即 $\frac{\mathrm{d}}{\mathrm{d}x}e^{rx}=re^{rx}$ ，它具有很好的微分性质所以我们可以做一个合理的假设，假设我们关注的解 y 具有指数形式 $y(x)=e^{rx}$ ，代回原方程，我们称其为特征方程。

$$a_nr^ne^{rx}+a_{n-1}r^{n-1}e^{rx}+\cdots+a_1r^ne^{rx}+a_0e^{rx}=0$$

当我们找到了特征方程的解 $r_1,r_2,\cdots,r_n$ ，并代回去，微分方程的通解经构建后形如：

$$y=C_1e^{r_1x}+C_2e^{r_2x}+\cdots+C_ne^{r_nx}$$

二阶特征方程的判别式

讲点初中的内容（防止某些离开初中就忘记这些的大笨蛋不会用）

当我们为一个二阶常系数线性齐次微分方程套用特征方程法时，我们会这样代入：

$$ar^2+br+c=0$$

公式很明显是一个二次方程，咱那会（？）有一个二次方程根的判别式 $\Delta=b^2-4ac$ 来判断方程是否有实数根，据我所知初中的教材有讨论这种情况：

- 若 $\Delta>0$ 则方程有两个实数根 $r=\frac{-b\pm\sqrt{\Delta}}{2a}$
- 若 $\Delta=0$ 则方程有一个实数根 $r=\frac{-b}{2a}$
- 若 $\Delta<0$ 则方程无实数根

拓展一下初中课本忽略的部分，当 $\Delta < 0$ 时，特征方程的根是共轭复数对，这样的情况经常出现在振动系统中，比如弹簧振子、摆动、机械振动系统等，或者在量子力学中波函数的时间演化方程也可以由类似的微分方程描述，它用于描述粒子的波动特征。α, β 为实数， i 为虚数单位，则我们有共轭复数对

$$\begin{aligned} r_1 &= \alpha + i\beta \\ r_2 &= \alpha - i\beta \end{aligned}$$

根据欧拉公式给出的 $e^{ix} = \cos(x) + i\sin(x)$ ，我们可以简单直接地套进去得到：

$$e^{(\alpha+i\beta)x} = e^{\alpha x}(\cos(\beta x) \pm i\sin(\beta x))$$

这样我们就能简单地合并它的实部和虚部，代入简化得到通解公式：

$$\begin{aligned} y &= C_1 e^{r_1 x} + C_2 e^{r_2 x} \\ &= C_1 e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) + C_2 e^{\alpha x}(\cos(\beta x) - i\sin(\beta x)) \\ &= (C_1 + C_2)e^{\alpha x}(\cos(\beta x)) + i(C_1 - C_2)e^{\alpha x}(sin(\beta x)) \end{aligned}$$

微分算子法

对于高阶的常系数线性微分方程，微分算子法经常是一个很有效的方法，我们考虑这样形式的 n 阶常系数线性微分方程：

$$a_n \frac{\mathrm{d}^n y}{dx^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = f(x)$$

我们引入算子 $\mathcal{D}$ 如下：

$$\mathcal{D} = \frac{\mathrm{d}}{\mathrm{d}x}, \mathcal{D}^2 = \frac{\mathrm{d}^2}{\mathrm{d}x^2}, \cdots, \mathcal{D}^n = \frac{\mathrm{d}^{\mathrm{n}}}{\mathrm{d}x^n}$$

直接代回去就变成：

$$\begin{aligned} a_n\mathcal{D}^n y + a_{n-1}\mathcal{D}^{n-1} y + \cdots + a_1\mathcal{D}y + a_0 y &= f(x) \\ (a_n\mathcal{D}^n + a_{n-1}\mathcal{D}^{n-1} + \cdots + a_1\mathcal{D} + a_0)y &= f(x) \end{aligned}$$

我们定义 $\mathcal{D}^0 = I$ ，其中 I 为单位算子表示 $\mathcal{D}$ 没有进行导数操作，然后我们就能保持微分算子幂运算的一致性的前提下定义 $P(\mathcal{D}) = \sum_{i=0}^{n} a_i \mathcal{D}^i$ ，这时我们称 $P(\mathcal{D})$ 为 n 阶算子多项式，这种时候方程可以表示为：

$$P(\mathcal{D})y = f(x)$$

显而易见的我们就能得到其特解，其中 $\frac{1}{P(\mathcal{D})}$ 我们称其为 $P(\mathcal{D})$ 的逆算子。

$$y^* = \frac{1}{P(\mathcal{D})}f(x)$$

指数输入定理

就此，我们可以引入一个与微分算子法相关的（或者说延伸出来的）定理叫指数输入定理，专门用于解决输入为指数函数的方法，当方程是 $f(x) = e^{ax}$，即 $P(\mathcal{D})y = e^{ax}$ 时，其特解为：

$$y^* = \frac{e^{ax}}{P(a)}\ (P(a) \neq 0)$$

在实数域的微分方程如何证明显而易见，就不多说了。啊对，我们还有复数域上的微分方程要面对，但是这很简单，我们考虑一个有一个复变函数 $f(z) = e^{az}$ ，我们就能先表示出它的实部和虚部：

$$f(z) = P(\mathcal{D})y = \Re(f(z)) + i\cdot\Im(f(z))$$

由于微分算子对于加法和数乘操作保持线性性质，所以我们可以将实部和虚部分别独立地处理：

$$f(z) = P(\mathcal{D})y = \Re(f(z))$$

$$f(z) = P(\mathcal{D})y = \Im(f(z))$$

所以指数输入定理同样适用于复数域的微分方程，即可以用于解三角函数。

拉普拉斯变换

> 其实这里自变量我还是比较想用 x（因为我一直很注重统一），但是拉普拉斯变换涉及的更多是时间相关的问题以至于我看过的大部分资料都是用 t表示，这里为了让读者看其他资料时能保持亲切感故使用 t

拉普拉斯变化是一种积分变换，属于线性变换，将一个在 $[0, \infty)$ 上局部可积的，有实数变量的函数 $f(t)$ 转换为一个变量为复数 s 的函数，用符号 $\mathcal{L}\{f(t)\}$ 表示，另外，在不混淆的情况下我们可以有 $F(s) = \mathcal{L}\{f(t)\}$ 来表示拉普拉斯变换，所以：

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty f(t)e^{-st}\mathrm{d}t$$

要注意的是，只有 $\int_0^\infty f(t)e^{-st}\mathrm{d}t$ 收敛时，拉普拉斯变换才存在。
我们有上述等式，当我们对 $f'(t)$ 应用拉普拉斯变换：

$$\mathcal{L}\{f'(t)\} = \int_0^\infty e^{-st} f'(t)\,\mathrm{d}t$$

通过套用分部积分法，我们得到：

$$\mathcal{L}\{f'(t)\} = e^{-st} f(t)\big|_0^\infty - \int_0^\infty f(t)(-se^{-st})\mathrm{d}t$$

对于指数函数 e^{-st} 当 t 趋于正无穷大时，指数的幂次部分 $-st$ 中的 t 会使指数函数的值趋近于零，即 $\lim_{a\to\infty} e^{-st} f(a) = 0$，所以边界项 $e^{-st} f(t)|_0^\infty = -f(0)$。另外一个项 $-\int_0^\infty f(t)(-se^{-st})\,\mathrm{d}t$ 我们可以简化，再反过来把那坨 $\int_0^\infty e^{-st} f(t)\,\mathrm{d}t$ 变换回去 $F(t)$，得到：

$$\begin{aligned}
\mathcal{L}\{f'(t)\} &= e^{-st} f(t)\big|_0^\infty - \int_0^\infty f(t)(-se^{-st})\mathrm{d}t \\
&= -f(0) + s\int_0^\infty e^{-st} f(t)\,\mathrm{d}t \\
&= sF(t) - f(0)
\end{aligned}$$

于是我们得到了拉普拉斯变化一阶导数定理。

> 这好像和文章主题跑远了，我想show一下证明只是单纯因为这是我以前自己瞎证的
>
> 为了读者更好运用，请自行前往Wiki查看常见的单边拉普拉斯变换性质：
> https://en.wikipedia.org/wiki/Laplace_transform
> 后面的内容和习题如果会用到相关性质便不再特别强调

好，反正无论怎样现在你也知道这是什么了，所以我们讨论一下怎么把拉普拉斯变换应用在常系数线性微分方程，还是一样考虑如下形式：

$$a_n \frac{\mathrm{d}^n y}{dt^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y = f(t)$$

我们可以对方程的两边同时进行拉普拉斯变换，即将上式中的各项转换为复频域上的代数表达式：

$$a_n\mathcal{L}\{\frac{\mathrm{d}^n y}{dt^n}\} + a_{n-1}\mathcal{L}\{\frac{\mathrm{d}^{n-1} y}{dt^{n-1}}\} + \cdots + a_1\mathcal{L}\{\frac{\mathrm{d}y}{\mathrm{d}t}\} + a_0\mathcal{L}\{y\} = \mathcal{L}\{f(t)\}$$

这样说太复杂又太抽象，所以我们以一个一阶常微分方程 $\frac{\mathrm{d}y}{\mathrm{d}t} + ay = 0$ 为例，$Y(s) = \mathcal{L}\{y(t)\}$。我们对其进行拉普拉斯变换得到：

$$\begin{aligned}
\int_0^\infty f(t)e^{-st}\mathrm{d}t &= ye^{-st}|_0^\infty - \int_0^\infty y(t)(-se^{-st}\mathrm{d}t) \\
&= sY(s) - y(0)
\end{aligned}$$

代回去得到：

$$sY(s) - y(0) - aY(s) = 0$$
$$Y(s) = \frac{y(0)}{s-a}$$

然后现在我们要引入拉普拉斯逆变换的概念，显而易见就是拉普拉斯变换的逆运算，它把复数域上的函数转换回到时域，形式如下，其中 γ 是一个包含 $F(s)$ 所有极点的实数，i 是虚数单位。

$$f(t) = \mathcal{L}^{-1}\{F\}(t) = \frac{1}{2\pi i}\lim_{T\to\infty}\int_{\gamma-iT}^{\gamma+iT} e^{st}F(s)\mathrm{d}s$$

其中，$\mathcal{L}^{-1}\frac{1}{s-a} = e^{at}u(t)$ 是一个关于拉普拉斯变换逆定理很常见的公式（因为网上推导比较多我就不写了，主要是今天比较忙x），所以我们应用就得到：

$$y(t) = y(0)e^{at}$$

例题

(1) 判断 $y'' + 2y' + 2y = 0$ 特征方程的根并写出其解
(2) 使用指数输入定理求解 $\frac{\mathrm{d}y}{\mathrm{d}x} + y = e^x$
(3) 对函数 $f(x) = e^{2x}$ 在区间 $[0, \infty]$ 上展开拉普拉斯变换
(4) 对拉普拉斯变换的函数 $F(s) = \frac{1}{s^2+4s+5}$ 求解它的原函数

齐次线性常微分方程组

定义

一个 n 阶的标准的线性常微分方程组可以用矩阵和向量紧凑地表示为：

$$\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x} = \mathbf{A}\mathbf{y}(x) + \mathbf{f}$$

其中 $\mathbf{y}(x)$ 就是我们关注的解，是未知函数的向量。$\mathbf{A}$ 是系数矩阵，显然当 $\mathbf{f} = 0$ 时，这是一个齐次方程组，一个齐次的常系数线性常微分方程组的未知函数的向量 $\mathbf{y}(x)$ 和系数矩阵 $\mathbf{A}$ 形式为：

$$
\mathbf{A}=\begin{bmatrix}a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{bmatrix},\mathbf{y}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}
$$

当常系数线性微分方程组是齐次的时候，我们可以简单地使用矩阵指数来表示这样的方程组的通解，引入定义，对于 n 阶的实矩阵 $\mathbf{A}$ ，其矩阵指数为：

$$
e^{\mathbf{A}}=\mathbf{E}+\mathbf{A}+\frac{\mathbf{A}^2}{2!}+\cdots=\sum_{n=0}^{+\infty}\frac{\mathbf{A}^n}{n!}
$$

我们定义一个单位矩阵 $\mathbf{E}$，然后我们可以定义这个矩阵指数函数为：

$$
e^{\mathbf{A}x}=\mathbf{E}+\mathbf{A}x+\frac{(\mathbf{A}x)^2}{2!}+\cdots=\sum_{n=0}^{+\infty}\frac{(\mathbf{A}x)^n}{n!}
$$

然后我们把 $e^{\mathbf{A}x}$ 代入 $\frac{\mathrm{d}\mathbf{y}(x)}{dx}$

$$
\begin{aligned}
\frac{d\mathbf{y}(x)}{dx}&=\frac{\mathbf{d}}{\mathbf{d}x}\left(\mathbf{E}+\mathbf{A}x+\frac{\mathbf{A}^2}{2!}+\cdots+\frac{\mathbf{A}^i}{i!}+\cdots\right)\\
&=\mathbf{A}+\mathbf{A}^2x+\frac{\mathbf{A}^3x^2}{2!}+\cdots+\frac{\mathbf{A}^ix^{i-1}}{(i-1)!}+\cdots\\
&=\mathbf{A}\left(\mathbf{E}+\mathbf{A}x+\frac{(\mathbf{A}x)^2}{2!}+\cdots+\frac{(\mathbf{A}x)^{i-1}}{(i-1)!}\right)\\
&=\mathbf{A}e^{\mathbf{A}x}\\
&=\mathbf{A}\mathbf{y}(x)
\end{aligned}
$$

类似地，$e^{\mathbf{A}x}$ 是一个常系数线性微分方程组的矩阵形式的通解，也就是所谓基解矩阵，我们之后可以将其写作

$$
\exp\mathbf{A}x=\sum_{n=0}^{+\infty}\frac{(\mathbf{A}x)^n}{n!}
$$

但是我们还是应该明确地下定义，我们在第一章便已经了解了用 Wronskian行列式 来判断解的线性无关性的概念，也就是当 Wronskian行列式 不为零时，即 $W(y_1,y_2,\ldots,y_k)(x)\neq 0$ ，其中 $\forall x\in[a,b]$ 时，这组函数就是线性无关的。而对于微分方程组，当这个解组 $\mathbf{y}$ 它们之间线性无关时，我们称其构成的解矩阵 $\mathbf{Y}$ 为基解矩阵，而如果存在某点 $x_0\in[a,b]$ 使得 $\mathbf{Y}(x_0)=\mathbf{E}$，则这个 $\mathbf{Y}$ 是一个标准基解矩阵。

对于一个基解矩阵 $\mathbf{Y}$，初值条件由 $\mathbf{y}(x_0)=y_0$ 给出，则满足初值条件的解为：

$$\mathbf{y}(x)=\mathbf{Y}(x)\mathbf{Y}^{-1}(x_0)y_0$$

然后我们还需要重新考虑 $\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x}=\mathbf{A}\mathbf{y}(x)$ 的复值解，我们将解 $\mathbf{y}$ 分为实部和虚部：

$$\mathbf{y}(x)=\Re(\mathbf{y}(x))+i\cdot\Im(\mathbf{y}(x))$$

再把这坨让我眼睛快酸死的Latex公式带回方程：

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(\Re(\mathbf{y}(x))+i\cdot\Im(\mathbf{y}(x))\right)=\mathbf{A}\left(\Re(\mathbf{y}(x))+i\cdot\Im(\mathbf{y}(x))\right)\\ \frac{\mathrm{d}}{\mathrm{d}x}\Re(\mathbf{y}(x))+i\frac{\mathrm{d}}{\mathrm{d}x}\cdot\Im(\mathbf{y}(x)=\mathbf{A}\Re(\mathbf{y}(x))+i\mathbf{A}\Im(\mathbf{y}(x)\end{aligned}$$

因为复数方程中的实部和虚部之间是相互独立的，所以我们直接分别对实部和虚部进行比较即可：

$$\frac{\mathrm{d}\Re(\mathbf{y}(x))}{\mathrm{d}x}=\mathbf{A}\Re(\mathbf{y}(x))$$

$$\frac{\mathrm{d}\Im(\mathbf{y}(x))}{\mathrm{d}x}=\mathbf{A}\Im(\mathbf{y}(x))$$

因此，这对复值矩阵函数也适用。

特征向量和特征分解

但是显然地，矩阵指数函数 $e^{\mathbf{A}x}$ 是无穷级数，搞得很乱阿，因此我们需要用特征向量和特征分解来更直观地解决问题，再不行就解决提出问题的那个人吧（？）尽管这东西我在线性代数已经写过了，但这里还是再适用地提一遍。

我们考虑矩阵 $\mathbf{A}$ 的特征值问题（前提是 $\mathbf{A}$ 可对角化，否则下面的解就不能被定义为基解矩阵，但是我们也有处理方法），有特征向量矩阵 $\mathbf{V}=[\mathbf{v_1},\mathbf{v_2},\cdots,\mathbf{v_n}]$ 和与之对应的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，因为我们对 $\frac{\mathrm{d}}{\mathrm{d}x}e^{\lambda_i x}\mathbf{v}_i$ 应用链式法则可以得到：

$$\frac{\mathrm{d}}{\mathrm{d}x}e^{\lambda_i x}\mathbf{v}_i=\lambda_i e^{\lambda_i x}\mathbf{v}_i=\mathbf{A}e^{\lambda_i x}\mathbf{v}_i$$

所以，显而易见地，我们可以定义基解矩阵为：

$$\mathbf{Y}=[e^{\lambda_1 t}\mathbf{v}_1,e^{\lambda_2 t}\mathbf{v}_2,\cdots,e^{\lambda_n t}\mathbf{v}_n]$$

而在线性代数I中，我们知道了我们可以求解特征方程 $|\mathbf{A}-\lambda I|=0$ 来得到特征值，因此在这种情况下我们直接解出来套回去就行。

Jordan标准型

对于比较叛逆无法对角化的矩阵 $\mathbf{A}$ 我们需要用到Jordan标准型来操作，其中 $\mathbf{P}$ 是由特征向量组成的矩阵，这样的矩阵 $\mathbf{A}$ 可以写成：

$$\mathbf{A}=\mathbf{PJP}^{-1}$$

其中 $\mathbf{J}$ 是Jordan矩阵 $\mathbf{J}=\operatorname{diag}(\mathbf{J}_1,\mathbf{J}_2,\cdots,\mathbf{J}_n)$，令 $\mathbf{J}_i$ 表示对 $\mathbf{J}$ 主对角线的分块，例如用 $\mathbf{J}_1,\mathbf{J}_2$ 对 $\mathbf{J}$ 分块：

$$\mathbf{J}=\begin{bmatrix}\mathbf{J}_1 & \\ & \mathbf{J}_2\end{bmatrix}$$

每个Jorden块的形式为：

$$\mathbf{J}_i=\begin{bmatrix}\lambda_i & 1 & \cdots & 0\\ 0 & \lambda_i & \ddots & 0\\ \vdots & \vdots & \ddots & 1\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}$$

我们可以分来计算每个分块 $\mathbf{J}^{\mathbf{k}}$ 和分块的矩阵指数 $e^{\mathbf{J}x}$：

$$\mathbf{J}^{\mathbf{k}}=\begin{bmatrix}\mathbf{J^k}_1 & \\ & \mathbf{J^k}_2\end{bmatrix},\ e^{\mathbf{J}x}=\begin{bmatrix}e_1^{\mathbf{J}x} & \\ & e_2^{\mathbf{J}x}\end{bmatrix}$$

混一下，得到 $\mathbf{J^k}_i$：

$$\mathbf{J}_i^k=\begin{bmatrix}\lambda_i^k & k\lambda_i^{k-1} & \frac{k(k-1)}{2}\lambda_i^{k-2} & \dots & \frac{k!}{(k-m)!}\lambda_i^{k-m}\\ 0 & \lambda_i^k & k\lambda_i^{k-1} & \dots & \frac{(k-1)!}{(k-m-1)!}\lambda_i^{k-m-1}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & \lambda_i^k\end{bmatrix}$$

我们可以对指数函数 $e^{\mathbf{J}x}$ 进行展开：

$$e^{\mathbf{J}x}=\mathbf{E}+\mathbf{J}x+\frac{(\mathbf{J}x)^2}{2!}+\cdots=\sum_{m=0}^{+\infty}\frac{(\mathbf{A}x)^m}{m!}$$

$$e^{\mathbf{J}_ix}=\begin{bmatrix}e^{\lambda_ix} & xe^{\lambda_ix} & \frac{x^2}{2!}e^{\lambda_ix} & \dots & \frac{x^{m-1}}{(m-1)!}e^{\lambda_ix}\\ 0 & e^{\lambda_ix} & xe^{\lambda_ix} & \dots & \frac{x^{m-2}}{(m-2)!}e^{\lambda_ix}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & e^{\lambda_ix}\end{bmatrix}$$

这样，我们可以计算每个Jordan块的指数形式，进而得到整个矩阵。

常微分方程的数值方法

差分

差分是一种很基础的数值计算方法，当然除了求解微分方程，还能用于类似数值逼近、图像处理等问题，重点就是通过有限的插值来逼近导数或者积分，将一坨连续的问题给离散化。

其中，前向差分是通过当前点和下一个点的信息来逼近，在等距的节点 $x_k = x_0 + kh,\ (k = 0, 1, \cdots, n)$ 中的差分是函数值的增量 $f(x)_{n+1}$ 与当前函数值的差，我们记为：

$$\Delta f(x)_n = f(x)_{n+1} - f(x)_n$$

我们可以直接推广到高阶，$\Delta^k f(x)$ 为 $f(x)$ 的 k 阶差分，差分的差分是更高一阶的差分，即一个 k 阶的前向差分公式为 $\Delta^k f(x)_n = \Delta^{k-1}(f(x)_{n+1} - f(x)_n)$，于是我们有：

$$\begin{aligned}\Delta^k f(x) &= \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} f(x+i) \\ &= \sum_{i=0}^{k} \binom{k}{i} (-1)^{i} f(x+k-i)\end{aligned}$$

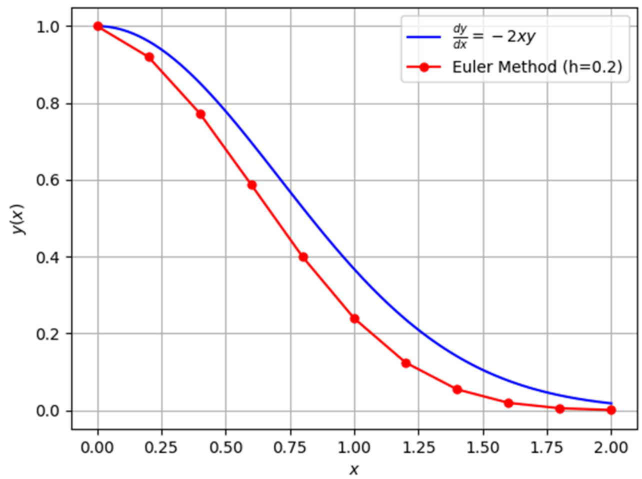
类似地，后向差分无非就是定义差分为当前函数值与前一函数值的差，即 $\Delta y_n = y_n - y_{n-1}$，一个 k 阶的后向差分公式为 $\Delta^k f(x)_n = \Delta^{k-1}(f(x)_n - f(x)_{n+1})$，注意，上述的公式同样适用于后向差分，因为我们包含了 -1 的幂，而决定了前向差分还是后向差分。

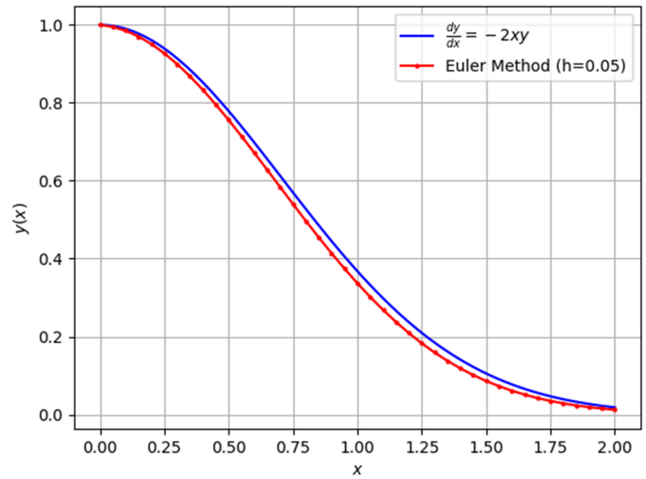
欧拉法

欧拉法是我们经常用于近似解常微分方程的初值问题的一种数值方法，基于线性近似，可以说是最简单的一种数值方法了。我们通过在当前点处切线的斜率来估计下一个点的函数值，也就是假设在一个步长内，解的变化可以用当前时刻的导数乘以时间步长来近似。为了比较紧凑地表示我们定义 $f(x) = y$ 若估计的步长为 h，则迭代公式为：

$$y_{n+1} = y_n + h \cdot f(x_n, y_n)$$

下图是选择两个不同的步长的欧拉法对微分方程 $\frac{\mathrm{d}y}{\mathrm{d}x} = -2xy$ 的近似，很显然地，步长越小，迭代的次数越多，近似程度会越高。





注意，欧拉法是很基础的一阶数值方法，其误差通常随着步长的增加而增加。因此，对于需要高精度数值解的问题，欧拉法可能不够准确，也不适用于变化很剧烈的曲线，但是在简单的初值问题经常是一个很好的选择。

上述形式的叫做『显式欧拉法』，对应地就有隐式欧拉法，形如：

$$y_{n+1} = y_n + h \cdot f(x_{n+1}, y_{n+1})$$

我们要做那之前先求解这个方程，例如使用牛顿迭代法将其转换为一个根查找问题，将这个方程变形并定义一个符号来表示就变成了：

$$G(y_{n+1}) = y_{n+1} - y_n - h \cdot f(x_{n+1}, y_{n+1}) = 0$$

牛顿迭代法的公式形如：

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

我们就能把这坨东西带进去，比较紧凑地表示为：

$$y_{n+1}^{k+1} = y_n - \frac{F(y_{n+1}^k)}{F'(y_{n+1}^k)}$$

类似地进行迭代运算即可。

Runge-Kutta方法

Runge-Kutta方法（音译好像是龙格-库塔，更难打了我还是就这样吧（） 也是一种数值求解常微分方程的数值方法。有点类似的就是不断调整斜率来估计解，并通过适当的权重对它们进行加权平均，所以可以说相比欧拉法它可能会更精确。

如果说我们在欧拉法的 $y_{n+1} = y_n + h \cdot f(x_n, y_n)$ 基础上做一些更改，变得使用两个斜率来估计

下一段的解，我们再次考虑步长 h，然后有如下形式的常微分方程和初值问题：

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y), y(x_0),\ y_0 = y(x_0)$$

类似地，我们可以分别计算我们两个目标斜率：

$$k_1 = h \cdot f(x_n, y_n)$$
$$k_2 = h \cdot f(x_n + h, y_n + k_1)$$

然后我们就能使用平均值加权来得到估计的解：

$$y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2)$$

这种取了两个斜率均值的 Runge-Kutta 我们叫二阶Runge-Kutta，简称RK2，而在实际应用中RK4会更为常用，取四个斜率的均值使得中点的斜率有更大的权重，这是同理的，它们都被称为『显式Runge-Kutta』，取决于阶数，一个 s 阶的Runge-Kutta（抱歉我习惯的k和n都被占掉了讨厌），其中 c_i 是权重系数，表达式为：

$$\begin{aligned}
k_1 &= hf(t_n, y_n) \\
k_2 &= hf(t_n + c_2 h, y_n + a_{21} k_1) \\
k_3 &= hf(t_n + c_3 h, y_n + a_{31} k_1 + a_{32} k_2) \\
&\ \vdots \\
k_s &= hf(t_n + c_s h, y_n + \sum_{i=1}^{s-1} a_{si} k_i)
\end{aligned}$$

定义权重系数 b_i ，所以我们有：

$$y_{n+1} = y_n + \sum_{i=1}^{s} b_i k_i$$

注意不要混淆，c_i 是对于局部调整的权重系数，而 b_i 是根据阶数做调整的权重系数，参考以下表格：

$$\begin{array}{c|ccccc}
0 & & & & \\
c_2 & a_{21} & & & \\
c_3 & a_{31} & a_{32} & & \\
\vdots & \vdots & \vdots & \ddots & \\
c_s & a_{s1} & a_{s2} & \dots & a_{s,s-1} \\
\hline
 & b_1 & b_2 & \dots & b_s
\end{array}$$

例题

> 我自己写答案会使用Haskell，读者自行选择即可。

(1) 编写一个函数，接受一个列表作为输入，返回该列表中每个元素与其后一个元素的差分值

(2) 编写一个函数，接收一个表示区间的二元组，求解常微分方程 $\frac{\mathrm{d}y}{\mathrm{d}x}=x-y$，步长 $h=0.1$。

(3) 使用四阶Runge-Kutta方法求解常微分方程 $\frac{\mathrm{d}y}{\mathrm{d}x}=x+y$，步长 $h=0.1$。

参考资料

Wiki-微积分学栏目（及子栏目）：https://en.wikipedia.org/wiki/Calculus

Differential Equations: Techniques, Theory, and Applications by Barbara D. MacCluer, Paul S. Bourdon, and Thomas L. Kriete.

Differential Equations(MIT open courses)

本文会含有我自编的例题，在博客发布后会以一篇独立的篇章发布包含解题思路的solutions，这取决于我什么时候写完。对于涉及编程的例题我会以Haskell标准库的实现为标准答案。

MAKIROR gzanan@gmail.com 23:24 06/12/2023