科学与工程计算基础

任课教师：黄忠亿

清华大学数学科学系

参考书目

1. 《数值分析基础》，关治、陆金甫编著，高等教育出版社
2. 《Elements of Numerical Analysis》, Peter Henrici, John Wiley & Sons, Inc.
3. 《An Introduction to Numerical Analysis》, Endre Süli and David F. Mayers, Cambridge University Press

目录

引言

《科学与工程计算基础（数值分析）》这门课是科学与工程计算领域的入门课程，也是极为重要的一门课。

这门课主要研究各种科学与工程中的数学物理问题的求解，包括数值方法的设计、分析以及有关的数学理论和具体实现等。

科学界目前公认的有三大类科学研究方法：理论方法、实验方法和科学计算方法。这三类方法相辅相成、互为补充。

可以说，基础理论只解决了适定性，但没给出（尤其是复杂问题）具体求解方法, 这里面仍然包含了大量数学上的 困难和挑战. 实验方法中某些特定条件实现起来过于昂贵、有些条件过于苛刻难于达到，也使得目前许多实验都用数值模拟实验来代替。

引言

下面我们通过几个例子来看看本课程所研究的内容有哪些。

例 1.1 (定积分的计算)

我们知道定积分的计算可以使用 *Newton-Leibnitz* 公式， 即假设 $F(x)$ 为 $f(x)$ 的一个原函数，那么

$$\int_a^b f(x)dx = F(b) - F(a).$$

但仅有极少数函数的原函数是可以由简单的初等函数来表示的，换句话说，几乎所有的函数的原函数都没有简单的解析表达式.

因此可以说 Newton-Leibnitz 公式虽然具有理论意义，但对于大量需要计算的定积分，我们仍需要设计数值积分方法。

引言

例 1.2 (线性方程组求解)

对于n个未知数的线性方程组 $Ax=b$，只要$\det A \neq 0$, 该方程组存在唯一解.

我们当然可用*Cramer's*法则来计算其解, 但需要计算 $n+1$ 个行列式. 若按定义来计算，其计算量为 $(n+1)!$

例如 $n=100$, 那么我们需要约9.4×10^{159}次浮点运算.

即便用神威太湖之光来计算(持续计算性能为每秒 9.3×10^{16} 次双精度浮点运算), 也需要1.01×10^{143}秒$\approx 3.2\times10^{127}$亿年（目前公认为宇宙年龄不过*100*多亿年）。

因而大规模线性方程组的求解必须要设计快速算法，而不能使用传统教科书中的简单方法。

引言

例 1.3 (多项式求根)

代数基本定理告诉我们，复系数一元 n 次多项式有 n 个复根．然而我们也知道，五次以上的代数方程便没有一般求根公式. 这就告诉我们，即便是求一元多项式的根，《线性代数》课程也没有告诉我们可行的求解方法，必须设计新的办法来求解, 而许多科学与工程计算问题中都需要求解多项式的根。

从这几个例子我们可以看到，大量实际问题我们无法得到其解析解的表达式（有时候即便可以写成积分或者级数形式，实际计算时还需要设计快速算法）. 因此我们这门课需要研究如何有效（用合理的计算量）并高精度地（相对小的误差）得到一个问题的近似解 (原因是其解析解无法显式给出)。

目录

误差来源及分类

从上面例子我们已经看到, 通常我们在科学与工程计算问题中需要在事先给定的精度要求下，通过构造性的方法给出在一定计算量下的近似解. 由于解析解通常无法得到，所以我们得到的解总是存在一定的误差。

有以下两个问题需要特别关注：

1. 如何估计近似解与原精确解之间的误差大小？
2. 如何保证算法的数值稳定性：即控制误差的传递，使得不会由于计算步骤的增加，误差无限制地增长。

误差来源及分类

下面我们先看看误差的来源（或者说分类）. 误差通常有以下四类（前两类在本课程中暂不考虑，我们主要考虑后两类）：

1. 模型误差（即建模时会舍弃部分次要因素而产生的误差）
2. 测量误差（即测量数据如温度、压强、速度等带来的误差）
3. 截断误差(truncation error)：这是由于使用近似问题代替原问题或者使用近似方法求解而产生的误差, 如

   $|x| \ll 1$时，可以用 x 近似代替 $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$

4. 舍入误差(roundoff error)：这是由于计算机字长有限带来的浮点数（实数）运算产生的误差， 如 $\pi = 3.1415926535897\cdots$ 是无理数，计算机字长有限，无法准确表示它，只能截断， 如果用6位数字表示即 3.14159，那么误差即为 $error = 0.0000026535\cdots$

目录

误差的基本概念

定义 1.1 (误差的定义)

设 x^* 为 x 的近似值，称 $e^* = x - x^*$ 为近似值 x^* 的误差, $|x - x^*|$ 称为 x^* 绝对误差. 如果 $x \neq 0$，称 $\dfrac{|x - x^*|}{|x|}$ 为 x^* 的相对误差。

一般来说 x 的值并不知道, 因此只能估计出误差绝对值的一个范围, 即 $|x - x^*| \leq \eta$, 此时称 η 为 x^* 的绝对误差限(界), 记为 $\varepsilon(x^*)$. 同样，若 $x^* \neq 0$，称 $\varepsilon_r(x^*) = \dfrac{\varepsilon(x^*)}{|x^*|} = \dfrac{\eta}{|x^*|}$ 为 x^* 的相对误差限。

看一个例子：

例 1.4

取 $\pi^* = 3.14159$, 有 $e^* = \pi - \pi^* = 0.0000026535\cdots$ 可以取其一个绝对误差限为 $\eta = 3 \times 10^{-6}$, 相对误差限为 1×10^{-6}.

误差的基本概念

定义 1.2 (有效数字)

设 x 的一个近似值 x^* 可写成

$$x^* = \pm 10^m \times 0.d_1 d_2 \cdots d_i \cdots$$

其中 d_i 为 $0 \sim 9$ 之间的数字. 设 $d_1 \neq 0$. 如果 $n \geq 0$ 是使得 $|x - x^*| \leq \frac{1}{2} \times 10^{m-n}$ 成立的最大非负整数, 则称x^*具有n位有效数字. （即其绝对误差不超过其第 n 位数位所表示数值的一半）

例 1.5

取 $\pi_3^* = 3.14$，显然有 $|\pi - \pi_3^*| \leq \frac{1}{2} \times 10^{-2}$, 因为 $m = 1$, 即 $n = 3$, 即 π_3^* 有三位有效数字. 如果取 $\tilde{\pi} = \frac{22}{7} = 3.142857\cdots$, 由于 $|\pi - \tilde{\pi}| \leq 0.002 < \frac{1}{2} \times 10^{-2}$, 即 $\tilde{\pi}$ 也有三位有效数字.

若取 $\pi_5^* = 3.1416$, 显然 π_5^* 有五位有效数字。

相对误差限与有效数字的关系

关于相对误差限与有效数字的关系，我们有以下定理

定理 1.1 (有效位与相对误差的关系)

设 x 的近似值 $x^* = \pm 10^m \times \sum_{j=1}^{l} (\alpha_j \times 10^{-j})$, 且 $\alpha_1 \neq 0$. 若 x^* 有n 位有效数字，则其有相对误差限

$$\varepsilon_r^* \leq \frac{1}{2\alpha_1} \times 10^{1-n};$$

反之，若 x^* 的相对误差限满足

$$\varepsilon_r^* \leq \frac{1}{2(\alpha_1 + 1)} \times 10^{1-n},$$

则 x^* 至少有n 位有效数字.

相对误差限与有效数字的关系

看一个例子：

例 1.6

要使 $\sqrt{2}$ 的近似值相对误差小于 0.1%，需要取几位有效数字？

$\triangleleft$ 假设取 n 位有效数字，由上述定理有 $\varepsilon_r^* \leq \frac{1}{2\alpha_1} \times 10^{1-n}$.

因为 $\alpha_1 = 1$, 可知取 $n = 4$ 即可.

此时 $\varepsilon_r^* \leq \frac{1}{2} \times 10^{-3} < 0.1\%$.

即相当于取 $\sqrt{2} \approx 1.414$ $\triangleright$

误差的运算与传递

若知道了 x_1 与 x_2 的两个近似值 x_1^* 与 x_2^* 的误差限 $\varepsilon(x_1^*)$ 与 $\varepsilon(x_2^*)$, 便可以计算其四则运算的误差限:

- $\varepsilon(x_1^* \pm x_2^*) = \varepsilon(x_1^*) + \varepsilon(x_2^*)$;
- $\varepsilon(x_1^* \cdot x_2^*) \approx |x_1^*| \cdot \varepsilon(x_2^*) + |x_2^*| \cdot \varepsilon(x_1^*)$;
- $\varepsilon\left(\dfrac{x_1^*}{x_2^*}\right) \approx \dfrac{|x_1^*| \cdot \varepsilon(x_2^*) + |x_2^*| \cdot \varepsilon(x_1^*)}{|x_2^*|^2}$, 若 $x_2^* \neq 0$.

$\lhd$

$$
\begin{aligned}
x_1x_2 - x_1^*x_2^* &= x_1(x_2 - x_2^*) + (x_1 - x_1^*)x_2^*;\\
\frac{x_1}{x_2} - \frac{x_1^*}{x_2^*} &= \frac{x_1x_2^* - x_2x_1^*}{x_2x_2^*} = \frac{x_1^*(x_2^* - x_2) + (x_1 - x_1^*)x_2^*}{x_2x_2^*}
\end{aligned}
$$

$\rhd$

误差的运算与传递

更一般的，设 f 为一多元函数，$A = f(x_1, x_2, \cdots, x_n)$，$A^* = f(x_1^*, x_2^*, \cdots, x_n^*)$. 我们有

$$e(A^*) = A - A^* \approx \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x_1^*, x_2^*, \cdots, x_n^*) \cdot (x_k - x_k^*).$$

即

$$\varepsilon(A^*) \approx \sum_{k=1}^{n} |\frac{\partial f}{\partial x_k}(x_1^*, x_2^*, \cdots, x_n^*)| \cdot \varepsilon(x_k^*).$$

类似地，关于相对误差限有

$$\varepsilon_r(A^*) \approx \frac{\varepsilon(A^*)}{|A^*|} \approx \sum_{k=1}^{n} |\frac{\partial f}{\partial x_k}(x_1^*, x_2^*, \cdots, x_n^*)| \cdot \frac{\varepsilon(x_k^*)}{|A^*|}$$

算法稳定性

从前面的分析中已经看到，误差在运算过程中会传播.

而且不同算法的效果会有很大差别（从下例可以看出）：

例 1.7

计算积分 $I_n = \int_0^1 x^n e^{x-1} dx$, $n = 0, 1, \cdots, N$, 并估计误差。

◁ 由分部积分公式

$$I_n = \int_0^1 x^n de^{x-1} = x^n e^{x-1}\Big|_0^1 - \int_0^1 e^{x-1} n x^{n-1} dx = 1 - nI_{n-1} \qquad (1.1)$$

另外易得 $I_0 = \int_0^1 e^{x-1} dx = 1 - e^{-1}$. 这样我们可以利用递推公式 $I_n = 1 - nI_{n-1}$, $n = 1, 2, \cdots, N$ 进行计算. 利用Taylor展开式 $I_0 = 1 - e^{-1} \approx 1 - [1 + (-1) + \frac{(-1)^2}{2!} + \cdots + \frac{(-1)^k}{k!}]$, 取 $k = 7$.

算法稳定性

(续前页)其误差约为 $\frac{1}{8!} \le \frac{1}{4}\times 10^{-4}$, 即 $I_0 \approx 0.6321$ 有四位有效数字.

通过计算会得到 $I_8^* = -0.7280,\ I_9^* = 7.552.$

而由 I_n 的表达式，显然被积函数在 0,1 之间，因而有 $0 < I_n < 1$.

这说明此算法中舍入误差影响增长迅速，不稳定。

我们也可把递推公式(1.1)反过来用，即 $I_{n-1} = \frac{1-I_n}{n}, n = N, \cdots, 1$

这样我们需要先给出 I_N 的估计值。 可利用下式

$$\frac{e^{-1}}{N+1} = e^{-1} \min_{0\le x\le 1} e^x \int_0^1 x^N dx \le \int_0^1 x^N e^{x-1} dx \le e^{-1} \max_{0\le x\le 1} e^x \int_0^1 x^N dx = \frac{1}{N+1}$$

例如我们可取一个很粗糙的初值 $\widetilde{I}_9 = \frac{1}{2}(\frac{e^{-1}}{10} + \frac{1}{10}) \approx 0.0684,$

可以得到 $\widetilde{I}_8 = 0.1035, \cdots, \widetilde{I}_1 = 0.3679,\ \widetilde{I}_0 = 0.6321.$

算法稳定性

(续前页)事实上 $I_0 = 0.6321,\ I_1 = 0.3679,\ I_8 = 0.1009,\ I_9 = 0.0916.$ 也就是说我们第二个算法得到的 $\widetilde{I}_0$ 与 I_0 之间的误差小于 10^{-4}.

分析一下这两个算法的误差我们可以得到:

$$|e_n^*| = n!|e_0^*|, \quad |\widetilde{e}_0| = \frac{1}{n!}|\widetilde{e}_n|.$$

也就是说第一个算法的初始误差的影响会按阶乘倍数恶性增长, 而第二个算法的初始误差的影响会按阶乘倍数缩小。

所以说，第一个算法是不稳定的，而第二个算法是稳定的。 ▷

定义 1.3 (算法稳定性)

若在计算过程中，存在一个与计算步数无关的常数 $C > 0$, 使得计算 n 步之后的误差可以被初始数据的误差的常数倍控制, 即 $|\varepsilon_n| \le C|\varepsilon_0|$，则称该算法数值稳定；否则就称之为不稳定。

如何规避误差带来的危害

我们有以下几个基本原则：

1. 避免将绝对值很大的数除以绝对值很小的数，由于计算机字长有限，否则会丢失有效数位；
2. 避免将两个相近的数相减，否则也会丢失有效数位；
   例如求 $x^2-16x+1=0$ 的小正根: $x_{1,2}=8\pm\sqrt{63}$. 如果用只有三位字长的计算器, $x_2=8-\sqrt{63}\approx 8.00-7.94=0.06=x_2^*$ 会只有一位有效数字. 若改成

$$x_2=\frac{1}{8+\sqrt{63}}\approx\frac{1}{8+7.94}=0.0627$$

仍然可以有三位有效数字。

如何规避误差带来的危害

3. 避免将绝对值很大的数与绝对值很小的数直接相加（减）；

例如计算 $A = 10000 + \sum_{i=1}^{1000} \delta_i$，假设每个 $|\delta_i| \le 0.4$

如果使用五位有效数字计算，取 $\delta_i = 0.4 = 0.000004 \times 10^5$，假如依次相加

$A = 0.10000 \times 10^5 + 0.000004 \times 10^5 + \cdots = 0.10000 \times 10^5$，

显然严重失真. 若改变次序，先算求和，再与前面大数相加

$A = 0.10000 \times 10^5 + \sum_{i=1}^{1000} \delta_i =$

$0.10000 \times 10^5 + 1000 \times 0.000004 \times 10^5 = 0.10400 \times 10^5$，

这样就不会失真。

如何规避误差带来的危害

4. 注意简化步骤，减少运算次数（以减少舍入误差影响）

如计算 $P_n(x) = \sum_{k=0}^{n} a_k x^k$. 若直接逐项计算，计算 x^k 与 a_k 相乘再相加，则需要 $\frac{n(n+1)}{2}$ 次乘法，n 次加法。

若使用秦九韶算法：

$$S_n = a_n, \quad S_{k-1} = xS_k + a_{k-1}, \ k = n, n-1, \cdots, 1.$$

最终有 $P_n(x) = S_0$. 这样只需要 n 次乘法与 n 次加法。

当 n 较大时，这样可以大大减少计算量。

如何规避误差带来的危害

再看一个例子以说明不同算法计算量的巨大差别:

例 1.8

我们可以利用泰勒展开式

$$\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}$$

的前 m 项部分和来计算 $\ln(1+x)$ 的近似值.

例如, 欲计算$\ln 2$, 令$x=1$, 若想精确到10^{-5}, 则需要十万项求和.

而如果我们利用公式

$$\ln\frac{1+x}{1-x}=2\sum_{n=1}^{\infty}\frac{x^{2n-1}}{2n-1}$$

同样计算$\ln 2$, 取 $x=\frac{1}{3}$, 只需计算前9项, 截断误差已经小于 10^{-10}.

本课程的特点与学习要点

总体来说，本课程致力于

- 介绍各类专题中的基本问题与求解方法;
- 强调分析与理解各种方法的本质特点。

因此我们重点介绍各专题中最常用的部分算法及其原理，将更进一步的讨论留给学生课后进行。

要学好本课程, 不能沉迷于背记各种公式, 而要注意以下几点:

- 熟练掌握相关数学知识(微积分、线性代数、常微分方程等);
- 明确基本概念及常用算法的推导技巧、特点、共性等;
- 充分理解各部分内容之间的相关性，做到融会贯通.

目录

矩阵特征值问题

设 $A \in \mathbb{C}^{n\times n}$，若$\exists\lambda \in \mathbb{C}$ 及 $0 \neq \mathbf{x} \in \mathbb{C}^n$, *s.t.*

$$A\mathbf{x} = \lambda\mathbf{x}$$

则称 λ 为 A 的一个特征值, $\mathbf{x}$ 称为属于特征值 λ 的特征向量。计算矩阵 A 的特征值 λ 与特征向量 $\mathbf{x}$ 的问题称为矩阵特征值问题.

由 $A\mathbf{x} = \lambda\mathbf{x} \Longleftrightarrow (A-\lambda I)\mathbf{x} = 0$ 有非零解 $\mathbf{x} \in \mathbb{C}^n$, 其充要条件是

$$\det(A - \lambda I) = 0. \tag{1.2}$$

矩阵特征值问题

上述方程(1.2)称为 A 的特征方程, 行列式 $\det(A-\lambda I)$ （是一个 λ 的 n 次多项式）称为 A 的特征多项式. 显然特征方程(1.2)是 λ 的 n 次方程, (考虑重数意义下) 它有 n 个根 $\lambda_i \in \mathbb{C}, i=1,\cdots,n$.

A 的所有特征值的集合称为 A 的谱，记为 $\sigma(A)$:

$$\begin{aligned}\sigma(A) &= \{\lambda_i \in \mathbb{C} \,|\, \exists\, 0 \neq \mathbf{x}_i \in \mathbb{C}^n,\ s.t.\ A\mathbf{x}_i = \lambda_i \mathbf{x}_i\} \\ \text{(或者说)} &= \{\lambda_i \in \mathbb{C} |\, \det(\lambda_i - A) = 0\}.\end{aligned}$$

令 $\rho(A) = \max\limits_{\lambda \in \sigma(A)} |\lambda|$, 称之为 A 的谱半径.

矩阵特征值问题

A 的对角元之和称为 A 的迹 (trace), 记为 $tr(A) = \sum_{i=1}^{n} a_{ii}$.

由上述定义可知

$$tr(A) = \sum_{i=1}^{n} \lambda_i, \quad \det(A) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \cdots \lambda_n.$$

设 $A, B \in \mathbb{C}^{n\times n}$, 若∃非奇异矩阵 $P \in \mathbb{C}^{n\times n}$, s.t. $B = P^{-1}AP$, 则称 A, B 相似. 相似矩阵具有相同的特征多项式和谱。

显然 A^T 与 A 也具有相同的谱。

矩阵特征值问题

如果存在 $P \in \mathbb{C}^{n\times n}$, s.t. $A = P\Lambda P^{-1}$, 其中 $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, 则称 A 可对角化。 我们有以下定理

定理 1.2

A 可对角化的充要条件是 A 有 n 个线性无关的特征向量。

$\triangleleft$ $AP = P\Lambda$. 记 $P = (P_1\ P_2\ \cdots P_n)$ 写成列向量的形式，即有 $AP_i = \lambda_i P_i$. 而 P 可逆即 $P_1, \cdots, P_n$ 线性无关。 $\triangleright$

如果 A 为实对称矩阵或者 Hermite 阵 ($\bar{A}^T = A$), 则 A 可对角化, 且其特征值均为实数.

矩阵对角化

定理 1.3

属于 A 的不同特征值的特征向量是线性无关的. 若 A 有 n 个不同的特征值（即其特征方程全是单根）, 则 A 可对角化.

$\triangleleft$ 设 $\mathbf{x}, \mathbf{y}$ 是矩阵 A 对应不同特征值的特征向量，即

$$A\mathbf{x} = \lambda\mathbf{x},\ A\mathbf{y} = \mu\mathbf{y},\ \lambda \neq \mu,\ \mathbf{x} \neq 0,\ \mathbf{y} \neq 0.$$

若 $\exists\alpha, \beta$ 不全为零, *s.t.* $\alpha\mathbf{x} + \beta\mathbf{y} = 0$, 显然 $\alpha \neq 0 \neq \beta$, 即 $\mathbf{x} = -\frac{\beta}{\alpha}\mathbf{y}$,

$$\text{又}\quad 0 = A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y} = \alpha\lambda\mathbf{x} + \beta\mu\mathbf{y}$$

由 $\lambda \neq \mu$，则 $\mathbf{x} = -\frac{\beta}{\alpha}\frac{\mu}{\lambda}\mathbf{y}$，与上面矛盾了. $\triangleright$

特征值的代数重数与几何重数

若矩阵 A 的特征值有重根，即

$\det(\lambda I - A) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$, $n = n_1 + \cdots + n_s$.

亦即 λ_i 为 A 的 n_i 重根. 这里 n_i 称为特征值 λ_i 的代数重数。

设 λ_i 对应的最大线性无关特征向量的个数为 m_i，即 m_i 为齐次方程 $(\lambda_i I - A)\mathbf{x} = 0$ 的基础解系中所含最多线性无关解的个数. m_i 称为特征值 λ_i 的几何重数。

由上述定义易见 $m_i \le n_i$, $i = 1, \cdots, s$.

看一个例子:

特征值的代数重数与几何重数

例 1.9

矩阵 $A=\begin{pmatrix}8&1&0\\0&8&0\\0&0&8\end{pmatrix}$ 的特征值 8 的代数重数为 3, 几何重数为 2.

它有两个线性无关特征向量 $\mathbf{v}_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$ 和 $\mathbf{v}_2=\begin{pmatrix}0\\0\\1\end{pmatrix}$.

定理 1.4

若 A 有重特征值，则 A 可对角化的充要条件是每个特征值的代数重数和几何重数都相等，即 $n_i=m_i, i=1,\cdots,s$.

Jordan标准型

在 A 的特征值几何重数与代数重数不相等情形, 有以下定理:

定理 1.5 (Jordan标准型)

设 $A \in \mathbb{C}^{n\times n}$ 有 s 个不同特征值 $\lambda_1, \cdots, \lambda_s$, 则 A 可相似为其 *Jordan* 标准型矩阵 J, 即 $\exists P \in \mathbb{C}^{n\times n}$ 可逆, *s.t.*

$$J = P^{-1}AP = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix} \text{为分块对角阵}$$

其中每个对角块 J_i 为 $n_i \times n_i$ 阶方阵, n_i 为 λ_i 的代数重数, 且

$$J_i = \begin{pmatrix} J_{i1} & & \\ & \ddots & \\ & & J_{im_i} \end{pmatrix}, \text{其中} J_{ik} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix}, k = 1, \cdots, m_i.$$

线性空间

数域的概念：设 $\mathcal{A} \subset \mathbb{C}$ 为 $\mathbb{C}$ 之子集，满足 $0 \in \mathcal{A}, 1 \in \mathcal{A}$, 且 $\mathcal{A}$ 中任意两个数的和、差、积、商 (除数不为零)仍属于 $\mathcal{A}$, 则称 $\mathcal{A}$ 为一个数域。我们常用的是实数域 $\mathbb{R}$ 和复数域 $\mathbb{C}$.

定义 1.4 (线性空间)

设 $\mathcal{A}$ 是一个数域，V 是一个非空集合，在 V 上定义了两种运算

1. 加法：
   $\forall \mathbf{u}, \mathbf{v} \in V, \exists! \mathbf{w} \in V$, *s.t.* $\mathbf{w} = \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ *(交换律).*
   而且有结合律：$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}, \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
   且 V 中有唯一的零元素 $\theta \in V$, *s.t.* $\mathbf{u} + \theta = \mathbf{u}$.
   对于 V 中每个元素 $\mathbf{u}$，存在其负元素 $-\mathbf{u}$, *s.t.* $\mathbf{u} + (-\mathbf{u}) = \theta$.

线性空间

定义 1.4 (线性空间(续))

2. 数乘:
   $\forall \alpha \in \mathcal{A}, \forall \mathbf{u} \in V, \alpha\mathbf{u}$ 仍为 V 中一元素， 且
   $1 \cdot \mathbf{u} = \mathbf{u}, \forall \mathbf{u} \in V.$
   $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}, \forall \alpha, \beta \in \mathcal{A}, \forall \mathbf{u} \in V.$
   $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}, \forall \alpha \in \mathcal{A}, \forall \mathbf{u}, \mathbf{v} \in V.$
   $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}, \forall \alpha, \beta \in \mathcal{A}, \forall \mathbf{u} \in V.$

则称 V 为数域 $\mathcal{A}$ 上的一个线性空间.

类似于 $\mathbb{R}^n, \mathbb{C}^n$ 中一样，我们也可以定义一般线性空间中的线性无关（相关）性、基向量、子空间等。

线性空间

我们来看几个常见的例子：

例 1.10 ($\mathbb{R}^n$ 和 $\mathbb{C}^n$)

$\mathbb{R}^n$是n维实向量的全体，按照向量加法及实数与向量的数乘构成了数域 $\mathbb{R}$ 上的线性空间. $\mathbb{R}^n$ 中零元素即为 $\boldsymbol{0}=(0,\cdots,0)^T$.

可以取 $\mathbb{R}^n$ 中一组基为 $\{\mathbf{e}_1,\cdots,\mathbf{e}_n\}$, 其中

$$\mathbf{e}_1=(1,0,\cdots,0)^T,\ \cdots,\mathbf{e}_n=(0,\cdots,0,1)^T.$$

$\mathbb{R}^n$ 中任意一个向量$\mathbf{x}$可以表示为这组基的线性组合，即

$$\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n,$$

其中的实数 $x_1,\cdots,x_n$ 是向量 $\mathbf{x}$ 在这组基下的坐标.

类似的，$\mathbb{C}^n$ 为 n 维复向量的全体，它为复数域 $\mathbb{C}$ 上的线性空间.

线性空间

例 1.11 ($\mathbb{R}^{n\times m}$ 和 $\mathbb{C}^{n\times m}$)

$\mathbb{R}^{n\times m}$是所有 n 行 m 列实元素矩阵的全体, 按照矩阵加法及实数与矩阵的数乘构成了数域 $\mathbb{R}$ 上的线性空间.

$\mathbb{R}^{n\times m}$ 中零元素即为 n 行 m 列元素全为零的矩阵.

$\mathbb{C}^{n\times m}$ 类似，它为复数域 $\mathbb{C}$ 上的线性空间。

线性空间

例 1.12 (连续函数空间$C[a,b]$)

$C[a,b]$是定义在区间 $[a,b]$ 上的实值连续函数全体， 按照函数加法及实数与函数的数乘构成了数域 $\mathbb{R}$ 上的线性空间.

对于 $f,g\in C[a,b], \forall\alpha\in\mathbb{R}$,

$$\begin{aligned}(f+g)(x) &= f(x)+g(x), \quad \forall x\in[a,b],\\ (\alpha f)(x) &= \alpha\cdot f(x), \quad \forall x\in[a,b].\end{aligned}$$

$C[a,b]$ 中零元素即为在 $[a,b]$ 区间上恒为零的函数.

线性空间

$C[a,b]$ 中函数 $\phi_1,\cdots,\phi_m$ 线性无关指的是如果存在实数 $c_1,c_2,\cdots,c_m$ 使得

$$c_1\phi_1(x)+\cdots+c_m\phi_m(x)=0,\quad \forall x\in[a,b],$$

那么有所有$c_1=c_2=\cdots=c_m=0.$

$C[a,b]$ 是无穷维线性空间，在里面可以找到任意多线性无关的函数。

类似地，$C^n[a,b]$是定义在区间 $[a,b]$ 上具有n阶连续导数的实值函数全体，按照函数加法及实数与函数的数乘构成了数域 $\mathbb{R}$ 上的线性空间.

线性空间

例 1.13 (多项式函数空间$\mathcal{P}_n[a,b]$)

最简单的连续函数即为多项式， 多项式函数空间$\mathcal{P}_n[a,b]$指的是定义在区间 $[a,b]$ 上的不超过 n 次的多项式全体，按上面的加法和数乘定义构成数域 $\mathbb{R}$ 上的线性空间。

$\forall p_n(x) \in \mathcal{P}_n[a,b]$, 存在实系数 $a_0, a_1, \cdots, a_n$ 使得

$$p_n(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad \forall x \in [a,b].$$

也就是说$\mathcal{P}_n[a,b]$中的一组基可以取成 $\phi_i(x) \equiv x^i, i = 0, 1, \cdots, n$, 即空间$\mathcal{P}_n[a,b]$的维数是 $n+1$, 它是 $C^n[a,b]$ 的一个 $n+1$维子空间，我们一般记作 $\mathcal{P}_n[a,b] = \mathrm{span}\{1, x, \cdots, x^n\}$.

范数的定义及其性质

先来看看一般线性空间中范数的定义.

定义 1.5 (范数、赋范空间)

设 X 为一个复（实）线性空间，一个函数 $\|\cdot\|: X \to \mathbb{R}$ 如果满足以下三个条件

1. （非负性）$\forall \mathbf{x} \in X, \|\mathbf{x}\| \geq 0$, 且 $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$;
2. （齐次性）$\forall \alpha \in \mathbb{C}(\mathbb{R}), \forall \mathbf{x} \in X, \|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;
3. （三角不等式）$\forall \mathbf{x}, \mathbf{y} \in X, \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;

则称 $\|\cdot\|$ 为 X 上的一个范数. $(X, \|\cdot\|)$ 就称为一个赋范线性空间.

范数的定义及其性质

例 1.14

对 $X=\mathbb{R}^n$ 或 $\mathbb{C}^n$, 有以下一些常用范数： $\mathbf{x}=(x_1,\cdots,x_n)^T\in X$,

$$\infty-\text{范数：}\qquad \|\mathbf{x}\|_\infty=\max_{1\le i\le n}|x_i|,$$

$$1-\text{范数：}\qquad \|\mathbf{x}\|_1=\sum_{i=1}^n|x_i|,$$

$$2-\text{范数：}\qquad \|\mathbf{x}\|_2=\left(\sum_{i=1}^n|x_i|^2\right)^{1/2},$$

更一般地，对 $p\ge 1$ 可以定义

$$p-\text{范数：}\qquad \|\mathbf{x}\|_p=\left(\sum_{i=1}^n|x_i|^p\right)^{1/p}.$$

范数的定义及其性质

定理 1.6 (范数的连续性)

设 $A \in \mathbb{R}^{n\times n}(\mathbb{C}^{n\times n})$, 对 $\mathbb{R}^n(\mathbb{C}^n)$ 上任意一种向量范数 $\|\cdot\|$, $\mathbf{x} = (x_1, \cdots, x_n)^T \in \mathbb{R}^n(\mathbb{C}^n)$, 有 $\|A\mathbf{x}\|$ 为 $x_1, \cdots, x_n$ 的连续函数.

$\lhd$ 只需证明 $\forall \varepsilon > 0, \exists \delta > 0$, 当 $\mathbf{h} = (h_1, \cdots, h_n)^T$, $|h_j| < \delta$ 时 $(1 \le j \le n)$, 有 $\big|\|A(\mathbf{x}+\mathbf{h})\| - \|A\mathbf{x}\|\big| < \varepsilon$.

将 A 写成列向量形式 $A = (\vec{a}_1, \cdots, \vec{a}_n)$, 有

$$\big|\|A(\mathbf{x}+\mathbf{h})\| - \|A\mathbf{x}\|\big| \overset{\text{三角不等式}}{\le} \|A\mathbf{h}\| = \left\|\sum_{j=1}^n h_j \vec{a}_j\right\|$$

$$\overset{\text{三角不等式}}{\le} \sum_{j=1}^n |h_j| \|\vec{a}_j\| \le M \max_{1\le j\le n} |h_j|, \quad \text{其中 } M = \sum_{j=1}^n \|\vec{a}_j\|.$$

这样取 $\delta = \varepsilon/M$, 那么 $|h_j| < \delta$ 时, 有 $\big|\|A(\mathbf{x}+\mathbf{h})\| - \|A\mathbf{x}\|\big| < \varepsilon$. $\rhd$

范数的定义及其性质

推论 1.1

$\|\mathbf{x}\|$ 是 $x_1, \cdots, x_n$ 的连续函数.

定义 1.6 (范数等价性)

称两个范数$\|\cdot\|_a$ 与 $\|\cdot\|_b$等价, 是指存在两个正常数 C_1 和 C_2, s.t.

$\forall \mathbf{x} \in X, \quad C_1\|\mathbf{x}\|_a \le \|\mathbf{x}\|_b \le C_2\|\mathbf{x}\|_a.$

定理 1.7 (有限维赋范空间范数等价性定理)

任意有限维线性空间上的范数均等价.

$\lhd$ 我们只需证明任意有限维线性空间上的范数与一个特定范数等价即可，然后利用等价的传递性即得定理结论.

设 X 的维数为 n, 取一组基为 $\mathbf{e}_1, \cdots, \mathbf{e}_n$. 即 $\forall \mathbf{x} \in X, \mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i$.

范数的定义及其性质

欲证明 $\exists C_1, C_2 > 0$, s.t. $\quad C_1\|\mathbf{x}\|_\infty \le \|\mathbf{x}\| \le C_2\|\mathbf{x}\|_\infty$,
其中 $\|\mathbf{x}\|_\infty = \max\limits_{1\le i\le n} |\alpha_i|$. 易验证这是一种向量范数.

$$\|\mathbf{x}\| = \left\|\sum_{i=1}^n \alpha_i \mathbf{e}_i\right\| \overset{\text{三角不等式}}{\le} \sum_{i=1}^n |\alpha_i|\|\mathbf{e}_i\| \le C_2\|\mathbf{x}\|_\infty.$$

其中 $C_2 = \sum\limits_{i=1}^n \|\mathbf{e}_i\|$. 这样证明了右半部份.

下面用反证法来证明存在 $C_1 > 0$ s.t. $\forall \mathbf{x} \in X$, $\|\mathbf{x}\|_\infty \le \frac{1}{C_1}\|\mathbf{x}\|$.
假设上述 C_1 不存在, 即 $\forall m \in \mathbb{N}$, $\exists \mathbf{x}_m \in X$, s.t. $\|\mathbf{x}_m\|_\infty \ge m\|\mathbf{x}_m\|$,
这里我们无妨设 $\|\mathbf{x}_m\| = 1$.

范数的定义及其性质

令 $\mathbf{y}_m = \frac{\mathbf{x}_m}{\|\mathbf{x}_m\|_\infty}$, 这样就有 $\|\mathbf{y}_m\|_\infty = 1$. 且由 $\|\mathbf{x}_m\|_\infty \geq m\|\mathbf{x}_m\|$, 知 $\|\mathbf{y}_m\| = \frac{\|\mathbf{x}_m\|}{\|\mathbf{x}_m\|_\infty} \leq \frac{1}{m}$.

设 $\mathbf{y}_m = \sum_{i=1}^{n} \alpha_{m,i}\mathbf{e}_i$, 那么由 $\|\mathbf{y}_m\|_\infty = 1$ 知 $\max_{1\leq i\leq n} |\alpha_{m,i}| = 1$.

这样序列 $\{\alpha_{m,i}\}_{i=1,m=1}^{n,\infty}$ 的任意子序列都是有界的.

因此 $\forall i = 1, \cdots, n$, 上述序列存在（收敛）子序列 $\{\alpha_{m_l,i}\}_{l=1}^{\infty}$, s.t. $\alpha_{m_l,i} \longrightarrow \beta_i$, 当 $l \to \infty$.

即，令 $\mathbf{y} = \sum_{i=1}^{n} \beta_i\mathbf{e}_i$, 就有 $\|\mathbf{y}_{m_l} - \mathbf{y}\|_\infty \to 0$, 当 $l \to \infty$.

即在 $\|\cdot\|_\infty$ 范数意义下，$\left\{\mathbf{y}_{m_l} = \sum_{i=1}^{n} \alpha_{m_l,i}\mathbf{e}_i\right\}_{l=1}^{\infty}$ 收敛到点 $\mathbf{y}$.

范数的定义及其性质

前面已证 $\forall \mathbf{x} \in X, \|\mathbf{x}\| \le C_2 \|\mathbf{x}\|_\infty$, 自然

$\|\mathbf{y}_{m_l} - \mathbf{y}\| \le C_2 \|\mathbf{y}_{m_l} - \mathbf{y}\|_\infty \to 0,$

另一方面，$\|\mathbf{y}_{m_l}\| \le \frac{1}{m_l} \to 0$, 当 $l \to +\infty$.

因此必有 $\|\mathbf{y}\| = 0$, 即 $\mathbf{y} = \mathbf{0}$.

另外由 $\|\mathbf{y}_{m_l} - \mathbf{y}\|_\infty \to 0$, 立刻有 $\|\mathbf{y}_{m_l}\|_\infty \to 0$, 当 $l \to +\infty$.

这显然与前面 $\|\mathbf{y}_{m_l}\|_\infty = 1$ 矛盾。

即我们用反证法证明了：应该存在 $C_1 > 0$, s.t. $C_1 \|\mathbf{x}\|_\infty \le \|\mathbf{x}\|$.

这样完成了定理的证明. $\triangleright$

几种常见矩阵

- 对称、正定: 若 $A \in \mathbb{R}^{n\times n}$ 满足 $A^T = A$，则称 A 为对称阵。
  若 $\forall 0 \neq \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T A\mathbf{x} > 0$ 则称 A 为正定阵。
  若 $\forall \mathbf{x}, \mathbf{x}^T A\mathbf{x} \geq 0$ 则称 A 为半正定(非负定)阵。

  对于实对称矩阵 A，我们有

  1. A 的特征值均为实数，且 A 有n个线性无关的特征向量;
  2. A 对应不同特征值的特征向量相互正交，A 的特征向量可以构成 $\mathbb{R}^n$ 的一组标准正交基;
  3. 利用前面（关于矩阵对角化）的结果知，存在正交阵 P, s.t. $P^{-1}AP = \Lambda$ 为对角阵.

几种常见矩阵

- Hermite 阵：若 $H \in \mathbb{C}^{n\times n}$ 满足 $\bar{H}^T = H$，称 H 为Hermite阵.
  以后记 $H^* \equiv \bar{H}^T$.
- 正交(酉)矩阵：若 $A \in \mathbb{R}^{n\times n}$ 满足 $A^T A = I$, 则称 A 为正交阵;
  若 $U \in \mathbb{C}^{n\times n}$ 满足 $\bar{U}^T U = I$, 则称 U 为酉矩阵.

  正交阵(酉矩阵)有以下性质：
  1. $A^{-1} = A^T(A^*)$, 且 A^T (A^*) 也是正交阵 (酉矩阵);
  2. $|\det A| = 1$;
  3. 若 $A, B \in \mathbb{R}^{n\times n}$ 为正交阵，则 AB, BA 均为正交阵;
  4. 若 $A, B \in \mathbb{C}^{n\times n}$ 为酉矩阵，则 AB, BA 均为酉矩阵。

几种常见矩阵

- 初等矩阵：对 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $0 \neq \sigma \in \mathbb{R}$, 称
  $E(\mathbf{u}, \mathbf{v}; \sigma) = I - \sigma \mathbf{u}\mathbf{v}^T$ 为实初等矩阵.
  如 $I_{ij} = I - (\mathbf{e}_i - \mathbf{e}_j) \cdot (\mathbf{e}_i - \mathbf{e}_j)^T$ 为初等排列阵.
  这里 $\mathbf{e}_i = (\cdots, 0, 1, 0, \cdots)^T$ (仅第 i 个元素为1).
  相当于将 I 的第i行与第j行对换，得到 I_{ij}.
  且 $\det I_{ij} = -1, \quad I_{ij}^{-1} = I_{ij}$.
  若干个初等排列阵之积 P 称为排列阵.
  它每行每列恰有一个非零元素，且为1.
  PA相当于把A的各行进行了一个新的排列。

几种常见矩阵

- 对角占优：设 $A=(a_{ij})\in\mathbb{C}^{n\times n}$, 如果 $|a_{ii}|>\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|$, 对 $i=1,\cdots,n$ 都成立, 称 A 为严格对角占优;
  如果 $|a_{ii}|\geq\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|,\ i=1,\cdots,n,$
  且至少有一个为严格不等号，称 A 为弱对角占优.

定理 1.8 (严格对角占优矩阵非奇异)

若 $A\in\mathbb{C}^{n\times n}$ 严格对角占优, 则 $a_{ii}\neq 0$, $i=1,\cdots,n$. 且 A 非奇异.

严格对角占优矩阵

◁ 如果 A 严格对角占优，即

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}| \geq 0, \quad \text{对 } i = 1, \cdots, n \text{ 都成立} \Longrightarrow a_{ii} \neq 0.$$

下面证 $\det A \neq 0$. 用反证法，假设 $\det A = 0$，即存在非零向量 $\mathbf{x} = (x_1, \cdots, x_n)^T \in \mathbb{C}^n$ 使得 $A\mathbf{x} = 0$.

无妨设 $|x_k| = \max_{1\leq i\leq n} |x_i| = 1$, $A\mathbf{x} = 0$ 的第 k 个方程即为

$$a_{kk}x_k = -\sum_{j=1, j\neq i}^{n} a_{kj}x_j \Longrightarrow |a_{kk}| \leq \sum_{j=1, j\neq i}^{n} |a_{kj}||x_j| \leq \sum_{j=1, j\neq i}^{n} |a_{kj}|,$$

与 A 严格对角占优矛盾. 故必有 $\det A \neq 0$. ▷

（不）可约矩阵

- 可约矩阵：设 $A \in \mathbb{C}^{n\times n}$. 如果存在排列阵 P 使得
  $$P^TAP = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix},$$ 其中 $\bar{A}_{11}, \bar{A}_{22}$
  分别为 $r \times r$ 和 $(n-r) \times (n-r)$ 阶方阵,
  则称 A 为可约矩阵。
  否则就称 A 为不可约矩阵.

不可约矩阵

如果 A 是可约矩阵，则可经过行与列的重排，使 P^TAP 具有 $P^TAP = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix}$ 的形式. 即求解方程组 $A\mathbf{x} = \mathbf{b}$ 可变换为

$$P^TAP\mathbf{y} = P^T\mathbf{b}, \quad \mathbf{y} = P^T\mathbf{x}.$$

利用上面分块形式可得

$$\begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$$

这样可以先解 $\bar{A}_{22}\mathbf{y}_2 = \mathbf{b}_2$, 再解 $\bar{A}_{11}\mathbf{y}_1 = \mathbf{b}_1 - \bar{A}_{12}\mathbf{y}_2$.
显然这样可以节约计算量。

不可约矩阵

怎样考察矩阵是否可约呢？ 可以利用矩阵的图来分析.

设 $A=[a_{ij}]\in\mathbb{C}^{n\times n}$, 在平面上作出 n 个不同的点 $P_1,\cdots,p_n$, 称为节点. 对于 A 的每个非零元素 a_{ij}, 建立一条有向路径 $\overset{\frown}{P_iP_j}$. 如果 $a_{ii}\neq 0$, 也作出一条从节点 P_i 指向自己的有向路径. 这样每个 $n\times n$ 的矩阵 A 都对应一个方向图 $G(A)$.

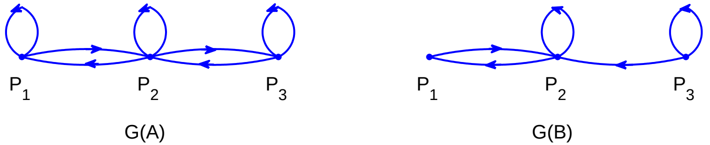
$G(A)$ 是一个有节点 $P_1,\cdots,P_n$ 的图， 当且仅当 $a_{ij}\neq 0$ 时， 节点 P_i 到 P_j 有一条有向路径.

看一个例子:

不可约矩阵

例 1.15

设 $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 1 & 2 \end{pmatrix}$, 则它们的方向图如下所示:



不可约矩阵

定义 1.7 (强连接)

设 $A \in \mathbb{C}^{n\times n}$, 在 A 的方向图 $G(A)$ 中，如果对于 $P_1, \cdots, P_n$ 中任意有序的一对节点 P_i 和 P_j，都存在连接 P_i 到 P_j 的有向路径: $\overset{\frown}{P_i P_k}, \overset{\frown}{P_k P_l}, \cdots, \overset{\frown}{P_m P_j},$, 则称方向图 $G(A)$ 是强连接的.

例 1.15 中 $G(A)$ 是强连接的，但是 $G(B)$ 不是强连接的, 因为 $G(B)$ 中例如 P_2 到 P_3 就没有有向路径.

定理 1.9

$A \in \mathbb{C}^{n\times n}$ 是不可约矩阵的充要条件是方向图 $G(A)$ 是强连接的.

即例 1.15 中 A 是不可约的，但 B 是可约的.

目录

极值点与费马定理

定义 1.8 (极值点)

设函数 $f(x)$ 在点 x_0 的某个邻域中有定义，如果存在其一个邻域 $N(x_0)$, 使得对所有的 $x \in N(x_0)$, 都有

$$f(x) \leq f(x_0), \text{ 或者 } f(x) \geq f(x_0),$$

则称函数 $f(x)$ 在点 x_0 取得极大值（或者极小值），x_0 称为 $f(x)$ 的极大指点（极小值点）. 极大指点和极小值点统称为极值点。

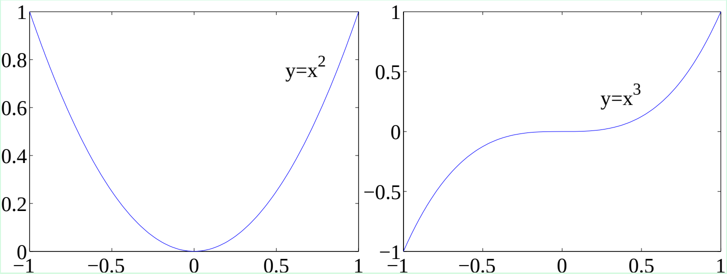
关于函数极值，有以下必要条件:

定理 1.10 (费马(Fermat)定理)

设函数 $f(x)$ 在 x_0 取得极值, 如果 $f'(x_0)$ 存在, 则必有 $f'(x_0) = 0$.

费马定理（续）

费马定理说明，在可导的前提下，导数为零是极值点的必要条件。当然反过来不一定成立，即导数为零的点不一定是极值点。 举例来说，函数 $f(x)=x^2$ 和 $g(x)=x^3$ 在 $x=0$ 点的导数都是零，显然 $x=0$ 是 $f(x)=x^2$ 的极小值点, 但不是 $g(x)=x^3$ 的极值点. 我们把导数为零的点称为函数的驻点或者临界点.



微分中值定理

定理 1.11 (罗尔(Roll)定理)

设函数 $y = f(x)$ 满足条件：

1. 在闭区间 $[a, b]$ 上连续；
2. 在开区间 (a, b) 上可导；
3. 在区间端点处函数值相等，即 $f(a) = f(b)$.

则存在点 $\xi \in (a, b)$ 使得 $f'(\xi) = 0$.

定理 1.12 (拉格朗日(Lagrange)定理)

设函数 $y = f(x)$ 满足条件：

1. 在闭区间 $[a, b]$ 上连续；
2. 在开区间 (a, b) 上可导.

则存在点 $\xi \in (a, b)$ 使得 $f'(\xi) = \dfrac{f(b) - f(a)}{b - a}$.

函数极值点的充分条件

定理 1.13 (极值点第一充分条件)

设函数 $f(x)$ 在点 x_0 的某个邻域内有一阶导数，并且 $f'(x)$ 在 x_0 两侧有不同的符号， 则 $f(x)$ 在点 x_0 取得极值. 即

1. 如果存在 $\delta > 0$, 使得在 $(x_0 - \delta, x_0)$ 内 $f'(x) > 0$, 在 $(x_0, x_0 + \delta)$ 内 $f'(x) < 0$, 则 $f(x)$ 在 x_0 取极大值;
2. 如果存在 $\delta > 0$, 使得在 $(x_0 - \delta, x_0)$ 内 $f'(x) < 0$, 在 $(x_0, x_0 + \delta)$ 内 $f'(x) > 0$, 则 $f(x)$ 在 x_0 取极小值。

函数极值点的充分条件

定理 1.14 (极值点第二充分条件)

设函数 $f(x)$ 在点 x_0 的某个邻域内有一阶导数，并且 $f'(x_0)=0$. 又设 $f''(x_0)$ 存在，那么

1. 如果 $f''(x_0)<0$, 则 $f(x)$ 在 x_0 取极大值；
2. 如果 $f''(x_0)>0$, 则 $f(x)$ 在 x_0 取极小值。

例 1.16 (例如求 $f(x)=x^3-x$ 的极值点)

解: 先求驻点 $f'(x)=0 \Longrightarrow 3x^2-1=0 \Longrightarrow x_{1,2}=\pm\frac{\sqrt{3}}{3}$. 显然

x	$(-\infty,-\frac{\sqrt{3}}{3})$	$-\frac{\sqrt{3}}{3}$	$(-\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3})$	$\frac{\sqrt{3}}{3}$	$(\frac{\sqrt{3}}{3},+\infty)$
$f'(x)$	$+$	0	$-$	0	$+$

可见 $x_1=-\frac{\sqrt{3}}{3}$ 是极大值点, $x_2=\frac{\sqrt{3}}{3}$ 是极小值点. 求二阶导数 $f''(x_1)<0$, $f''(x_2)>0$, 得到与前面一样的结论.

函数的凸性

定义 1.9 (凸函数与凹函数)

如果 $\forall x_1, x_2 \in [a,b]$ 及 $\forall \lambda \in [0,1]$, 对于函数 $f(x)$ 都有

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

成立，那么称 $f(x)$ 在 $[a,b]$ 上为**凸函数(convex function)**.

反之, 如果 $\forall x_1, x_2 \in [a,b]$ 及 $\forall \lambda \in [0,1]$, 对于函数 $f(x)$ 都有

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2)$$

成立，那么称 $f(x)$ 在 $[a,b]$ 上为**凹函数(concave function)**.

显然，如果 $f(x)$ 可微, 那么 f 凸 $\Longleftrightarrow$ $f'(x)$ 单调增.

如果 $f(x)$ 二阶可微, 那么 f 凸 $\Longleftrightarrow$ $f''(x) \geq 0$.

泰勒(Taylor)公式

反复运用微分中值定理，可以证明以下定理

定理 1.15 (带有拉格朗日余项的泰勒公式)

设函数 $f(x)$ 在 x_0 的某个邻域 $N(x_0)$ 内存在 $n+1$ 阶导数, 那么 $\forall x \in N(x_0)$, 有

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1},$$

其中 ξ 介于 x 与 x_0 之间.

积分中值定理与分部积分公式

定理 1.16 (积分中值定理)

设 $f(x)$ 在 $[a,b]$ 上连续, $g(x)$ 在 $[a,b]$ 上可积且不变号，那么一定存在 $\xi \in (a,b)$ 使得 $\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx$.

特别当 $g(x) \equiv 1$ 时, 有 $\xi \in (a,b)$ 使得 $\int_a^b f(x)dx = f(\xi)(b-a)$.

定理 1.17 (分部积分公式)

设 $f(x)$ 在 $[a,b]$ 上连续, $g(x)$ 在 $[a,b]$ 有连续导数，那么对 $f(x)$ 在 $[a,b]$ 上的任一原函数 $F(x)$ 有

$$\int_a^b f(x)g(x)dx = F(x)g(x)\Big|_{x=a}^{x=b} - \int_a^b F(x)g'(x)dx.$$