计算方法 Computing Method

姚遥

南京大学智能科学与技术学院

课程信息

- **课程名称**：计算方法（Computing Method）
- **课程网站**：https://garnet-larch-cd5.notion.site/cm2025
- **教室地点**：南雍楼 西311
- **授课教师**：姚遥
  - **办公室**：南雍楼 东509
  - **Office hour**：Thursday 1:00 pm - 2:00 pm
  - **教师主页**：https://yoyo000.github.io/

教师简介

Work Experience

- **Nanjing University**, Associate Professor *2023.06 - now*
- **Apple**, Senior Researcher *2020 - 2023*
- **Altizure**, Founding Member *2015 - 2020*

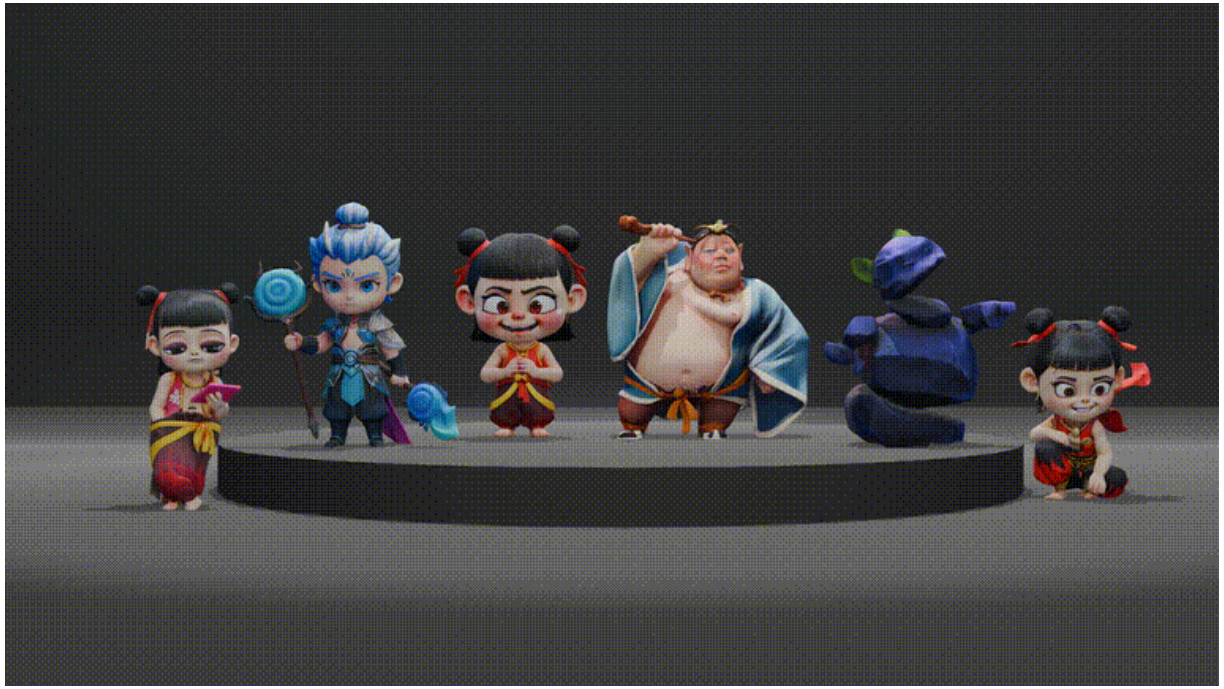
Education

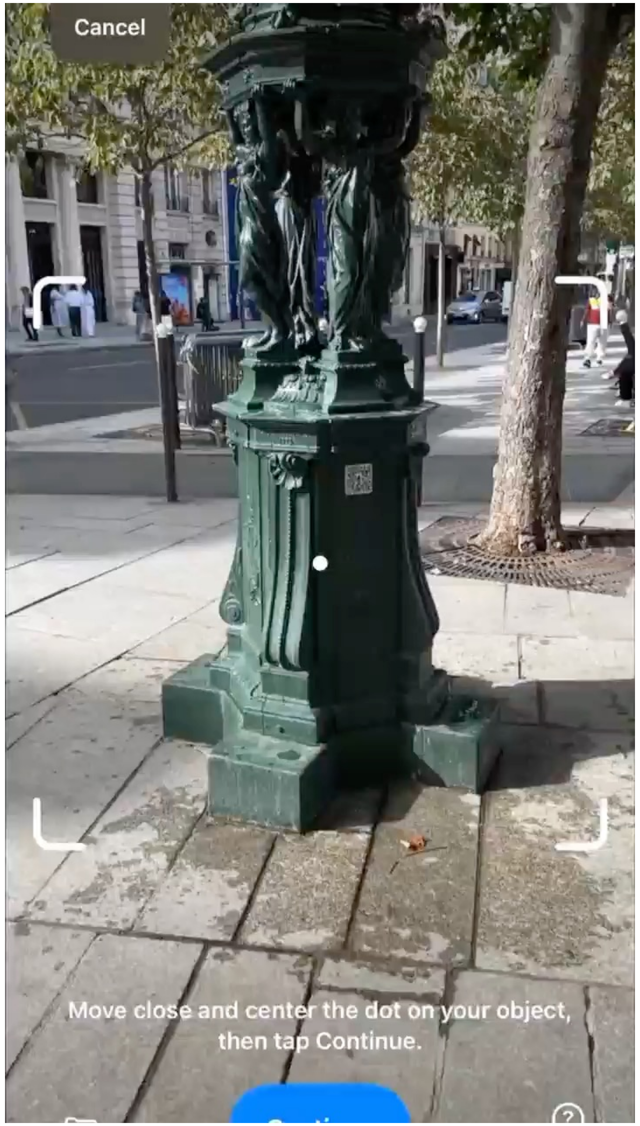
- **HKUST**, CSE, PhD *2015 - 2020*
- **Nanjing University**, ESE, BSc *2011 - 2015*

教师简介

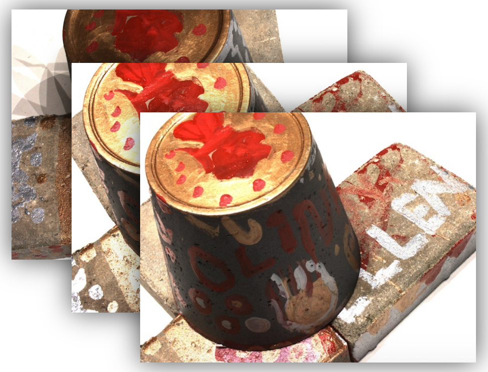
Research Interests: 3D Computer Vision

- 3D Reconstruction
- 3D Generation





三维重建与渲染



Input Images

Reconstruction
(Vision)

Geometry: Point cloud / Mesh / SDF …

Appearance: Texture map / BRDF / …

Lighting: Environment map / NeILF / …

Rendering
(Graphics)

Rendered Images

可微分渲染：大规模数值优化问题

Bridging vision & graphics!

Geometry: Point cloud / Mesh / SDF …

Input Images

Reconstruction
(Vision)

Appearance: Texture map / BRDF / …

Rendering
(Graphics)

Rendered Images

Lighting: Environment map / NeILF / …

先行课程

- **微积分**（Calculus）
- **线性代数**（Linear Algebra）
- **C++** (for implementing numerical algorithms)
- **Python** (for using numerical algorithms in real apps)

课程名称

中文名称：

- 计算方法 / 数值分析 / 数值计算 / …

英文名称：

- Computing Method
- Numerical Analysis
- Numerical Computation/Computing

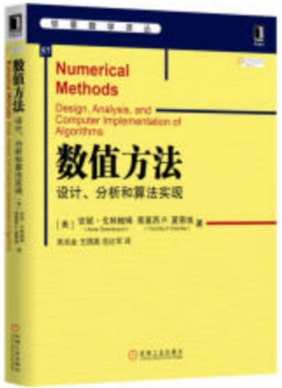
区分：

- 解析法 Analytical methods
- 数值法 Numerical methods

参考教材及课程

参考教材：

- 《数值分析》第5版，李庆扬等著，华中科技大学出版社，2018
- 《数值分析》原书第2版，裴玉茹等译，机械工业出版社，2014
  - ➢ 《Numerical Analysis》Second Edition，Timothy Sauer，2011



参考课程

- NJU IST 2024 Spring《计算方法》，黄卫华教授
- NJU AI 2023 Spring《计算方法》，张利军教授
  - ➢ https://www.lamda.nju.edu.cn/zhuangzh/cm2023.html

特别感谢**黄卫华**老师与**张利军**老师对本课程建设的支持！

助教简介

- **助教**:
  - 王嘉豪（ jiah.wang@smail.nju.edu.cn ）
  - 袁萌启（ mqyuan@smail.nju.edu.cn ）
  - 杨忆康（ yik.yang@smail.nju.edu.cn ）

Jiahao Wang (class 2024)

Yikang Yang (class 2024)

Mengqi Yuan (class 2024)

cm2025
群号：768238555

课程 QQ 群

课程总览

L1. 第 1 章：绪论
L2. 第 2 章：插值法（1）
L3. 第 2 章：插值法（2）
L4. 第 3 章：函数逼近与计算（1）
L5. 第 3 章：函数逼近与计算（2）
L6. 第 4 章：数值积分与数值微分（2）
L7. 第 4 章：数值积分与数值微分（2）

L8. 第 6 章：方程求根（1）
L9. 第 6 章：方程求根（2）
L10. 第 7 章：解线性方程组的直接方法（1）
L11. 第 7 章：解线性方程组的直接方法（2）
L12. 第 8 章：解线性方程组的迭代方法
L13. 第 9 章：矩阵的特征值与特征向量计算（1）
L14. 第 9 章：矩阵的特征值与特征向量计算（2）

课程考核

暂定，可能动态调整：

- **平时作业 16%**（共8次）
- **编程作业 24%**（共2次，10% + 14%）
- **期末考试 60%**（半开卷，10% + 50%）
- **出勤考核**（随机抽查，每次缺勤 -1%）

第 1 章 绪论

姚遥

yaoyao@nju.edu.cn

本章目录

- 研究对象与特点
- 误差来源与分析
- 误差的基本概念
- 误差分析的方法与原则

绪论 | 研究对象与特点

数值分析：

- 研究各种数学问题求解的数值计算方法
  - $\sqrt{\ }$ ∇ $\int$ $\mathcal{F}$ min max …
- 研究适合于计算机使用的数值计算方法
  - & | + - * /

实际问题 → 数学模型 → 数值计算方法 → 程序设计 → 编程实现

绪论 | 研究对象与特点

研究用计算机解决数学问题的数值方法及其理论

- 研究内容：函数的数值逼近、数值微分与数值积分、非线性方程数值解、数值线性代数、微分方程数值解等；
- 是数学的分支但不止于研究数学理论本身，理论与计算紧密结合，着重研究数学问题的数值方法及其理论

也称计算方法

- 有自身理论体系，内容丰富，研究方法深刻
- 既有纯数学的**高度抽象性与严密科学性**，又有**应用的广泛性**与实际试验的**高度技术性**
- 与计算机应用密切结合、实用性很强的数学课程

绪论 | 研究特点

- **面向计算机**：根据计算机特点提供实际可行的有效算法，只能包括加、减、乘、除运算和逻辑运算；
- **有可靠的理论分析**：能任意逼近并达到精度要求，近似算法要保证收敛性和数值稳定性，有误差分析；
- **有好的计算复杂性**：时间复杂性好是指节省计算时间，空间复杂性好是指节省存储空间；
- **有数值实验**：需通过数值实验证明算法有效性。

绪论 | 举个例子

求平方根倒数

- 图形引擎雷山之锤3 Quake-III Arena:
- Evil floating point bit level hacking?
- 4x faster than (float)(1.0/sqrt(x)) !

References:
https://zh.wikipedia.org/zh-hans/平方根倒数速算法
https://www.gongdear.com/articles/2018/06/13/1528869084699.html

```
float Q_rsqrt( float number ) {
    long i;
    float x2, y;
    const float threehalfs = 1.5F;
  x2 = number * 0.5F;
    y   = number;
    i   = * ( long * ) &y;
  // evil floating point bit level hacking
    i   = 0x5f3759df - ( i >> 1 ); // what the fuck?
    y   = * ( float * ) &i;
    y   = y * ( threehalfs - ( x2 * y * y ) );
// 1st iteration
    // y   = y * ( threehalfs - ( x2 * y * y ) );
 // 2nd iteration, this can be removed
  #ifndef Q3_VM
    #ifdef __linux__
      assert( !isnan(y) );
 // bk010122 - FPE?
    #endif
 #endif
return y; }
```

绪论丨国内发展

- **1955**：周恩来领导10年科技规划，提出发展几个新技术，包括计算技术（计算机，程序设计，计算数学），半导体技术，自动化技术。
- **1956**：成立计算技术研究所筹备处，主任华罗庚，手下有两个组（计算机与计算数学）。在其领导下于数学所成立计算方法讨论班。
- **1958**：计算所成立，带半军事性质，并且与苏联合作在中科院计算研究所造出104机。 北大、吉大、复旦、南大相继成立计算数学专业。

绪论丨国内发展

- **1958**：“计算数学 ” 专业
- **1984**：“计算数学及其应用软件”专业
- **1998**：“信息与计算科学”专业

绪论 | 国内发展

冯康（1920.09.09 - 1993.08.17）：

- 1944年毕业于国立中央大学
- 有限元方法
- 辛几何算法

https://baike.baidu.com/item/%E5%86%AF%E5%BA%B7/47515?fr=aladdin

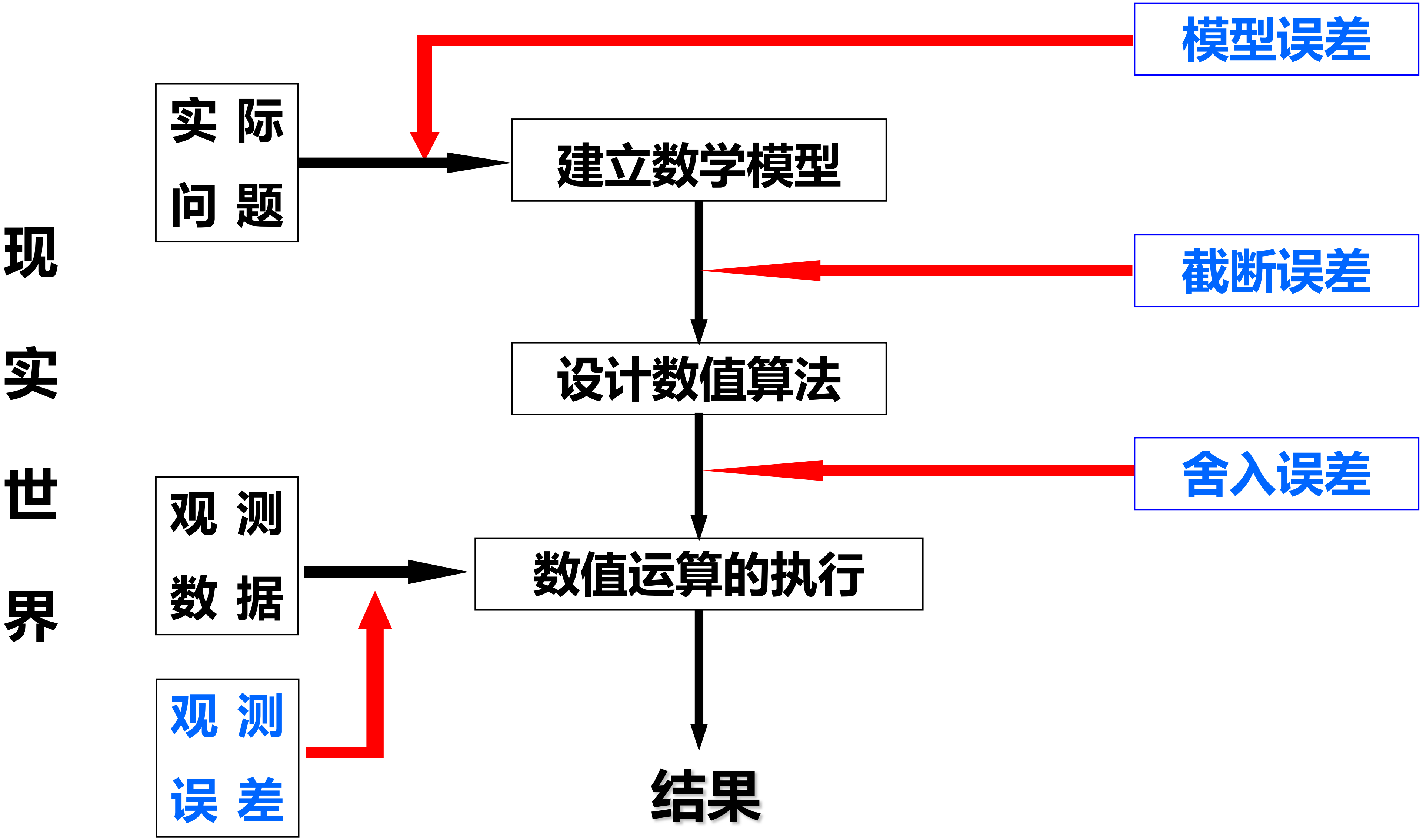
“中国近代数学能够超越西方或与之并驾齐驱的主要原因有三个，主要是讲能够在数学历史上很出名的有三个：一个是陈省身教授在示性类方面的工作，一个是华罗庚在多复变函数方面的工作，一个是冯康在有限元计算方面的工作。”

- 1997年春，丘成桐教授在清华大学所作题为“中国数学发展之我见”的报告

本章目录

- 研究对象与特点
- 误差来源与分析
- 误差的基本概念
- 误差分析的方法与原则

绪论｜误差来源与分析



模型误差

- 在用计算机解决科学计算问题时，首先要建立数学模型，它是对被描述的实际问题进行抽象、简化而得到的，因而它是近似的。数学模型与实际问题之间出现的这种误差称为模型误差。

观测误差

- 在数学模型中往往还有一些根据观测得到的物理量，如温度、长度、电压等，这些参量显然也包含误差，这种由观测产生的误差称为观测误差。

本课程不讨论“模型误差”和“观察误差”

截断误差/方法误差

- 当数学模型不能得到精确解时，通常要用数值方法求它的近似解，其近似解与精确解之间的误差称为截断误差或方法误差。

例：用Taylor多项式近似函数$f(x)$

$$P_n(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n$$

根据泰勒余项定理得截断误差：

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$$

其中ξ在0与x之间

截断误差/方法误差

指数函数：

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

近似函数：

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

截断函数：

$$e^x - P_n(x) = \frac{x^{n+1}}{(n+1)!} e^{\theta x}, \qquad 0 < \theta < 1$$

舍入误差：

- 由于计算机的字长有限，原始数据在计算机上表示会产生误差，计算过程又可能产生新的误差，这种误差称为舍入误差。
- FP64 -> FP32 -> FP16 -> FP8

例：用3.14159近似代替π，产生舍入误差

$$R=\pi-3.14159=0.0000026\ldots$$

其他例子：

$$\sqrt{2}=1.41421356237\cdots\approx 1.4142$$

$$\ln(2)=0.69314718056\cdots\approx 0.6931$$

DeepSeek量化模型部署：

The original DeepSeek R1 671B model is 720GB in size, which is huge. Even a $200k monster like NVIDIA DGX H100 (with 8xH100) can barely hold it. Here I use Unsloth AI's dynamically quantized version, which selectively quantize a few important layers to higher bits, while leaving most of the MoE layers for lower bits. As a result, the model can be quantized to as small as 131GB (1.58-bit), making it much more accessable for local users. It can even run on a single Mac Studio ($5.6k)!

I choose the following two models based on the spec of my workstation:

- `DeepSeek-R1-UD-IQ1_M` (671B, dynamically quantized 1.73-bit, 158 GB, HuggingFace)
- `DeepSeek-R1-Q4_K_M` (671B, standard 4-bit, 404 GB, HuggingFace)

There are four dynamically quantized models from 131GB (1.58-bit) to 212GB (2.51-bit), and you can choose according to your own spec. A detailed introduction of the four models can be found here, which I strongly suggest you to read before the selection.

https://snowkylin.github.io/blogs/a-note-on-deepseek-r1.html

当准确值 x 有多位数时，常四舍五入得到 x 的前几位近似值 x^*

例：$x = \pi = 3.14159265\ ...$

- 取前3位，$x_3^* = 3.14$，$\varepsilon_3^* \le 0.002$

$$e_3^* = |\pi - 3.14| \le 0.002 \le \frac{1}{2}\times 10^{-2}$$

- 取前5位，$x_5^* = 3.1416$，$\varepsilon_5^* \le 0.000008$

$$e_5^* = |\pi - 3.1416| \le 0.000008 \le \frac{1}{2}\times 10^{-4}$$

误差都不超过末位数字的半个单位

综合例子：近似计算 $\int_0^1 e^{-x^2}\,dx \quad (= 0.747\ldots)$

- 将e^{-x^2}作Taylor展开后再积分：

$$\int_0^1 e^{-x^2}\,\mathrm{d}x = \int_0^1 \left(1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots\right)\mathrm{d}x$$

$$= \underbrace{1 - \frac{1}{3} + \frac{1}{2!}\times\frac{1}{5} - \frac{1}{3!}\times\frac{1}{7}}_{S_4} + \underbrace{\frac{1}{4!}\times\frac{1}{9} - \cdots}_{R_4}$$

综合例子：近似计算 $\int_0^1 e^{-x^2}\,dx$ $(= 0.747\ \dots)$

- 取$\int_0^1 e^{-x^2}\,\mathrm{d}x \approx S_4$，则$R_4 = \frac{1}{4!}\times\frac{1}{9} - \frac{1}{5!}\times\frac{1}{11} + \cdots$称为截断误差
- 这里 $|R_4| < \frac{1}{4!}\times\frac{1}{9} < 0.005$

$$S_4 = 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42}$$
$$\approx 1 - 0.333 + 0.1 - 0.024 = 0.743$$

- 舍入误差 $< 0.0005\times 2 = 0.001$
- 总体误差 $< 0.005 + 0.001 = 0.006$

误差分析/估计

- 研究计算结果的误差是否满足精度要求
- 本课程主要讨论算法的截断误差与舍入误差

例1.1 计算$\mathrm{I_n} = \mathrm{e}^{-1}\int_0^1 \mathrm{x^n e^x\, dx}\ (\mathrm{n} = 0, 1, \ldots)$，并估计误差

- 分部积分公式：

$$\int_a^b u(x)v'(x)\,dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)\,dx$$

- 令 $u(x) = x^n, v(x) = e^x$ ，有：

$$I_n = e^{-1} \int_0^1 x^n e^x \,\mathrm{d}x$$

$$= e^{-1} \left([x^n e^x]_0^1 - n \int_0^1 x^{n-1} e^x \,dx \right)$$

$$= e^{-1} \left(e^1 - n \int_0^1 x^{n-1} e^x \,dx \right)$$

$$= 1 - nI_{n-1}$$

- **方案 A**：计算I_n的递推公式：

$$I_n = 1 - nI_{n-1} \quad (n = 1, 2, \dots)$$

$$I_0 = \mathrm{e}^{-1}\int_0^1 e^x \,\mathrm{d}x = 1 - e^{-1}$$

- 为计算I_0，先计算e^{-1}

$$e^{-1} \approx 1 + (-1) + \frac{(-1)^2}{2!} + \cdots + \frac{(-1)^k}{k!}$$

并取$k = 7$，用四位小数计算，则得$e^{-1} \approx 0.3679$，截断误差

$$R_7 = |e^{-1} - 0.3679| \le \frac{1}{8!} < \frac{1}{4}\times 10^{-4}$$

- **方案 A**：计算I_n的递推公式：

$$I_n = 1 - nI_{n-1} \ (n = 1, 2, \dots)$$

$$I_0 = \mathrm{e}^{-1} \int_0^1 e^x \, \mathrm{d}x = 1 - e^{-1}$$

- 初值 $\tilde{I}_0$ 误差：$E_0 = I_0 - \tilde{I}_0$
- 各步误差$E_n = I_n - \tilde{I}_n$满足关系：

$$E_n = -nE_{n-1} \ (n = 1,2, \dots)$$

$$E_n = (-1)^n n! \, E_0$$

误差爆炸💥

n	$\tilde{I}_n$ (A)
0	0.6321
1	0.3679
2	0.2642
3	0.2074
4	0.1704
5	0.1480
6	0.1120
7	0.2160
8	−0.728
9	7.552

$$\frac{e^{-1}}{n+1} = e^{-1}\left(\min_{0\le x\le 1} e^{x}\right)\int_0^1 x^n\,\mathrm{d}x < I_n < e^{-1}\left(\max_{0\le x\le 1} e^{x}\right)\int_0^1 x^n\,\mathrm{d}x = \frac{1}{n+1}$$

- 当$\mathrm{n} = 9$时有$\frac{e^{-1}}{10} < I_9 < \frac{1}{10}$
- 粗略取$I_9 \approx \frac{1}{2}\left(\frac{1}{10} + \frac{e^{-1}}{10}\right) = 0.0684 = I_9^*$

方案 B：

$$\begin{cases} I_9^* = 0.0684, \\ I_{n-1}^* = \dfrac{1}{n}\left(1 - I_{\mathrm{n}}^*\right)\ (n = 9, 8, \ldots 1) \end{cases}$$

绪论 | 误差来源与分析

- 方案（B）由于$|E_0^*| = \frac{1}{n!}|E_n^*|$，$E_0^*$比$E_n^*$缩小了$n!$倍。尽管$E_9^*$较大，但误差逐步缩小
- 方案（A）计算时，初值$\tilde{I}_0$较准确，但误差逐步扩大

n	$\tilde{I}_n$ （A）	I_n^* （B）	n	$\tilde{I}_n$ （A）	I_n^* （B）
0	0.6321	0.6321	5	0.1480	0.1455
1	0.3679	0.3679	6	0.1120	0.1268
2	0.2642	0.2643	7	0.2160	0.1121
3	0.2074	0.2073	8	−0.728	0.1035
4	0.1704	0.1708	9	7.552	0.0684

本章目录

- 研究对象与特点
- 误差来源与分析
- 误差的基本概念
- 误差分析的方法与原则

定义1.1 设x为准确值，x^*为x的近似值，称 $e^* = x^* - x$ 为近似值的绝对误差，简称误差

- e^*可正可负：当绝对误差为正时，叫做强近似值；当绝对误差为负时，叫做弱近似值
- 通常准确值x和误差e^*都是未知的

估计出误差的绝对值不超过某正数 ε^*，称为近似值的误差限

- 总是正数
- $|x - x^*| \le \varepsilon^*$，也可表示为$x = x^* \pm \varepsilon^*$

例：用米刻度的米尺测量一长度x（单位：mm）

- 读出和该长度接近的刻度x^*，x^*是x的近似值，它的误差限是0.5：

$$|x^* - x| \le 0.5$$

- 如读出的长度为765，则有：

$$|765 - x| \le 0.5$$

$$764.5 \le x \le 765.5$$

$$x \in [764.5, 765.5]$$

绪论 | 误差与误差限

- 误差限的大小不能完全表示近似值的好坏
- 在判断近似值好坏时，除考虑误差大小外，还应考虑准确值x本身的大小

例：有两个量 $x = 10 \pm 1$，$y = 1000 \pm 5$，则

$$x^* = 10,\ \varepsilon_x^* = 1,\ y^* = 1000,\ \varepsilon_y^* = 5$$

虽然 ε_y^* 比 ε_x^* 大 4 倍，但

$$\frac{\varepsilon_y^*}{y^*} = \frac{5}{1000} = 0.5\% < \frac{\varepsilon_x^*}{x^*} = \frac{1}{10} = 10\%$$

说明 y^* 近似 y 的程度比 x^* 近似 x 的程度好

- 近似值的误差 e^* 与准确值 x 的比值，称为近似值 x^* 的**相对误差**，记作 e_r^* ：

$$e_r^* = \frac{e^*}{x} = \frac{x^* - x}{x}$$

- 由于 x 未知，通常取

$$e_r^* = \frac{e^*}{x^*} = \frac{x^* - x}{x^*}$$

- 如果 $e_r^* = e^*/x^*$ 较小，这个近似是合理的

$$\frac{e^*}{x} - \frac{e^*}{x^*} = \frac{e^*(x^* - x)}{x^* x} = \frac{(e^*)^2}{x^*(x^* - e^*)} = \frac{(e^*/x^*)^2}{1 - (e^*/x^*)}$$

- 相对误差可正可负，它的绝对值上界称为相对误差限，记作 ε_r^*，若 ε^* 为近似值的误差限，则：

$$\varepsilon_r^* = \frac{\varepsilon^*}{|x^*|}$$

例：有两个量 $x = 10 \pm 1$，$y = 1000 \pm 5$，则其相对误差限为

$$\frac{\varepsilon_y^*}{y^*} = \frac{5}{1000} = 0.5\% < \frac{\varepsilon_x^*}{x^*} = \frac{1}{10} = 10\%$$

说明 y^* 近似 y 的程度比 x^* 近似 x 的程度好

若近似值 x^* 的误差限是某一位的半个单位，该位到 x^* 的第一位非零数字共有 n 位，就说x^*有n位有效数字：

- $x_3^* = 3.14$，3位有效数字；$x_5^* = 3.1416$，5位有效数字

可得标准形式，其中a_1是1到9中的数字；a_2至a_n是0到9中的数字

$$x^* = \pm 10^m \times \left(a_1 + a_2 \times 10^{-1} + \cdots + a_n \times 10^{-(n-1)}\right)$$

$$= \pm a_1 . a_2 \ldots a_n \times 10^m$$

有效数字的误差为:

$$e^* = |x - x^*| \le \frac{1}{2} \times 10^{m-n+1}$$

例1.2 按四舍五入原则写出下列各数具有五位有效数字的近似数

187.9325	0.03785551	8.000033	2.7182818
187.93	0.037856	8.0000	2.7183

例：$x = \sqrt{3} \approx$ 1.732050808，判断下面的有效数字有几位

近似值	有效数字	原因
1.73	3	$\lvert\sqrt{3} - 1.73\rvert < 0.5\times10^{-2} = 0.5\times10^{0-3+1}$
1.7321	5	$\lvert\sqrt{3} - 1.7321\rvert < 0.5\times10^{-4} = 0.5\times10^{0-5+1}$
1.7320	4	$\lvert\sqrt{3} - 1.7320\rvert > 0.5\times10^{-4} = 0.5\times10^{0-5+1}$

例1.2 按四舍五入原则写出下列各数具有五位有效数字的近似数

187.9325	0.03785551	8.000033	2.7182818
187.93	0.037856	8.0000	2.7183

例：$x=\sqrt{3}\approx$ 1.732050808，判断下面的有效数字有几位

近似值	有效数字	原因
1.73	3	$\lvert\sqrt{3}-1.73\rvert<0.5\times10^{-2}=0.5\times10^{0-3+1}$
1.7321	5	$\lvert\sqrt{3}-1.7321\rvert<0.5\times10^{-4}=0.5\times10^{0-5+1}$
1.7320	4	$\lvert\sqrt{3}-1.7320\rvert<0.5\times10^{-4}=0.5\times10^{0-4+1}$

绪论 | 有效数字

- 有效位数与小数点后有多少位数无关
  - 如 101 m 与 0.101 km，都是有3位有效数字
- 绝对误差与误差限是有量纲的
- 相对误差与相对误差限是无量纲的

定理1.1 对于下面的近似数

$$x^* = \pm 10^m \times \left(a_1 + a_2 \times 10^{-1} + \cdots + a_n \times 10^{-(n-1)}\right)$$

若 x^* 具有 n 位有效数字，则其相对误差限为

$$\varepsilon_{\mathrm{r}}^* \le \frac{1}{2a_1} \times 10^{-(n-1)}$$

反之，若 x^* 的相对误差限满足

$$\varepsilon_{\mathrm{r}}^* \le \frac{1}{2(a_1+1)} \times 10^{-(n-1)}$$

则 x^* 至少具有 n 位有效数字

证明： 由x^*的形式可知

$$a_1 \times 10^m \le |x^*| \le (a_1+1)\times 10^m$$

当 x^* 有 n 位有效数字时，有

$$\frac{|x-x^*|}{|x^*|} \le \frac{0.5\times 10^{m-n+1}}{a_1\times 10^m} = \frac{1}{2a_1}\times 10^{-(n-1)}$$

反之，若$\varepsilon_r^* \le \frac{1}{2(a_1+1)}\times 10^{-(n-1)}$，有

$$|x-x^*| \le |x^*|\varepsilon_r^*$$

$$\le (a_1+1)\times 10^m \times \frac{1}{2(a_1+1)}\times 10^{-(n-1)} = \frac{1}{2}\times 10^{m-n+1}$$

- 两个近似数x_1^*和x_2^*，其误差限分别为$\varepsilon(x_1^*)$和$\varepsilon(x_2^*)$，则

$$\varepsilon(x_1^* \pm x_2^*) = \varepsilon(x_1^*) + \varepsilon(x_2^*)$$

$$\varepsilon(x_1^* x_2^*) \approx |x_1^*|\varepsilon(x_2^*) + |x_2^*|\varepsilon(x_1^*)$$

$$\varepsilon\left(\frac{x_1^*}{x_2^*}\right) \approx \frac{|x_1^*|\varepsilon(x_2^*) + |x_2^*|\varepsilon(x_1^*)}{|x_2^*|^2} \quad (x_2^* \neq 0)$$

- 当自变量有误差时，计算函数值误差限可用Taylor展开式来估计

绪论｜一元函数的误差估计

- 设 $f(x)$ 是一元函数，x 的近似值为 x^*，根据Taylor展开

$$f(x)-f(x^*)=f'(x^*)(x-x^*)+\frac{f''(\xi)}{2}(x-x^*)^2$$

- 其中 ξ 介于 x 与 x^* 之间，上式取绝对值得

$$|f(x)-f(x^*)|\leqslant|f'(x^*)|\varepsilon(x^*)+\frac{|f''(\xi)|}{2}\varepsilon^2(x^*)$$

- $f'(x^*)$ 与 $f''(x^*)$ 的比值不太大，误差限$\varepsilon(f(x^*))$近似为

$$\varepsilon(f(x^*))\ \approx|f'(x^*)|\varepsilon(x^*)$$

绪论｜多元函数的误差估计

- 当f为多元函数时，计算 $A = f(x_1, x_2, \cdots, x_n)$
- 如果 $x_1, x_2, \cdots, x_n$ 的近似值为 $x_1^*, x_2^*, \cdots, x_n^*$，则 A 的近似值为

$$A^* = f(x_1^*, x_2^*, \cdots, x_n^*)$$

- 函数值 A^* 的误差 $e(A^*)$ 由Taylor展开得

$$e(A^*) = A^* - A = f(x_1^*, x_2^*, \cdots, x_n^*) - f(x_1, x_2, \cdots, x_n)$$

$$\approx \sum_{k=1}^{n} \left(\frac{\partial f(x_1^*, x_2^*, \cdots, x_n^*)}{\partial x_k} \right) (x_k^* - x_k) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k} \right)^* e_k^*$$

绪论 | 多元函数的误差估计

- 函数值 A^* 的误差 $e(A^*)$ 由Taylor展开得

$$e(A^*) = A^* - A = f(x_1^*, x_2^*, \cdots, x_n^*) - f(x_1, x_2, \cdots, x_n)$$

$$\approx \sum_{k=1}^{n} \left(\frac{\partial f(x_1^*, x_2^*, \cdots, x_n^*)}{\partial x_k} \right) (x_k^* - x_k) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k} \right)^* e_k^*$$

- A^*的误差限为：

$$\varepsilon(A^*) \approx \sum_{k=1}^{n} \left| \left(\frac{\partial f}{\partial x_k} \right)^* \right| \varepsilon(x_k^*)$$

- A^*的相对误差限：

$$\varepsilon_r^* = \varepsilon_r(A^*) = \frac{\varepsilon(A^*)}{|A^*|} \approx \sum_{k=1}^{k} \left| \left(\frac{\partial f}{\partial x_k} \right)^* \right| \frac{\varepsilon(x_k^*)}{|A^*|}$$

绪论 | 多元函数的误差估计

例1.5：测得某场地长 l 的值为 $l^* = 110\ \mathrm{m}$，宽 d 的值为 $d^* = 80\ \mathrm{m}$，已知 $|l - l^*| \le 0.2\ \mathrm{m}$，$|d - d^*| \le 0.1\ \mathrm{m}$，试求面积 $S = ld$ 的绝对误差限和相对误差限

- 由多元函数误差限公式，知

$$\varepsilon(S^*) \approx \left|\left(\frac{\partial S}{\partial l}\right)^*\right| \varepsilon(l^*) + \left|\left(\frac{\partial S}{\partial d}\right)^*\right| \varepsilon(d^*)$$

- 其中$\left(\frac{\partial S}{\partial l}\right)^* = d^* = 80\ \mathrm{m}$，$\left(\frac{\partial S}{\partial d}\right)^* = l^* = 110\ \mathrm{m}$，$\varepsilon(l^*) = 0.2\ \mathrm{m}$，$\varepsilon(d^*) = 0.1\ \mathrm{m}$

- 故绝对误差限为：$\varepsilon(S^*) \approx (80\times 0.2 + 110\times 0.1)\ \mathrm{m}^2 = 27\ \mathrm{m}^2$

- 相对误差限为：$\varepsilon_{\mathrm{r}}(S^*) = \dfrac{\varepsilon(S^*)}{|S^*|} = \dfrac{\varepsilon(S^*)}{l^* d^*} \approx \dfrac{27}{8800} = 0.31\%$

本章目录

- 研究对象与特点
- 误差来源与分析
- 误差的基本概念
- 误差分析的方法与原则

绪论 | 误差分析的方法

实际工程或科学计算问题：

- 往往要运算千万次，每步都作误差分析是不可能的
- 误差积累有正有负，绝对值有大有小，都按最坏情况估计得到的结果比实际误差大得多

概率分析法

- 将数据/运算中的舍入误差视为适合某种分布的随机变量，确定结果的误差分布

误差定性分析

- 数值稳定性：运算过程舍入误差不增长的计算公式是数值稳定的，否则不稳定

例1.1中 方案（A）是数值不稳定的

例1.1中 方案（B）是数值稳定的

研究一个计算公式是否稳定

- 假定初始值有误差 ε_0，中间不再产生新误差，考察由 ε_0 引起的误差积累是否增长，如不增长就认为是稳定的，否则是不稳定的
- 对于稳定的计算公式，不具体估计舍入误差积累也可相信它是可用的，误差限不会太大
- 不稳定的公式通常就不能使用，如要使用，其计算步数也只能很少，并且注意控制累计误差

原则1：要避免除数绝对值远小于被除数绝对值

- 用绝对值小的数作除数，舍入误差会增大
- 计算 $\frac{x}{y}$，若$0 < |y| \ll |x|$

$$\varepsilon\left(\frac{x}{y}\right) \approx \frac{|x|\varepsilon(y) + |y|\varepsilon(x)}{|y|^2} = \frac{|x|\varepsilon(y)}{|y|^2} + \frac{\varepsilon(x)}{|y|}$$

- 表明当$|y|$相对太小时，商的绝对误差可能很大

例1.6 线性方程组 $\begin{cases} 0.00001x_1 + x_2 = 1 \\ 2x_1 + x_2 = 2 \end{cases}$ **的准确解为**

$$x_1 = \frac{200000}{399999} = 0.50000125,$$
$$x_2 = \frac{199998}{199999} = 0.999995.$$

- 若用四位浮点十进制数下用消去法求解

$$\begin{cases} 10^{-4}\times 0.1000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.1000 \\ 10^1\times 0.2000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.2000 \end{cases}$$

绪论 | 误差分析的原则：原则1

$$\begin{cases} 10^{-4}\times 0.1000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.1000 & ① \\ 10^1\times 0.2000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.2000 & ② \end{cases}$$

- 如用 ① $/(10^{-4}\times 0.1000/2)$， 此处“大数除小数”，得

$$10^1\times 0.2000x_1 + 10^6\times 0.2000x_2 = 10^6\times 0.2000 \qquad ③$$

- ③ − ② ，消除x_1，得

$$10^6\times 0.2000x_2 - 10^1\times 0.1000x_2 = (10^6 - 10^1)\times 0.2000$$

- 精度有限，忽略小数字

$$\begin{cases} 10^{-4}\times 0.1000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.1000 \\ 10^6\times 0.2000x_2 = 10^6\times 0.2000 \end{cases}$$

- 解得$x_1 = 0$，$x_2 = 1$，严重失真

绪论丨误差分析的原则：原则1

$$\begin{cases} 10^{-4}\times 0.1000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.1000 & ① \\ 10^1\times 0.2000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.2000 & ② \end{cases}$$

- 如用 ② − ① ，消除x_2，得

$$10^1\times 0.2000x_1 - 10^{-4}\times 0.1000x_1 = 10^1\times 0.1000$$

- 精度有限，忽略小数字

$$\begin{cases} 10^1\times 0.2000x_1 = 10^1\times 0.1000 \\ 10^1\times 0.2000x_1 + 10^1\times 0.1000x_2 = 10^1\times 0.2000 \end{cases}$$

- 由此解出$x_1 = 0.5$，$x_2 = 1$，基本正确

原则2：要避免两相近数相减

- $x=532.65$，$y=532.52$有五位有效数字，但$x-y=0.13$只有两位有效数字
- 令$z=y-x$，则$\varepsilon(z)=\varepsilon(y)+\varepsilon(x)$

$$\varepsilon_{\mathrm{r}}(z)=\frac{\varepsilon(z)}{|z|}=\frac{\varepsilon(y)+\varepsilon(x)}{|z|}$$

$$=\frac{|y|}{|z|}\varepsilon_{\mathrm{r}}(y)+\frac{|x|}{|z|}\varepsilon_{\mathrm{r}}(x)$$

- 当$y\approx x$时，$z\approx 0$，相对误差限会很大

例1.7：计算 $A = 10^7(1 - \cos 2^\circ)$

- 若利用 $\cos 2^\circ = 0.9994$，直接计算得

$$A = 10^7\left(1 - \cos 2^\circ\right) = 10^7(1 - 0.9994) = 6\times10^3$$

只有1位有效数字

- 若利用$1 - \cos x = 2\sin^2\frac{x}{2}$计算，则有

$$A = 10^7\left(1 - \cos 2^\circ\right) = 2\times\left(\sin 1^\circ\right)^2\times10^7 = 6.13\times10^3$$

有3位有效数字，其中 $\sin 1^\circ = 0.0175$

解决方法1：改变计算公式

- 如果x_1和x_2很接近

$$\lg x_1 - \lg x_2 = \lg \frac{x_1}{x_2}$$

- 当x很大时

$$\sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1} + \sqrt{x}}$$

- 若$f(x) \approx f(x^*)$，可用 Taylor 展开

$$f(x) - f(x^*) = f'(x^*)(x - x^*) + \frac{f''(x^*)}{2}(x - x^*)^2 + \cdots$$

解决方法2：增加有效位数进行运算

- half 变 float，float 变 double

原则3：要防止大数“吃掉”小数

大数“吃掉”小数的现象，影响结果可靠性

- 参加运算的数有时数量级相差很大
- 计算机位数有限，如不注意运算次序

解决方案

- 按绝对值由小到大的顺序累加
- 合理分组，保证数量级大致相同

例1.8：在五位十进制计算机上，计算

$$A = 52492 + \sum_{i=1}^{1000} \delta_i$$

其中$0.1 \le \delta_i \le 0.9$

- 把运算的数写成规格化形式，有

$$A = 0.52492 \times 10^5 + \sum_{i=1}^{1000} \delta_i$$

- 计算时要对阶，若取 $\delta_i = 0.9$，对阶时$\delta_i = 0.000009 \times 10^5$
- 在五位的计算机中表示为0，计算结果为52492，不可靠

绪论 | 误差分析的原则：原则3

- 如果先把数量级相同的1000个δ_i相加，最后再加上52492，这时有

$$0.1\times10^3 \le \sum_{\mathrm{i=1}}^{1000} \delta_i \le 0.9\times10^3$$

- 则有

$$0.001\times10^5 + 0.52492\times10^5 \le A \le 0.009\times10^5 + 0.52492\times10^5$$

$$52592 \le A \le 53392$$

原则4：注意简化计算步骤，减少运算次数

减少运算次数

- 节省计算机的计算时间，还能减小舍入误差
- 是数值计算必须遵从的原则

例1.9 计算x^{255}的值

- 逐个相乘，要用254次乘法
- 若用$x^{255} = x \cdot x^2 \cdot x^4 \cdot x^8 \cdot x^{16} \cdot x^{32} \cdot x^{64} \cdot x^{128}$，则只需要14次运算

绪论丨误差分析的原则：秦九韶算法

例：计算多项式 $P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$

- 直接结算$a_k x^k$再逐项相加，一共需做n次加法及以下次数乘法

$$n + (n-1) + \cdots + 2 + 1 = \frac{n(n+1)}{2}$$

- 秦九韶算法： n次乘法和n次加法

$$P_n(x) = (\ldots((a_n x + a_{n-1})x + a_{n-2})x + \cdots + a_1)x + a_0$$

$$\begin{cases} S_n = a_n \\ S_k = xS_{k+1} + a_k, \qquad (k = n-1, n-2, \ldots, 0) \\ P_n(x) = S_0 \end{cases}$$

绪论丨误差分析的原则：秦九韶算法

- **秦九韶（约公元1202年 - 1261年），南宋末年人，出生于鲁郡（今山东曲阜一带人）**

秦九韶聪敏勤学，宋绍定四年（公元1231），秦九韶考中进士，先后担任县尉、通判、参议官、州守等职。

他在政务之余，以数学为主线进行潜心钻研，且应用范围至为广泛：天文历法、水利水文、建筑、测绘、农耕、军事、商业金融等方面。

秦九韶是我国古代数学家的杰出代表之一，他的《数书九章》概括了宋元时期中国传统数学的主要成就，尤其是系统总结和发展了高次方程的数值解法与一次同余问题的解法，提出了相当完备的“正负开方术”和“大衍求一术”。对数学发展产生了广泛的影响。

https://baike.baidu.com/item/%E7%A7%A6%E4%B9%9D%E9%9F%B6%E7%AE%97%E6%B3%95/449196

绪论 | 误差分析的原则：秦九韶算法

宋淳祐四至七年（公元1244至1247），秦九韶在湖州为母亲守孝三年期间，把长期积累的数学知识和研究所得加以编辑，写成了举世闻名的数学巨著《数书九章》。 书成后，并未出版。原稿几乎流失，书名也不确切。后历经宋、元，到明建国，此书无人问津，直到明永乐年间，在解缙主编《永乐大典》时，记书名为《数学九章》。又经过一百多年，经王应麟抄录后，由王修改为《数书九章》。

19世纪初，英国数学家威廉·乔治·霍纳重新发现并证明，后世称作霍纳算法（Horner's method、Horner scheme）。但是，19世纪英国传教士伟烈亚力Alexander Wylie. (1815–1887) 最早对霍纳的发明权提出质疑。……

他被国外科学史家称为是“他那个民族，那个时代，并且确实也是所有时代最伟大的数学家之一。”

研究对象与特点

- 面向计算机、有可靠理论、容易计算、有实验

误差来源与误差分析

- 截断误差、舍入误差

误差的基本概念

- 误差限、相对误差限、有效数字、误差估计

误差分析的方法与原则

- 数值稳定、四个原则、秦九韶算法

绪论 | 作业

第一章 习题 2, 5, 11, 13, 15