

数 学

集 合 论

【集合论】数学的一个重要分支，现代数学的基础。集合论是康托尔在 19 世纪末创立的。它经历了两个发展阶段：1908 年以前称为朴素集合论；1908 年以后称为公理集合论。

朴素集合论产生的背景是分析学，康托尔在研究“函数展开为三角级数的唯一性”问题时，首先提出了集合的概念，并且提出了对无穷集合分类的一种法则，即“一一对应”法则。如果两个集合的元素之间能建立一一对应的关系，则称这两个集合等势。一个无穷集合可与它的一个真子集等势，这与传统的观念“全量大于部分”矛盾，但康托尔认为这恰恰是无穷集合区别于有穷集合的特征。他称与全体自然数集合 $\mathbb{N}$ 等势的集合为可数集，当他证明了全体有理数集和全体代数数集都是可数集之后，又发现全体实数集 $\mathbb{R}$ 不是可数集，这说明无穷集合之间存在着差异，为了刻划这种差异，他引进了基数的概念，同时他还开始研究有序的集合，特别是良序集的结构，并引入了序数的概念，将数学归纳

法推广到自然数以外，提出了超限归纳法。

1900 年左右集合论中出现了悖论，其中最著名的是罗素悖论。悖论的出现表明集合论的理论是不协调的，也使人们对整个数学推理的正确性与结论的真理性产生怀疑，酿成了数学史上的第三次危机。

为了克服悖论所带来的困难，策梅洛于 1908 年提出了一个公理化方案，后经弗伦克尔和斯科朗的改进，形成了现代的公理集合论。

【集合】数学中最基本的概念，也是集合论研究的对象。我们把一定范围内、确定的、可区别的事物的整体称为集合，组成集合的事物叫做集合的元素，如果 a 是集合 A 的元素，则称 a 属于 A 或 A 包含 a ，记作 $a \in A$ 。如果一个集合只含有限个元素，则称为有限集，如果一个集合包含无限个元素，则称为无限集。

当一个集合的所有元素都已确定时，这个集合就被确定了。有限集合可将其元素全部列出。置于括弧之内来表示。例如全体小写英文字母的集合可表示为 $\{a, b, \dots, z\}$ ，对无限集合上述方法不行，这时只有用可以刻画所有元素 X 的某一性质 $P(X)$ 来加以概括，例如全体自然数的集合可表示为 $\{X | X \text{ 是自然数}\}$ 。

集合可以没有任何元素，这种集合只有一个，称为空集，记作 $\emptyset$ 。如果集合 B 的元素都是 A 的元素，就称 B 是 A 的子集，也称 A 包含 B 。记作 $B \subseteq A$ 或 $A \supseteq B$ 。空集是任何集合的子集，任何集合都是它自己的子集。如果集合 A, B 有相同的元素，那么这两个集合相等，记作 $A = B$ 。

以属于 A 或属于 B 的元素为元素的集合称为 A 与 B 的并集，记作 $A \cup B$ ，以属于 A 又属于 B 的元素为元素的集合称为 A

与 B 的交集, 记作 $A \cap B$, 以属于 A 但不属于 B 的元素为元素的集合称为 A 与 B 的差集, 记作 $A - B$ 。当 $B \subseteq A$ 时, $A - B$ 可称为 B 关于 A 的补集, 集合的并与交都适合交换律和结合律, 即 $A \cup B = B \cup A$, $(A \cup B) \cup C = A \cup (B \cup C)$, $A \cap B = B \cap A$, $A \cap (B \cap C) = (A \cap B) \cap C$ 。并与交还适合分配律, 即 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, 它们与差一起适合德·摩根律: $S - (A \cup B) = (S - A) \cap (S - B)$, $S - (A \cap B) = (S - A) \cup (S - B)$ 。如果对某个非空集合 I 的每个元素 $i \in I$ 都指定一个集合 A_i , 那么由 A_i 组成的集合 $\{A_i: i \in I\}$ 称为集合族。集合族 $\{A_i: i \in I\}$ 的交定义为 $\{X | \forall i \in I X \in A_i\}$, 记作 $\bigcap_{i \in I} A_i$, 集合族 $\{A_i: i \in I\}$ 的并定为 $\{X | \exists i \in I X \in A_i\}$, 记作 $\bigcup_{i \in I} A_i$ 。集合 A 的全部子集组成的集合称为 A 的幂集, 记作 $P(A)$, 对任意两个元素 a, b 按一定次序排列起来可构成有序对集合, 记作 $\langle a, b \rangle$ 。设 A, B 是两个非空集合, 从 A, B 中各取一个元素 a, b 所组成的有序对 $\langle a, b \rangle$ 的全体组成集合 $\{\langle a, b \rangle | a \in A, b \in B\}$ 称为 A 与 B 的直积, 记作 $A \times B$ 。

【映射】数学的基本概念。设 A 和 B 是两个非空集合, 如果按照某一法则将 A 中任一元素与 B 中唯一元素相对应, 就称该法则为从 A 到 B 的映射, 记作 $f: A \rightarrow B$, A 称为 f 的定义域, 记作 $dom(f)$, A 中元素 X 所对应的 B 中唯一元素 y 称为 X 在 f 下的象, 记作 $f(X)$, $f(A) = \{f(X) | X \in A\}$ 是 B 的子集, 称为 f 的值域, 记作 $ran(f)$ 。两个映射 f 和 g 只有当它们有相同的定义域, 并且对任意 $X \in A$ 有 $f(X) = g(X)$ 时才称为相等。当 $ran(f) = B$ 时称 f 是 A 到 B 上的映射, 也称满射。如果 f 将 A 中不同元素映到 B 中

不同的象,则称 f 是单射。如果 f 既是满射又是单射,则称 f 是双射,也称一一映射。对于双射 $f: A \rightarrow B$,使 $f(X)$ 对应于 X 的映射称为 f 的逆映射,记作 $f^{-1}: B \rightarrow A$ 。显然有 $(f^{-1})^{-1} = f$,对于映射 $f: A \rightarrow B, g: B \rightarrow C$,由 $h(X) = g(f(X))$ 所确定的映射 $h: A \rightarrow C$ 称为 f 与 g 的复合映射,记作 $h = g \circ f$ 。复合映射满足结合律,即 $h \circ (g \circ f) = (h \circ g) \circ f$ 。

当 $X \in A$ 时所有有序对 $\{\langle X, f(X) \rangle; X \in A\}$ 组成的直积 $A \times B$ 的子集称为 f 的图像。这个子集可以看作是 A 与 B 之间的关系,每个映射是一特殊的关系。

【关系】数学的基本概念。由有序对 $\langle X, y \rangle$ 组成的非空集合 $R \subseteq A \times B$,称为 A 与 B 之间的二元关系。如果 $\langle X, y \rangle \in R$ 称 X 与 y 有关系 R ,记作 XRy , R 中有序对的第一坐标组成的集合称为 R 的定义域。第二坐标组成的集合称为 R 的值域,如果 $A=B$,则 R 是 A 上的二元关系,一般地,如果 $R \subseteq A^n$,则称 R 是 A 上的 n 元关系。

A 上二元关系 R 满足: $\forall X \in A, XRX$ 称 R 具有自反性, R 满足 $XRy, yRz \Rightarrow XRz$,称 R 具有传递性。 R 满足 $XRy \Rightarrow yRX$,称 R 具有对称性,如果 R 满足 $XRy, yRX \Rightarrow X=y$ 则称 R 满足反对称性。如果 A 上的二元关系 R 满足自反性、对称性、传递性,则称 R 是 A 上的等价关系。如果 A 上的二元关系 R 满足自反性、反对称性、传递性则称 R 是 A 上的偏序关系,偏序关系 R 通常记作 $\leq$ 。集合 A 连同偏序关系 $\leq$ 称为偏序集,记作 $(A, \leq)$,如果偏序关系 $\leq$ 还满足 $\forall X, y \in A, X \leq y$ 或 $y \leq X$,则称 $\leq$ 是 A 上的线序, $(A, \leq)$ 称为线序集,如果对线序集 $(A, \leq)$ 的每个非空子集 B 都存在 $X_0 \in B$ 满足 $\forall X \in B, X_0 \leq X$,则称 $\leq$ 是 A 上的良序,

$(A, \leq)$ 称为良序集。

【序数】集合论的基本概念。是日常使用的第一、第二等表示次序的数的推广。如果两个偏序集 $(A, \leq)$ 和 $(B, \leq)$ 是序同构的, 记作 $(A, \leq) \cong (B, \leq)$, 其中 $\cong$ 是一个等价关系。在所有偏序集中可以按这个等价关系进行分类, 每个等价类称为序型。序数原来定义为良序集的序型, 即与良序集 $(A, \leq)$ 序同构的所有良序集组成的类。但是在 ZFC 系统中无法证明序型是一个集合, 因此原来的定义必须修改。1923 年, 冯·诺伊曼给出序数的一个新定义。如果线序集 $(\alpha, \in)$ 满足 $\forall X \in \alpha, \forall y \in X \Rightarrow y \in \alpha$, 则称 $(\alpha, \in)$ 是传递集。如果 $(\alpha, \in)$ 是传递的良序集, 则称 α 是序数, 可以证明每个良序集都有唯一的序数与之序同构, 因此每个序数可以看作良序集的序型的代表元。

在 ZFC 系统中, 自然数 0 定义为 $\emptyset$, 1 定义为 $0 \cup \{0\}$, $n+1$ 定义为 $n \cup \{n\}$, $\dots$ 所有自然数的集合 $\omega = \{0, 1, \dots\}$ 也是序数。序数有三种, 第一种是 0; 第二种是某一序数 α 的后继 $\alpha^1 = \alpha \cup \{\alpha\}$, 称为后继序数; 其它序数属于第三种, 称为极限序数。 ω 就是第一个极限序数, $\omega+1 = \omega \cup \{\omega\}$ 是 ω 的后继序数。

【基数】又称势。集合论的基本概念。是日常用以表示多少的数的概念的推广和发展。如果集合 A 与 B 的元素之间存在一一对应的关系, 则称 A 与 B 等势, 或称 A 与 B 有相同的基数, 记作 $A \sim B$ 。集合 A 的基数原来定义为 $\overline{A} = \{B | B \sim A\}$, 即一切与 A 等势的集合的类, 但是在 ZFC 系统中无法证明 $\overline{A}$ 是集合。1928 年冯·诺伊曼给出了新的定义, 根据选择公理, 每个集合 A 都可以良序化, 因此有一个序数 α 使得 $A \sim \alpha$, 冯·诺伊曼把这种 α 中最小的那个序

数定义为集合 A 的基数, 记作 $|A|$ 。这样定义的基数仍是集合, 而且如果 $A \sim B$, 则 $|A| = |B|$ 因此基数是序数的一部分, 每个基数都是不能与小于自己的序数等势的序数, 也称为初始序数。例如 $0, 1, 2, \dots$ 等自然数都是初始序数, ω 是初始序数, 但 $\omega+1$ 不是初始序数。所有势为自然数的集合称为有限集合, 与自然数集 ω 等势的集合称为可数集合, ω 作为基数常用 $\aleph_0$ 表示, 它是第一个无限基数, 非可数的无限集称为不可数集, 实数集是不可数集, 第一个不可数的基数常用 $\aleph_1$ 表示。

【超限归纳法】集合论的重要原理。是证明集合论中某种类型命题的重要方法, 超限归纳法可以叙述为: 设 φ 是关于全体序数的命题, 如果 (1) $\varphi(0)$ 成立; (2) 若 $\varphi(\alpha)$ 成立可推出 $\varphi(\alpha+1)$ 成立; (3) 若 α 为极限序数, 并且对一切 $\beta < \alpha$, $\varphi(\beta)$ 成立, 可推出 $\varphi(\alpha)$ 成立, 那么 φ 对于所有的序数都成立。

超限归纳法不仅是对于全体序数类成立, 还可以推广到一般的良序集, 设 $(X, \leq)$ 是一良序集, 对任意 $a \in X$, $X_a = \{b \in X \mid b < a\}$ 称为在 X 中由 a 所确定的前截段。对于 $E \subseteq X$, 如果 $\forall a \in X$ 只要 $X_a \subseteq E$ 就有 $a \in E$, 那么称 E 是 X 的归纳子集, 超限归纳法断言: 如果 E 为良序集 $(X, \leq)$ 的归纳子集, 则 $X = E$, 因为若 a 为 X 的最小元, 则 $X_a = \emptyset \subseteq E$, 所以 $a \in E$, 因此 E 非空。如果 E 是 X 的真子集, a 是 $X - E$ 的最小元, 那么 $X_a \subseteq E$, 因此 $a \in E$ 矛盾, 所以 $E = X$, 当 X 为自然数集 N 时, 即得到了上述定理的一个特殊情况, 即普通的数学归纳法。因此超限归纳法是数学归纳法的推广。

【悖论】自相矛盾的命题, 即如果承认这个命题, 就可以用逻辑推

理的方法推出它的否定。反之，如果承认它的否命题，又可以推出这个命题。在集合论中著名的悖论有罗素悖论、康托尔悖论和布拉里-福尔蒂悖论。

罗素悖论：设 S 为一切不属于自身的集合所组成的集合，问 S 是否属于 S ，如果 S 属于 S ，则 S 是 S 的元素，因此由 S 的定义，可知 S 不属于 S ，反之若 S 不属于 S ，则 S 是不属于自身的集合，由 S 的定义可知 S 属于 S ，无论如何都将导致矛盾，为了解释罗素悖论，他又给出了一个通俗的例子，即理发师悖论。据说在一个小岛上只有一个理发师，他声称只给那些不为自己理发的人理发，但是不给那些为自己理发的人理发，于是有人问他，是否给自己理发？按照他的声明，如果他给自己理发，他就不能为自己理发，如果他不给自己理发，那么他就应该给自己理发，这位理发师陷入了逻辑的窘境中。

数 理 逻 辑

【数理逻辑】数学的一个分支学科。用数学方法研究的逻辑。通常人们认为数理逻辑包含 4 个领域：(1) 公理集合论；(2) 模型论；(3) 递归论；(4) 构造主义和证明论。此外还有两个不太确定的领域，可以称为：(a) 逻辑和计算机；(b) 数学基础。纵观数理逻辑的发展史，它和哲学有多方面的相互作用，包括构造一些其它的逻辑系统，如模态逻辑、时态逻辑、道义逻辑、量子逻辑、多值逻辑和模糊逻辑等。

数理逻辑的先驱者是莱布尼茨和布尔，他们首先提出了用数学方法研究逻辑系统的思想。18 世纪末，弗雷格和康托尔的工作为现代数理逻辑奠定了基础。1879 年弗雷格出版了一部专著《概念演算》，第一次规定了一阶谓词演算的现代形式，几乎同时，康托尔创建了集合论，开始用一种严格的数学方法来处理无穷集合。他在 1874 年提出了著名的连续统假设 (CH)，这个假设猜想：实数的个数是自然数个数之后的第一个无穷数。虽然 CH 至今还没有解决，但是为解决它所提出的各种方法极大地推动了集合论的发展。1900 年左右集合论中出现了一些悖论，最著名的有罗素悖论、康托尔悖论、布拉里-福尔蒂悖论。特别是罗素悖论在当时的数学界和逻辑界引起了极大震动，它动摇了传统的数学概念、数学证明和数学方法的可信性标准，引起了第三次数学危机。

为了摆脱危机，希尔伯特提出了一个计划，即将所有古典数学理论公理化，并将它们的协调性归结到初等数论的协调性中，然后用有穷方法证明数论的协调性，希尔伯特计划后来发展为证明论。

与此同时关于集合论的公理化进程也在进行中，在 1908 年策梅罗提出了集合论的公理化方案，经过一系列改进和完善，最终形成了 ZFC 公理系统。

从 30 年代开始哥德尔发表了一系列重要的结果，开辟了数理逻辑的新纪元。1931 年他发表了著名的不完备性定理，证明了数论以及强于数论的任何形式系统都是不完备的。这个定理给希尔伯特计划以致命的打击。这个定理也是证明论中最重要的定理。

此后不久，哥德尔在埃尔布朗的启发下提出了给予算法以精确定义的方法，克林进一步把它具体化，提出了递归函数的概念。1936 年阿兰·图灵引进了著名的图灵机来描述算法，继而证明了

图灵可计算函数和递归函数的等价性,从而开始了递归论的研究。

哥德尔在 1930 年还证明了一阶谓词演算的完备性,即每个“在一切可能模型中为真”的语句在一阶逻辑中是可证的,或者说每个不可否证的语句都有模型。哥德尔的这个定理在模型论的发展中起了重要的作用,在一阶逻辑中可以建立许多数学公理系统,模型论主要研究这些公理系统的模型的一般性质,虽然在 20 年代模型论已有了一些零碎的结果,但对模型论的系统研究开始于 50 年代,模型论的最重要的方法是紧致性定理以及超积方法。这些方法在许多数学领域都有广泛的应用。

【公理集合论】数理逻辑的一个分支。现代数学的基础。19 世纪末,康托尔创立的朴素集合论直观地阐述了集合、序数、基数等概念。这些概念在当时都含有某种未被澄清的含糊性,从而出现了若干悖论。为了避免悖论,策梅洛于 1908 年提出了第一个集合论公理系统,在 20 年代弗伦克尔和斯科朗曾予以改进和补充,从而形成了策梅洛-弗伦克尔公理系统,简记为 ZF 。 ZF 是一个形式系统,它以“集合”和“属于”为仅有的不加定义的原始概念,它的公理有:外延公理、空集存在公理、无序对公理、并集公理、幂集公理、无穷公理、分离公理模式、替换公理模式、正则公理。如果另加选择公理,则所得到的公理系统简记为 ZFC 。

在 ZF 中,几乎所有的数学概念都能用集论语言表达。数学定理也大多都可以在 ZFC 系统内得到形式证明,因而作为整个数学的基础(至多范畴论例外), ZFC 是完备的。数学的协调性可以归结成 ZFC 的协调性。但是根据哥德尔不完备性定理,在 ZFC 中不能证明 ZFC 的协调性。所以在 ZFC 中只考虑相对协调性问题。

公理集合论中最著名的问题是 1874 年康托尔提出的连续统

假设,这个问题至今仍未解决。1938年哥德尔首先提出了可构成公理,从而证明了 CH 和选择公理(AC)与 ZF 的协调性,即 CH, AC 在 ZF 中不可否性。1963年科恩创立了力迫方法,并用力迫法证明了 CH, AC 与 ZF 的独立性,即 CH, AC 在 ZF 中不能证明。

力迫方法是一种非常复杂的构造模型的方法,应用力迫方法,人们证明了许多集合论命题的协调性和独立性,力迫方法极大的推动了公理集合论的发展与应用。

在公理集合论的研究中,有许多问题要用到 ZF (或 ZFC)以外的附加假设才能确定,这里常用的附加假设有:可构成公理;各种大基数公理;以及与 AC 不协调的决定性公理等。目前,力迫方法,大基数公理,决定性公理成为现代公理集合论研究的主流方向。公理集合论的现代成果对其它数学领域有着深远的影响。

【模型论】数理逻辑的一个分支,是研究形式语言及其解释(模型)之间的关系的理论。模型论中的概念和方法,不仅来源于数理逻辑,也有不少来源于代数,特别是泛代数。模型论的成果和方法不但对数理逻辑的其它分支,如逻辑演算、证明论、递归论、公理集合论有很大推动作用,也对数学的一些学科,如代数、数论、拓扑、概率等有相当影响。

由于所涉及的逻辑系统不同,模型论可分为:一阶模型论、高阶模型论、无穷长语言模型论、模态模型论、具有广义量词逻辑的模型论、多值模型论等,一阶模型论是应用最广泛、内容最丰富的部分,是模型论最重要的基础。

模型论中最早的重要结果是勒文海姆-斯科朗定理:设一阶语言 L 中所能表达的命题个数为 λ (超限数),如果 L 中的一个形式理论 T 有无限模型,则 T 有基数为任意 $\alpha \geq \lambda$ 的模型。模型论的最

基本定理是紧致性定理：如果一阶语言 L 的一个命题集 T 的任意有限部分有模型，则 T 自身有模型。紧致性定理有许多有趣的应用，实数理论的非标准模型的存在性就是根据紧致性定理，这是非标准分析的基础。模型论中还有许多构造模型的方法。如斯科朗函数方法、亨金常量方法、塔尔斯基初等链方法、超积方法、鲁宾孙力迫方法等。这些方法不仅推动了模型论的发展，而且有许多应用。例如阿克斯和科琴用超积方法解决了阿廷提出的关于 p 进数域 Q_p 的一个猜想。一个形式理论 T ，如果它的任何两个基数为 α 的模型都同构，则称 T 为 α 范畴理论。莫利关于范畴性的重要定理是：对于可数语言中的完备理论 T ，如果它对某个不可数基数 α 是范畴的，则对于任一不可数基数 β ， T 也是 β 范畴的。在研究范畴性的过程中，莫利提出了一个重要的概念：“稳定性”。谢拉赫对形式理论的稳定性作了深入的研究，提出了许多新的概念和方法。

模型论是一门年轻的学科，它还在发展中。可以预料，随着模型论的深入发展，它将为其他数学分支提供更多的新工具和新方法。

【递归论】数理逻辑的一个分支，是研究算法、可计算性、不可计算性的数学。在古代中国、希腊就有了算法概念，16 世纪，代数里已出现了大量算法，但是有些问题经过长时期的研究，还是找不到解决它们的算法。例如希尔伯特第九问题、半群上字的问题等等。因此人们怀疑是否这些问题没有一致的算法，为了证明解决这些问题的算法不存在，就要求把算法概念予以精确化，因此产生了建立算法的精确数学定义的问题，20 世纪 30 年代，这个问

题才得到了解决。出现了几个等价的定义，其中最重要的有递归函数、图灵机、正规算法等。

递归函数产生后，开始了递归论的研究，递归论中另一个重要概念是算法可产生集的精确描述递归可枚举集，可以证明一切递归集，即算法可判定集都是递归可枚举集，但是存在不是递归集的递归可枚举集。由于许多数学问题都可以化为某个递归可枚举集是否是递归集的问题。因此非递归的递归可枚举集的研究帮助人们证明了希尔伯特第十问题、半群上字的问题等都是不存在算法解的。

进而人们又开始研究是否一切递归可枚举集的不可解的程度是一样的，为了刻划这个问题，1944 年波斯特提出不可解度的概念，并且提出著名的波斯特问题。1957 年弗里德贝格和穆切尼克各自独立地创造了有穷损害优先方法解决了波斯特问题。后来人们又创造了无穷损害优先方法，以及 O^m 优先方法等，这些方法是递归论研究中的重要工具。

随着集合论的发展，递归论也从 ω 上的递归论发展成为在一般序数 α 上的递归论。从而形成了广义递归论。

在计算机科学中，人们也经常使用递归论的概念和方法，递归论主要应用于计算复杂性理论，起初人们把图灵机作为研究计算复杂性的模型考虑计算的时间和空间复杂性。后来在递归论的基础上又建立了抽象计算复杂性理论。近年来递归论的方法大量地用于 P 与 NP 问题的研究中，取得了一系列成果。

【哥德尔不完备性定理】数理逻辑中最重要的定理之一。哥德尔在

1931 年证明了这个定理,它包括两个事实,第一不完备性定理:设 S 是包含算术系统的任意形式系统,则存在命题 F 使得 F 和它的否命题都在 S 中不可证。 F 也称为形式系统 S 内的不可判定命题。第二不完备性定理:在上述形式系统 S 中不能证明它本身的协调性。

哥德尔不完备性定理表明,在一个足够强的公理系统中不能证明自身的协调性,而且在相容性的范围内无论怎样添加公理,它们仍然是不完备的。

哥德尔的工作从正面或反面解答了 20 世纪以来在数学基础方面争论的最根本的问题。同时也给希尔伯特计划以致命的打击。哥德尔定理说明要想证明数学理论的绝对的协调性是不可能,从此以后,人们开始了相对协调性的研究。哥德尔第一不完备性定理的证明方法对递归论的早期发展有很大影响,哥德尔的工作推动了整个数理逻辑的发展。

组 合 数 学

【组合数学】又称组合分析或组合学。研究将有限个元安排到适合(服从)某些限制条件的集合。有三个问题需解决:(1)组态问题,解决存在这种安排的条件,给出明确的结论;(2)组态存在时,确定其数目或将它们进行分类;(3)研究安排的性质和结构,包括最优化问题。

组合数学最早出现的是神话传说：大禹时代（前 2200）的神龟背上驮着的幻方，古代称为“九宫”，即

4	9	2
3	5	7
8	1	6

一般是将 $1, 2, \dots, n^2$ 放到 $n \times n$ 格子中，使每行每列各数之和相等，称 n 阶幻方。缺角棋盘的覆盖问题、柯克曼 15 女生散步问题、欧拉 36 名军官问题都是著名的组合学例子。现代科学技术中，又提出离散性问题及关系结构分析，图论、信息论、编码、实验设计、线性规划等领域也提了一系列问题，促进了组合学的发展，它们互相促进、充实。

$|A \times B| = |A| \cdot |B|$ ；若 $S = S_1 \cup S_2 \cup \dots \cup S_n$ ，且 $S_i \cap S_j = \emptyset$ ，则 $|S| = |S_1| + |S_2| + \dots + |S_n| = \sum_{i=1}^n |S_i|$ 。

n 为正整数， $1 \leq r \leq n$ ， $P(n, r) = n(n-1) \cdots (n-r+1)$ 表示 n 元集中 r 元排列的个数，特别， $P(n, n) = n(n-1) \cdots 2 \cdot 1 = n!$ ， n 元集中 r 元圆排列数为 $P(n, r)/r = n!/r(n-r)!$ ，特别， n 元圆排列数为 $(n-1)!$ 。

$r \leq n$ ， $P(n, r) = r! C(n, r)$ ， $C(n, r) = \begin{Bmatrix} n \\ r \end{Bmatrix} = n!/r!(n-r)!$ 表示 n 元集中 r 元的组合数。特别， $C(n, r) = C(n, n-r)$ ，而 n 元集的所有子集数为 $C(n, 0) + C(n, 1) + \dots + C(n, n) = 2^n$ 。

若 S 有 k 个不同元, 每个元有无限重数, 则 S 的 r 元排列数为 k^r ; S 的 r 元组合数为 $C(k-1+r, r) = \binom{k-1+r}{r}$, 若 S 为有有限重数 $n_1, n_2, \dots, n_k$ 的重集, $n=n_1+\dots+n_k$, 则 S 的排列数为 $n! / n_1! n_2! \dots n_k!$

容斥原理 (包含排斥原理) 指的是 $|\bigcup_{i=1}^n S_i| = \sum_{i=1}^n |S_i| - \sum_{1 \leq i < j \leq n} |S_i \cap S_j| + \dots + (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} |S_{i_1} \cap \dots \cap S_{i_k}| + \dots + (-1)^{n-1} |\bigcap_{i=1}^n S_i|$

帕斯卡公式 $C(n, k) = C(n-1, k) + C(n-1, k-1)$, $1 \leq k \leq n-1$ 。

二项式定理 $(X+y)^n = X^n + \binom{n}{1} X^{n-1}y + \binom{n}{2} X^{n-2}y^2 + \dots + \binom{n}{n-1} Xy^{n-1} + \binom{n}{n} y^n = \sum_{k=0}^n \binom{n}{k} X^{n-k} \cdot y^k$

若数列 $H(0), H(1), H(2), \dots, H(n) \dots$, 其中 $H(n)$ 与它前面的若干项联系的方程对所有大于某个 n_0 的整数 n 都成立, 该方程 (关系式) 称为递归关系。常系数线性齐次递归关系指的是

$H(n) = a_1 H(n-1) + a_2 H(n-2) + \dots + a_k H(n-k), n=k, k+1, \dots (*)$

其中 a_i 是常数, $a_k \neq 0$, 称为 k 阶的, 而各 H 均为一次且没有常数项, $(*)$ 可化为

$H(n) - a_1 H(n-1) - \dots - a_k H(n-k) = 0,$
 $n=k, k+1, \dots。$

$$X^k - a_1 X^{k-1} - a_2 X^{k-2} - \dots - a_k = 0$$

称为 $(*)$ 的特征方程。由 $a_k \neq 0$, $(*)$ 的全部特征根均不为零, 对非零数 q , 当且仅当 q 是 $(*)$ 的一个特征根时, $H(n) = q^n$ 是递归关系 $(*)$ 的一个解。而 $q_1, q_2, \dots, q_k$ 是互不相同的特征根时, $(*)$ 的一般解为

$$H(n) = c_1 q_1^n + c_2 q_2^n + \dots + c_k q_k^n.$$

数列 $a_0, a_1, a_2, \dots, a_n, \dots$ 的生成函数指的是形式幂级数

$$f(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n + \dots,$$

有限序列 $a_0, a_1, a_2, \dots, a_n$ 可看成无限序列 $a_0, a_1, a_2, \dots, a_n, a_{n+1} = 0, a_{n+2} = 0, \dots$, 因之也可称为有限数列的生成函数。二项式系数

$$\begin{bmatrix} n \\ 0 \end{bmatrix}, \begin{bmatrix} n \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} n \\ n \end{bmatrix} \text{ 的生成函数是 } f_n(X) = \begin{bmatrix} n \\ 0 \end{bmatrix} + \begin{bmatrix} n \\ 1 \end{bmatrix} X + \begin{bmatrix} n \\ 2 \end{bmatrix} X^2 + \dots + \begin{bmatrix} n \\ n \end{bmatrix} X^n = (1+X)^n.$$

设 a_n 是 $k(>0)$ 个不同元的多重集的 n 组合数, 则数列 $a_0, a_1, a_2, \dots, a_n, \dots$ 的生成函数 $g_k(X)$ 为

$$g_k(X) = \begin{bmatrix} k-1 \\ 0 \end{bmatrix} + \begin{bmatrix} k \\ 1 \end{bmatrix} X + \begin{bmatrix} k+1 \\ 2 \end{bmatrix} X^2 + \dots + \begin{bmatrix} k+n-1 \\ n \end{bmatrix} X^n + \dots = \frac{1}{(1+X)^k}.$$

特别, $1, 1, \dots, 1, \dots$ 的生成函数为

$$g(X) = \frac{1}{1-X}; 1, 2, \dots, n \text{ 的生成函数为 } g_2(X) = \frac{1}{(1-X)^2}.$$

$A_1, A_2, \dots, A_n$ 均为 G 的子集, 从每个 A_i 中选取一个相异元 $a_1, a_2, \dots, a_n$, 称为 $A_1, A_2, \dots, A_n$ 的一个相异代表组。

配偶定理内容为: $A_1, A_2, \dots, A_n$ 有相异代表组当且仅当对每个 $k=1, 2, \dots, n$, 与每次选取的 $i_1, i_2, \dots, i_k (1 \leq i_1 < \dots < i_k \leq n)$ 都有 $|A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}| \geq k$ 。

相异代表组及配偶定理与图论中二部图的匹配及匹配的有关定理只是两种不同的表述。

鸽笼原理内容为: $n+1$ 或更多只鸽子飞入 n 只笼子,则至少有一个笼子里有多于一只鸽子。又称为抽屉原理、迪里赫莱原理。

拉姆齐定理是鸽笼原理的深刻推广。即适合条件 $1 \leq t \leq q_1, q_2, \dots, q_n$ 的正整数组,必存在最小正整数 $N(q_1, q_2, \dots, q_n; t)$, 依赖于 $q_1, q_2, \dots, q_n, t$, 具有性质:若 $m \geq N(q_1, q_2, \dots, q_n, t)$ 且 S 是 m 元集,将 S 的 t 元子集分在 n 个盒子里,则或有 q_1 个元使它的全部 t 元子集均在第一个盒子里,或者有 q_2 个元使它的全部 t 元子集均在第二个盒子里, $\dots$, 或有 q_n 个元使它的全部 t 元子集均分在第 n 个盒子里。

鸽笼原理的加强形式为: $q_1, q_2, \dots, q_n$ 是正整数,将 $q_1 + q_2 + \dots + q_n - n + 1$ 或更多物件放入 n 个盒子里,则或第一个盒子里至少有 q_1 件,或第二个盒子里至少有 q_2 件, $\dots$, 或第 n 个盒子里至少有 q_n 件。

【图论】见“概率论”中的“图论”。

【树】连通无回路的图。一般无回路的图称为(森)林。 n 点 m 边的图 T , 则树有以下特征:

T 是树 $\Leftrightarrow T$ 连通且 $m = n - 1 \Leftrightarrow T$ 无回路且 $m = n - 1 \Leftrightarrow T$ 无回路,但任二点间添一条边得到一条且仅一条回路 $\Leftrightarrow T$ 连通,但删去任一边必不连通 $\Leftrightarrow T$ 的任二点间有且仅有一条路。

任何 $n (\geq 2)$ 点树 T 至少有 2 片树叶(度为 1 的点)。

G 的生成子图 T 是树,则 T 称为 G 的生成树。 G 有生成树 $\Leftrightarrow G$ 是连通图。

(边)赋权图 G 的生成树中,各边权和最小的,称为最小生成树。首先选边权最小的边 e_1 (当有相同的权,任选一条);当已选好 $e_1, e_2, \dots, e_k$, 则选 e_{k+1} , 使 $e_1, e_2, \dots, e_k, e_{k+1}$ 无回路, 且 e_{k+1} 是权最小者, 直到 $k=n-1$ 结束, 就得到 G 的最小生成树。

若干城镇间修建管线,选取造价最低的设计方案的数学模型就是在连通赋权图中选取最小生成树。

图 G 的回路 M 经过 G 的每条边一次且仅一次,称为欧拉回路,而 G 称为欧拉图。简称 G 为 E 图。若图 G 的路 M 经过每条边一次且仅一次,则称为欧拉路。

1736 年欧拉得到:连通图 G , 若恰有二个奇度点,则存在由一奇度点出发到另一奇度点的欧拉路;若所有点均为偶度点,则必是 E 图,可从任一点出发过每边一次且仅一次回到该点。因此可断定哥尼斯堡七桥问题无解。

类似的是中国邮递员问题。1958 年我国管梅谷提出并解决。邮递员每天投递报纸信函,设计从邮局出发走遍所负责街区的每条街巷再回到邮局的投递方案,使总行程最短。

图 G 的回路若经过每个点一次且仅一次,则称为哈密顿回路,简称 H 回路。有 H 回路的图称为哈密顿图,简称 H 图。若图 G 有过每个点一次且仅一次的路,则称为 H 路。

1859 年,哈密顿将正十二面体的 20 个顶点表示世界上 20 个大城市,棱表示交通线,在上面找出经过每个点一次且仅一次的“周游世界”游戏。他本人给出了方案,且获得了专利。与找到判别条件的欧拉图不同,哈密顿图的判别条件至今仍是图论著名难题之一。

类似的是货郎(推销员)问题。货郎要到若干城镇去,如何设计从驻地出发经过每个城镇至少一次而回到驻地的旅行方案,使旅

费最低(或旅程最短)? 至今也未找到满意的方法。

平面图是指能画在平面上而使得没有任何二边相交的图。平面图 G 的各边将平面划分成若干区域, 每个区域称为 G 的一个面, 包围面的各边称为该面的边界; 最外层各边界围成无限面。

有 n 点 m 边 f 面的连通平面图, 则有以下关系式成立, 称为欧拉公式: $n - m + f = 2$

点数 $n \geq 3$ 的简单连通平面图适合关系式: $m \leq 3n - 6$, 而平面图 G 的每个面均由 k 条边围成, 则 $m \leq k(n - 2) / (k - 2)$ 。

最重要的非平面图是 K_5 与 $K_{3,3}$ 。在图 G 的一条边上增加若干个 2 度点, 即将一条边看成若干条相连边, 并不影响 G 的平面性; 同样, 若一个面的边数多于 3, 当删去 2 度点, 即将相邻的两条边看成一条边, 也不影响 G 的平面性。增加或删去 2 度点, 由图 G 得到 G_1 , 则称 G, G_1 是同胚的, 库拉托夫斯基证明了以下重要结果:

G 是平面图当且仅当 G 不含与 K_5 或 $K_{3,3}$ 同胚的子图。

图 G 的点集 $V = V_1 \cup V_2, V_1 \cap V_2 = \emptyset$, 且 V_1 的点仅与 V_2 的点邻接, V_2 的点仅与 V_1 的点邻接, 则称 G 是二部图, 记成 $G = (V_1, V_2; E)$, 也称二分图或偶图。 V_1 的每个点与 V_2 的每个点邻接的二部图, 称为完全二部图。 $|V_1| = m, |V_2| = n$ 的完全二部图, 记作 $K_{m,n}$ 。

G 是二部图 $\Leftrightarrow G$ 没有长为奇数的回路。

二部图 $G = (V_1, V_2; E)$, 若对 V_1 的每个点 v_i 有 V_2 的不同点 u_i , 使 $v_i u_i \in E$, 则 $\{v_1 u_1, v_2 u_2, \dots, v_{|V_1|} u_{|V_1|}\}$ 称为 V_1 对 V_2 的一个匹配。

二部图 $G = (V_1, V_2; E)$, 存在 V_1 对 V_2 的匹配当且仅当对每个 $i, 1 \leq i \leq |V_1|$, V_1 中的 i 个点至少与 V_2 中的 i 个点邻接。本结

果常称为婚配定理。

【四色猜想】画在平面上或球面上的地图,每个国家是个连通区域,有共同边界的国家要着不同颜色,需要用多少种颜色?下图说明至少要 4 种。

1890 年证明 5 种一定够用。然而实践中 4 种已够用,人们猜想四色已够用,而一直未能证明它,称为“四色猜想”。1976 年美国有人用计算机证明“四色猜想”成立,但有人怀疑,有人仍希望得到直接证明。

【网络流】有向赋权图 G , 边 $v_i v_j$ 的权称为该边的容量, 记作 c_{ij} , G 称为容量网络。 G 有点 v_s 称为源(发点), 只有引出的边; 另有点 v_t 称为汇(收点), 只有引入的边, 其余点为中间点, 既有引入的也有引出的边。

$f = \{f_{ij}\}$ 为定义在边上的函数, 若 f 适合以下条件, 则称为 G 上的可行流: (1). 对每条有向边 $v_i v_j$, $0 \leq f_{ij} \leq c_{ij}$; (2). 对每个中间点

$v_i, \sum_{v_i v_j \in E} f_{ij} - \sum_{v_j v_i \in E} f_{ji} = 0$, 即流入量与流出量相同。

当 f 是可行流, $\sum_{v_s v_j \in E} f_{sj} - \sum_{v_j v_s \in E} f_{js} = e(f)$ 称为 f 的流量。

在流量网络上, 求使得流量 $e(f)$ 达到最大的可行流, 称为网络最大流问题。1955 年开始研究, 开辟了图论应用的新途径。

数 论

【数论】最古老的数学学科之一。是研究数的规律, 特别是整数性质的一门纯粹数学学科。古希腊人对数论的主要贡献是整数的整除性理论, 这是数论中最古老, 最重要的部分之一, 它的建立不晚于公元前 5 世纪末到 4 世纪初。尼可马修斯的《算术入门》是数学史上第一本数论书籍, 它的大多数主题与欧几里得《原本》中关于数论的部分相同。但风格迥异。该书总结了早期毕达哥拉斯学派的有关论述, 也提出了一些新理论, 丢番图在《算术》中讲述了许多数论命题, 成了费马、欧拉、高斯等人研究数论的出发点。高斯的《算术探讨》开始了现代数论。

数论按研究的方法又分成初等数论, 解析数论, 代数数论, 超越数论, 几何数论, 概率数论, 堆垒数论等。

数论是一门高度抽象的纯粹数学学科, 近来, 随着计算机科学及应用数学的发展, 它的应用也日益广泛, 如初等数论中的许多成果应用于组合论, 编码中, 数论的许多深刻成果在近似分析、快速变换等方面都得到应用。

中国在数论研究上有悠久的历史,《周髀》中已有数论启蒙思想,勾股定理即早期不定方程的研究,南宋秦九韶的解联立同余式组方法早于欧拉、高斯五百年。在现代,著名数学家华罗庚是中国解析数论的前驱,他的《数论导引》、《堆垒素数论》是世界有影响的名著。在他的领导下,中国数论工作者做了大量工作,最突出的当推陈景润在研究哥德巴赫猜想时的 $(1,2)$,它是目前这一工作的最佳成果。

【初等数论】以整数的整除理论,不定方程,同余式等为主要研究内容,以算术方法为主要研究方法的一个数论分支。近几十年来,初等数论在计算机科学,组合数学,代数编码,信号的数字处理等领域得到广泛应用。

整数的整除 整除是数论中一个基本概念,若 a, b, c 为整数,且 $a = bc$, 称 b 整除 a , 记 $b | a$, 或说 a 被 b 整除。 b 称为 a 的一个因数, a 称为 b 的一个倍数。对于 a, b , 若不存在整数 c 使 $a = bc$, 称 b 不整除 a , 记 $b \nmid a$ 。对于任意整数 a , $\pm 1, \pm a$ 是 a 的因数, 称为 a 的当然因数, 若 $|b| < |a|, b | a$, 称 b 为 a 的真因数。一个大于 1 的整数, 若只有当然因数, 称为素(质)数, 否则称为合数。整除有以下基本性质:

- (1) $\forall a, \pm 1 | a, \pm a | a$;
- (2) 若 $c | a, c | b$, 则 $c | a \pm b; c | ka + lb$,
 $c | am$, 其中 k, l, m 为任意整数;
- (3) 若 $b | a$, 则 $|b| \leq |a|$, 特别, 若 $a | b$ 且 $b | a$, 则 $|a| = |b|$ 。

带余除法定理 a, b 是两个任意整数, b 大于零, 则存在两个唯一的整数 q, r , 使得 $a = bq + r, 0 \leq r < b$ 成立, 称 q 为 b 除 a 所得的不完全商, r 为 b 除 a 所得的余数。利用带余除法定理可以对给

定的两个整数 a, b , 判断 b 是否整除 a 。

最大公因数 若 a, b, c 是整数, $c|a, c|b$, 称 c 是 a, b 的一个公因数。若 d 是 a, b 的一个公因数, 且 d 是 a, b 任意公因数的倍数, 称 d 是 a, b 的一个最大公因数, 若 a, b 不全为 0, a, b 的最大公因数 d 必不为 0, 若 $d \geq 0$, 记 $d = (a, b)$, 任意两个整数均存在最大公因数, 且 (a, b) 由 a, b 唯一确定。求两个整数的最大公因数, 可以累次利用带余除法, 这种方法称为辗转相除法, 是中国古代数学家创造的一种方法(也称欧几里得算法)。辗转相除法还在求最小公倍数、解一次不定方程、解一次同余式中有重要应用。若 $(a, b) = 1$, 称 a, b 互素(质), 两个整数 a, b 互素, 当且仅当存在整数 s, t , 使 $as + bt = 1$ 。若已知两个整数 a, b 的典型分解式: $a = \pm p_1^{k_1} \cdots p_t^{k_t}$, $b = \pm p_1^{l_1} \cdots p_t^{l_t}$, $k_i \geq 0, l_i \geq 0, p_i$ 是素数, 则 $(a, b) = p_1^{m_1} \cdots p_t^{m_t}$, 其中 $m_j = \min(k_j, l_j)$

最大公因数的概念及主要性质可以推广到 k 个整数, $k \geq 2$ 。

最小公倍数 若 a, b, c 是整数, 且 $a|c, b|c$, c 是 a, b 的一个公倍数, 若 c 是 a, b 的一个公倍数, 且 c 是 a, b 任意公倍数的因数, 称 c 是 a, b 的一个最小公倍数, 若 $c > 0$, 记 $[a, b]$ 。任意两个整数均存在最小公倍数, 且 $[a, b]$ 由 a, b 唯一确定, 若 a, b 全不为 0, $[a, b] = \frac{|ab|}{(a, b)}$ 。若 a, b 有典型分解式 $a = \pm p_1^{k_1} \cdots p_t^{k_t}$, $b = \pm p_1^{e_1} \cdots p_t^{e_t}$, 则 $[a, b] = p_1^{n_1} \cdots p_t^{n_t}$ 。其中 $n_j = \max(k_j, e_j)$, 最小公倍数的概念及主要性质可以推广到 k 个整数($k \geq 2$)。

算术基本定理 任意大于 1 的整数 a 均可唯一地表示为素数的乘积, 即若 $a = p_1 \cdots p_s = q_1 \cdots q_t$, 其中 p_i, q_j 均为素数, 且 $p_1 \leq p_2 \leq \cdots \leq p_s, q_1 \leq q_2 \leq \cdots \leq q_t$, 则 $s = t, p_i = q_i, 1 \leq i \leq s$, 上述表达式可以写为 $a = p_1^{k_1} \cdots p_e^{k_e}$, p_i 是素数, 且 $p_i \neq p_j$, 称为 a 的典型分解式。素数

作为整数分解的基本构件,有如下主要性质:(1)若 p 是素数,对任意整数 a ,或 $(p, a) = 1$,或 $p \mid a$;(2)若 p, q 是素数,则 $p = q$ 或 $(p, q) = 1$;(3)若 p 是素数, $p \mid ab$,且 $p \nmid a$,则 $p \mid b$ 。算术基本定理是整数理论的重要定理,也称为整数的唯一分解定理。初等数论中不少结果直接或间接与它有关。

筛法 数论中有广泛应用的一个初等方法,最早由古希腊厄拉多塞提出,现在常称这种筛法为厄拉多塞筛法,具体做法是:写出 1 到 n 的所有自然数,划掉 1,第一个剩下的 2 是素数,划掉 2 的倍数,(它们都是合数),此时第一个剩下的 3 是素数,划掉 3 的倍数(它们都是合数)……,如此下去划完为止。事实上,要求不超过 n 的所有素数,只需在从 1 到 $\sqrt{n}$ 的所有自然数中去筛。筛法之得名来源于这个求素数的过程。现有的素数表多是利用上述筛法稍加变化做出的。目前有了电子计算机,造出了更大的素数表。1977 年扎盖尔造出 $1-5 \times 10^7$ 的素数表,为目前较完善的一个。

早在公元前 3 世纪,欧几里得就证明了素数有无穷多个,不仅如此,一切 $4n-1, 6n-1$ 型的素数也有无穷多个。素数的分布是数论中极富吸引力的重要问题之一,现有的结果是:1—100 间有 25 个素数;1—200 间有 46 个素数,1—1000 间有 168 个素数,1001—2000 间有 135 个素数,2001—3000 间有 127 个素数,3001—4000 间有 120 个素数,4001—5000 间有 119 个素数,等等。若 $\pi(n)$ 表示大于 n 的素数的个数,欧几里得指出 $\lim_{X \rightarrow \infty} \pi(X) = \infty$,且 X 越大, $\frac{\pi(X)}{X}$ 越接近 0, X 越大, $\frac{\pi(X)}{X/\ln X}$ 越接近 1。大约 1800 年,勒让得,高斯猜测 $\lim_{X \rightarrow \infty} \frac{\pi(X)}{X/\ln X} = 1$,1896 年阿达玛等人独立证明了高斯这一猜测,这便是今日所说的素数定理。

【不定方程】数论中有悠久历史的一个分支,内容丰富。指未知数个数多于方程的个数,且它们的解受某种限制的方程(比如要求解是整数或正整数)。讨论不定方程,即决定何时解,解的个数以及求出所有适合条件的解。古希腊丢番图所著《算术》一书中,讲了一些具体不定方程的解法,内容丰富,方法巧妙,人们也称不定方程为丢番图方程来纪念他。在丢番图以前,中国古代的《周髀算经》中已有了不定方程。

二元一次不定方程 形如

(1) $aX + by = c$, 其中 a, b, c 是给定整数, $abc \neq 0$ 的方程。(1)有整数解 $\Leftrightarrow (a, b) | c$ 。若 $(a, b) = 1$, (X_0, y_0) 是(1)的一个解,则(1)的解集为 $\{X = X_0 + bt, y = y_0 - at \mid t \text{ 为任意整数}\}$ 。

三元一次不定方程 形如

(2) $aX + by + cz = n$, 其中 a, b, c, n 是给定整数, $abcn \neq 0$ 的方程。若 $(a, b, c) = 1$, $(a, b) = d$, $a = a_1d$, $b = b_1d$, (X_0, y_0, z_0) 是(2)的一个特解,则(2)的解集为 $\{X = X_0 + b_1t_1 - u_1ct_2, y = y_0 - a_1t_1 - u_2ct_2, z = z_0 + dt_2 \mid u_1, u_2 \text{ 满足 } a_1u_1 + b_1u_2 = 1, t_1, t_2 \text{ 为任意整数}\}$ 。

$s(\geq 2)$ 元一次不定方程 形如

(3) $a_1X_1 + \cdots + a_sX_s = n$, 其中 $a_1, \cdots, a_s, n$ 是给定整数, $a_1 \cdots a_s \neq 0$ 的方程。(3)有整数解 $\Leftrightarrow (a_1, \cdots, a_s) | n$ 。

佩尔方程 形如 $X^2 - Dy^2 = \pm 1$ 的二元二次不定方程,有时 $X^2 - Dy^2 = \pm 4$ 的二元二次不定方程也称佩尔方程。

对于佩尔方程(1): $X^2 - Dy^2 = 1$, 若 D 是非完全平方的正整数,则(1)有无穷多组整数解,并且,若 $X_0^2 - Dy_0^2 = 1, X_0 > 0, y_0 > 0$, 且满足对(1)的任一解 $(X, y), X > 0, y > 0, X_0 + y_0\sqrt{D} \leq X + y\sqrt{D}$, 称 (X_0, y_0) 为(1)的基本解,则(1)的任一解 (X, y) 满足 $X + y\sqrt{D} =$

$\pm(X_0+y_0\sqrt{D})^n$, 其中 n 是某一整数, 并且对任一 n , 上面所得 (X, y) 均是(1)的解。

勾股数(商高数) 是不定方程 $X^2+y^2=z^2$ 的正整数解, 也为最早的二次不定方程。 $X^2+y^2=z^2$ 满足 $(X, y)=1$,

$2 \nmid X, X, y, z > 0$ 的全部整数解可表为

$$X=2ab, y=a^2-b^2, z=a^2+b^2,$$

其中 $a > b > 0, a, b$ 奇偶不同, 且 $(a, b)=1$ 。中国古代《周髀算经》一书中已提到“故折矩以为勾广三, 股修四, 弦隅五”。这说明 $(3, 4, 5)$ 是一个解, 它是勾股定理最有名的一组正整数解。《九章算术》中给出了 $(5, 12, 13), (8, 15, 17), (7, 24, 25), (20, 21, 29)$ 等八组勾股数, 1844 年, 清朝陈杰在《算法大成》中将 $X^2+y^2=z^2$ 的正整数解做了归纳, 1906 年黄宗宪在《惔笑不计》一书中, 有“三角垛堆整数勾股术”, 写出了关于勾股正整数解的通式。

不定方程组 求若干个不定方程的公共解称为解不定方程组。 n 元一次不定方程组问题可由消元法化为一次方程求解。“五家共井”、“百鸡问题”都是中国最古老的不定方程组问题。

【同余式】同余是数论中一个重要的概念, 它是整数集的一个关系。 a, b 是两个整数, m 是一正整数, 若 $m \mid a-b$, 称 a 与 b 模 m 同余, 记 $a \equiv b \pmod{m}$ 。同余是一个等价关系, 具有自反性、对称性、传递性。同余有以下主要性质:

若 $a_1 \equiv b_1 \pmod{m}, a_2 \equiv b_2 \pmod{m}$, 则 $a_1 \pm a_2 \equiv b_1 \pm b_2 \pmod{m}$;

若 $a \equiv b \pmod{m}, k > 0$, 则 $ka \equiv kb \pmod{m}$;

若 $a \equiv b \pmod{m_i} (1 \leq i \leq n)$, 则 $a \equiv b \pmod{[m_1, \dots, m_n]}, ([m_1, \dots, m_n]$ 表 $m_1, \dots, m_n$ 的最小公倍数);

若 $a \equiv b \pmod{m}, d > 0, d \mid m$, 则 $a \equiv b \pmod{d}$;

若 $a \equiv b \pmod{m}$, n 正整数, 则 $a^n \equiv b^n \pmod{m}$;

若 $ac \equiv bc \pmod{m}$, 且 $(c, m) = 1$, 则 $a \equiv b \pmod{m}$;

若 $ac \equiv bc \pmod{m}$, 而 $(c, m) = d > 1$, 则 $a \equiv b \pmod{\frac{m}{d}}$ 等。

利用同余的性质可以简化求余数的计算得到整数整除的某些判别法。关于整数整除的某些判别法, 有:

若一个数 n 的末位数字为偶数, 则 $2 | n$;

若一个数 n 的末位数字为 0 或 5, 则 $5 | n$;

若一个数 n 的末两位数被 4 整除, 则 $4 | n$;

若一个数 n 的末两位数被 8 整除, 则 $8 | n$;

若 $n = a_m a_{m-1} \cdots a_2 a_1$ 是一个 m 位数, 因为 $10 \equiv 1 \pmod{9}$, 所以 $n \equiv a_m + a_{m-1} + \cdots + a_1 \pmod{9}$, 同样, 由 $10 \equiv 1 \pmod{3}$, 所以 $n \equiv a_m + a_{m-1} + \cdots + a_1 \pmod{3}$;

若 $n = a_m a_{m-1} \cdots a_2 a_1$, 令 $\alpha = a_2 a_1 + a_6 a_5 + \cdots$, $\beta = a_4 a_3 + a_8 a_7 + \cdots$, 因为 $100 \equiv 100^3 \equiv \cdots \equiv 100^{2k+1} \equiv -1 \pmod{101}$, $100^2 \equiv 100^4 \equiv \cdots \equiv 100^{2k} \equiv 1 \pmod{101}$, 所以 $n \equiv \alpha - \beta \pmod{101}$ 。

若 $ab = c$, 则 $ab \equiv c \pmod{m}$, 利用它, 可以得通常验算的弃九, 弃十一法等。

一个整系数多项式 $f(X) = a_0 + a_1 X + \cdots + a_n X^n$, m 是一正整数, 称 $f(X) \equiv 0 \pmod{m}$ 为模 m 的同余式, 若 $a_n \not\equiv 0 \pmod{m}$, 称 n 为这个同余式的次数, 若存在整数 a , $f(a) \equiv 0 \pmod{m}$, 则称 $X \equiv a \pmod{m}$ 为这个同余式的一个解, 讨论同余式的目的, 即判断是否有解, 有解时求出所有解来。

若 $a \not\equiv 0 \pmod{m}$, 形如 $aX \equiv b \pmod{m}$ 的同余式称一元一次同余式, $a \not\equiv 0 \pmod{m}$ 、 $aX \equiv b \pmod{m}$ 有解 $\Leftrightarrow (a, m) | b$, 有解时, 有 (a, m) 个模 m 互不同余的解。若 $(a, m) = d > 1$, 且若 X_0 是该同余

式的一个解,则 $X_0, X_0 + \frac{m}{d}, X_0 + \frac{2m}{d}, \dots, X_0 + \frac{(d-1)m}{d}$ 为模 m 互不同余的全部解。

若 $a_i \not\equiv 0 \pmod{m}$, $a_1X_1 + a_2X_2 + \dots + a_nX_n \equiv b \pmod{m}$ 称 n 元一次同余式, $a_1X_1 + \dots + a_nX_n \equiv b \pmod{m}$ 有解 $\Leftrightarrow (a_1, \dots, a_n, m) \mid b$, 有解时, 有 $m^{n-1}(a_1, \dots, a_n, m)$ 个模 m 互不同余的解。对于高次同余式, 若 $m = m_1m_2$, 且 $(m_1, m_2) = 1$, 则 $f(X) \equiv 0 \pmod{m}$ 有解 $\Leftrightarrow f(X) \equiv 0 \pmod{m_i} i = 1, 2$, 有解, 若 $f(X) \equiv 0 \pmod{m}$ 的解数为 $T(m)$, 则 $T(m) = T(m_1) + T(m_2)$ 。拉格朗日亦证明了模素数 p 的 n 次同余式 $f(X) \equiv 0 \pmod{p}$ 最多有 n 个解, 特别 $X^p - X \equiv 0 \pmod{p}$ 恰有 p 个解。

费马小定理 p 是素数, $p \nmid a$, 则 $a^{p-1} \equiv 1 \pmod{p}$ 。也可叙述为若 p 是素数, a 是任一整数, 则 $a^p \equiv a \pmod{p}$ 。这个定理可用来简化求余数的计算, 也可用来判定一个数是否合数等等, 应用很广泛, 因此是数论的一个重要结论。

威尔逊定理 若 p 是素数, 则 $(p-1)! \equiv (-1) \pmod{p}$ 。它的逆命题也是成立的, 即若 n 是大于 1 的整数, $(n-1)! \equiv (-1) \pmod{n}$, 则 n 是素数, 这个定理给出了一个判别整数 n 是否素数的重要条件, 理论上十分重要, 但当 n 充分大时, 计算量太大, 无实用价值。

剩余类与完全剩余系, 简化剩余系 给定正整数 m , 可将整数集 Z 分成 m 个子集, 令 $[i]$ 表示被 m 除余数是 i 的整数的集合, 则 $Z = [0] \cup [1] \cup \dots \cup [m-1]$, $[i]$ 称模 m 的一个剩余类, $[i]$ 中任二整数模 m 同余, 不在同一类的任二整数模 m 不同余, 对给定 m , 模 m 的剩余类共 m 个, 从每一类取一个数, 这 m 个数称以 m 为模的完全剩余系, 常取 $\{0, 1, \dots, m-1\}$ 。在与 m 互素的各类中各取一

数, 所得 $\varphi(m)$ 个数称以 m 为模的简化剩余系, 如 $m=4, \{0, 1, 2, 3\}$ 为模 4 的完全剩余系, 模 4 的简化剩余系为 $\{1, 3\}$ 。

二次剩余 m 为正整数, a 是整数, $(m, a) = 1$, 若 $X^2 \equiv a \pmod{m}$ 有解, 称 a 对模 m 二次剩余, 否则称 a 对模 m 二次非剩余。两个二次剩余(或二次非剩余)之积仍二次剩余, 一个二次剩余与一个二次非剩余之积为二次非剩余。若 p 是奇素数, 欧拉给出: 若 $(a, p) = 1$, 则 a 是模 p 的二次剩余的充要条件是 $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$, a 是模 p 的二次非剩余的充要条件是 $a^{\frac{p-1}{2}} \equiv (-1) \pmod{p}$ 。由此知在模 p 的简化剩余系中恰有 $\frac{p-1}{2}$ 个二次剩余, 它们是: $1^2, \dots, (\frac{p-1}{2})^2 \pmod{p}$, 有 $\frac{p-1}{2}$ 个二次非剩余。当 p 较大时, 欧拉判别条件不便使用。

设 p 是奇素数, a 是整数, 令

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & a \text{ 是模 } p \text{ 的二次剩余;} \\ -1, & a \text{ 是模 } p \text{ 的二次非剩余;} \\ 0, & p|a. \end{cases}$$

称 $\left(\frac{a}{p}\right)$ 为 a 对 p 的勒让得符号, 利用 $\left(\frac{a}{p}\right)$, 易知 $X^2 \equiv a \pmod{p}$ 是否有解。勒让得符号有以下性质:

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p};$$

若 $a \equiv a_1 \pmod{p}$, 则 $\left(\frac{a}{p}\right) = \left(\frac{a_1}{p}\right)$;

$$\left(\frac{a_1 a_2 \cdots a_n}{p}\right) = \left(\frac{a_1}{p}\right) \left(\frac{a_2}{p}\right) \cdots \left(\frac{a_n}{p}\right);$$

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$$

若 p, q 是不同的奇素数, 则 $\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} \left(\frac{p}{q}\right)$, 称为二次反转

定律,这个定律首先由高斯给出严密证明。二次反转定律是同余式中一个基本结果,高斯誉它为算术中的宝石,利用勒让得符号可判断 $X^2 \equiv a \pmod{p}$ 是否有解。

【特殊类型的数】一些具有特殊性质的数。若正整数 n 的一切小于自身的正因数之和恰为 n ,称 n 是完全数。公元前 6 世纪,毕达哥拉斯首先接触完全数问题。欧几里得证明了若 $p, 2^p - 1$ 是素数,则 $2^{p-1}(2^p - 1)$ 是完全数。两千年之后,欧拉证明了每个偶完全数都有上述形式。这样,求偶完全数问题,归结为求 $2^n - 1$ 是素数的问题 ($2^n - 1$ 是素数,必有 n 是素数), $p = 2, 3, 5, 7$ 时,对应的 6, 28, 496, 8128 是 4 个完全数。目前所知的完全数均为偶完全数,并且这些完全数与莫森素数有关,若 $2^p - 1$ 是莫森素数,则 $2^{p-1}(2^p - 1)$ 是偶完全数,反之亦真。关于完全数的两个尚待解决的难题是:是否存在奇完全数以及是否存在无穷多个偶完全数。

一对正整数 m, n , 若它们彼此等于对方的所有因数(不含本身)之和,称 m, n 是一对亲和数。毕达哥拉斯给出了第一对亲和数 220 和 284,事实上,220 的所有因数为 1, 2, 4, 5, 10, 11, 20, 22, 44, 55, 110, 其和为 284, 284 的所有因数为 1, 2, 4, 71, 142, 其和为 220。笛卡儿给出了求亲和数的一个规则:设有以 2 为底的一个乘幂,如果它的 3 倍减 1, 6 倍减 1, 以及它的平方的 18 倍减 1 都是素数,则令最后一个素数乘以所设 2 的乘幂的 2 倍,即得亲和数中的一个,华林也给过求亲和数的公式,目前,已知道一千对以上的亲和数。

若 p 是素数,称 $2^p - 1$ 为莫森数,若它是素数,称莫森素数,此数常记为 M_p , 当 M_p 是莫森素数时, $\frac{1}{2}M_p(M_p + 1)$ 是偶完全数,所

以莫森素数与偶完全数间有一一对应,至今所知最大莫森素数是 M_{216091} 。莫森素数是否有无穷多个,是数论的又一难题,通过发现莫森数,可以达到发现已知最大素数的目的,因此,引起数学家的兴趣。莫森素数有以下性质:

若 p 是奇素数, $q|M_p$, 则存在 k , $q=2kp+1$;

若 $p=4n+3$, p 是素数, $2p+1=4n+7$ 是素数 $\Leftrightarrow 2p+1|M_p$ 。

形如 $2^{2^n}+1$ 的正整数称费马数,记为 F_n ,其中 $F_0=3, F_1=5, F_2=17, F_3=257, F_4=65537$ 都是素数,1640 年费马曾猜想,一切费马数都是素数,但 1732 年欧拉指出 $641|F_5$; $F_5=641 \times 6700417$,从而否定了费马的这个猜想。但究竟有多少费马数是素数,是有限个还是无限个? 是否有无限多个费马数是合数? 这些问题都是没有解决的难题。已经知道了 48 个费马数不是素数, F_{17} 究竟是素数还是合数尚不得而知。费马数与尺规作图问题有关,高斯证明了,若 F_n 是素数,则正 F_n 边形能用尺规作出。

【数论函数】定义在整数集或自然数集上的实或复值函数,又称算术函数,常见的有:

欧拉函数 从 1 到 n 中与 n 互素的整数个数,记 $\varphi(n)$,若 $(m_1, m_2)=1$, 则 $\varphi(m_1 m_2)=\varphi(m_1)\varphi(m_2)$, 若 p 是素数, 则 $\varphi(p)=p-1$, 一般, 若 $n=p_1^{k_1} \cdots p_t^{k_t} (k_i > 0, p_i, p_j \text{ 互不相同})$, 则 $\varphi(n)=n(1-\frac{1}{p_1}) \cdots (1-\frac{1}{p_t})$ 。

除数函数 正整数 n 的正因数个数,记 $d(n)$ 。 $d(n)=\sum_{d|n} 1$, 如 $d(1)=1, d(2)=2, d(4)=3, d(p)=2, \dots$; 若 $n=p_1^{k_1} \cdots p_t^{k_t}$, 则 $d(n)=(k_1+1) \cdots (k_t+1)$ 。

除数和函数 也称因数和函数,指正整数 n 的所有正因数之

和,记

$$\tau(n), \tau(n) = \sum_{d|n} d, \tau(1)=1, \tau(2)=3, \tau(4)=7, \tau(p)=p+1, \dots, \text{若}$$

$$n=p_1^{k_1} \cdots p_t^{k_t}, \text{则 } \tau(n) = \frac{p_1^{k_1+1}-1}{p_1-1} \cdots \frac{p_t^{k_t+1}-1}{p_t-1}.$$

麦比乌斯函数 正整数 n 的麦比乌斯函数记为 $\mu(n)$,

$$\mu(n) = \begin{cases} 1, & n=1; \\ (-1)^r, & n \text{ 是 } r \text{ 个不同素数之积;} \\ 0, & n \text{ 被素数平方整除。} \end{cases}$$

$$\mu(2)=-1, \mu(4)=0, \mu(p)=-1, \dots.$$

可乘函数 以自然数集为定义域的函数 f , 若满足: (1) $\forall a, f(a) \neq 0$, (2) 若 $(a, b)=1$, 则 $f(ab)=f(a) \cdot f(b)$, 上述数论函数均为可乘函数。

【几个著名猜想】数论的几个著名难题, 被称为数学皇后皇冠上的明珠, 这些难题叙述清楚, 简单易懂, 即使外行人也很容易理解, 然而它们又都是非同一般的数学难题。要解决它们, 不但要有深广的数学知识做基础, 而且还需把现有的理论和方法用出新水平, 并创造新理论发现新方法, 因此, 对这些问题的研究, 推动了数学多分支的发展。

费马猜想 大约在 1637 年, 费马在阅读丢番图著《算术》一书的拉丁文译本时, 读到第 II 卷第八命题“将一个平方和分为两个平方数”, 在书的页边空白处写了一段话, 意思是说“将一个立方数分成两个立方数, 一个四次幂分成两个四次幂, 或者一般地将一个高于 2 次幂分成两个同次的幂, 这是不可能的, 关于此, 我确信已发现了一种奇妙的证法, 可惜这里的空白太小, 写不下”。用现代数学

语言叙述,费马猜想是说, $n > 2$ 时,方程 $X^n + y^n = z^n$ 没有正整数解。

费马猜想又常称费马大定理,要证费马猜想是对的,只需证明 $X^4 + y^4 = z^4$ 及 p 是奇素数时 $X^p + y^p = z^p$ 均无正整数解。费马说,他用无穷递降法证明了前者。1676 年,贝西也对 $n=4$ 给出了证明,欧拉对 $n=3,4$ 都给出了证明,此外勒让得与狄利克雷对 $n=5$ 给出了证明,19 世纪中期,库默对 $n < 100$ (除 37,59,67 外)的奇素数给出了证明。1908 年,德国数学家佛尔夫斯克尔遗言,将 10 万马克奖给第一个证明费马大定理的人。从费马提出这一猜想至今,已过去三个半世纪,问题仍未解决。近年来主要结果有:

(1)1977 年瓦格斯塔夫证明了,对于每一个素数 $p < 125000$,费马定理都是对的。

(2)1983 年,伐尔廷斯证明了 1922 年英国数学家莫德尔提出的猜想:如果 $F(X,y)$ 为有理多项式,代数曲线 $F(X,y)=0$ 的亏格 ≥ 2 ,则 $F(X,y)=0$ 至多只有有限多个有理解。这保证, $n \geq 4$ 时至多只有有限个 n 使 $X^n + y^n = z^n$ 有整数解。

(3)1985 年,爱德列曼和海斯·布朗用解析数论的方法,证明了存在无穷多个素数 p ,使不存在整数 X,y,z ,满足 $X^p + y^p = z^p$ 成立, $p \nmid Xyz$ 。

(4)1993 年 6 月 23 日英国数学家 *A. Wiles* 在剑桥大学牛顿数学研究所做题为“模形式,椭圆曲线和伽罗瓦表示”的学术报告。最后宣布“我证明了费马猜想”。有关专家和权威人士的初步反映大都持肯定态度。

哥德巴赫猜想 1742 年 6 月,哥德巴赫在与欧拉的通信中提出的关于素数的一个著名猜想,它包含两个问题:

问题 1:是否每一个不小于 6 的偶数均可表为两个奇素数之

和?

问题 2: 是否每个不小于 9 的奇素数均可表为三个奇素数之和?

19 世纪结束时, 猜想的研究只限于具体的验证工作。现在已知, 33×10^6 以内的偶数都是对的。从而相应的奇数也有同样结论。1937 年, 苏联数学家维诺格拉朵夫创造了“三角和方法”, 第一个证明了当奇数 N 充分大时, 问题 2 成立, 这个事实称哥德巴赫-维诺格拉朵夫定理, 也简称维氏定理。问题 1 的研究要困难得多。20 世纪 20 年代以后, 特别是 60 年代取得一系列进展: 1920 年, 挪威布龙证明了每一个充分大的偶数都可表为两个数之和, 其中每一个均不超过 9 个素因数之积, 常记为 $(9, 9)$ 。1924 年, 德国雷特马赫证明了 $(7, 7)$, 1932 年, 英国埃斯特曼证明了 $(6, 6)$, 1938—1940 年, 原苏联布赫夕太勃证明了 $(5, 5)$, $(4, 4)$, 中国数学家王元先后于 1956、1957、1962 年, 证明了 $(3, 4)$, $(3, 3)$, $(2, 3)$, $(1, 4)$, 中国另一位数学家潘承洞于 1962、1963 年先后证明了 $(1, 5)$, $(1, 4)$, 1965 年, 布赫夕太勃证明了 $(1, 3)$, 1966 年, 中国著名数学家陈景润得到 $(1, 2)$, 于 1973 年正式发表。这是目前最好的结果, 被命名为陈氏定理。但距问题的解决仍有漫长之路。这里值得特别提起的是, 中国数学家在这个著名数学难题的研究中, 有着突出的贡献。早在 1938 年, 著名数学家华罗庚就证明了几乎所有偶数都能表成两个奇素数之和, 1941 年, 他对维氏“三角和方法”做了非常深刻的研究和改进, 并对维氏定理做了重要推广。华罗庚证明了每一个充分大的奇数 N , 皆可表为三个奇素数的 k 次方之和: $N = p_1^k + p_2^k + p_3^k$, k 为任意给定的正整数, 维氏定理即 $k=1$ 时的特例。

孪生素数猜想 若 $p, p+2$ 都是素数, 称为一对孪生素数, 记作 $(p, p+2)$ 。孪生素数猜想认为有无穷多对孪生素数, 这是至今

未解决的一个难题。已知 10^5 以内共有 1224 对孪生素数, 10^6 以内共有 8164 对, 10^9 以内共有 3424506 对。1989 年发现了三对最大的孪生素数: $663777 \times 2^{7650} \pm 1$, $571305 \times 2^{7701} \pm 1$, $1706595 \times 2^{11235} \pm 1$, 关于这个问题, 目前最好的结果是中国数学家陈景润得到的: 存在无穷多个素数 p , 使 $p+2$ 为不超过两个素数之积。

从筛法来看, 孪生素数猜想与哥德巴赫猜想是两个难度相当的姐妹问题。

【代数数论】数论中研究代数数域和代数整数的一个分支。它的形成和发展主要由于不定方程的推动, 起源于费马猜想, 高斯和库默为两位主要的开拓者, 高斯复整数理论的建立是代数数论发展方向上的一个阶段。库默把高斯的理论推广到代数数, 进而建立理想数的理论, 为代数数论的产生奠定了基础。他证明了对于 100 以内的奇素数, 除 37、59、67 外, 费马猜想都是对的, 由此将数论的研究推到一个崭新阶段, 出现了代数数论。继库默之后, 对代数数论做出重大贡献的是希尔伯特, 他的〈数论报告〉总结了直到库默、迪德金的研究成果。1900 年在第二次国际数学家大会上, 希尔伯特提出了著名的 23 个问题, 其中 9—12 属于代数数论。目前, 代数数论已成为数学中内容异常丰富的分支, 是研究不定方程的主要工具。代数数论越来越深入地运用到计算机科学、信息科学等应用学科中。

【解析数论】以分析方法为研究工具的数论分支, 18 世纪欧拉奠基, 起源于素数的分布, 哥德巴赫猜想等著名问题。解析数论是解决数论中艰深问题的强有力工具, 陈景润正是运用了深刻的解析方法, 对筛法做了重要改进, 从而证明了哥德巴赫猜想的 $(1, 2)$ 。解

析数论的方法主要有复变积分法、圆法、筛法及指数和方法等。在数论中有这样一些问题,尽管问题本身不涉及分析概念,但若不应应用分析方法,就解决不了,如华林问题等,也有些问题可以不利用分析方法,但若应用分析方法证明,结果可以更深刻,或证明更简单,如整数分拆,当然也有的问题本身必须用分析方法才能表达清楚。

【概率数论】始于 1917 年哈代等人关于数论函数 $\omega(n)$ 的研究,这里 $\omega(n)$ 表示 n 的不同的素因数的个数;坐标为整数的点称整点,全部整点构成空间格网,以空间格网为基本研究对象的学科称几何数论,它由德国数学家闵可夫斯基等人开创;堆垒数论又称加性数论,是数论的一个分支,整数的分拆是它的一个基本问题,如平方和问题、哥德巴赫猜想问题等,是堆垒数论研究的著名问题;超越数论是目前活跃的一个数论分支,它是以超越数为研究对象的数论分支,若复数 α 是某个有理系数多项式的根,就称 α 为代数数,否则 α 称超越数。

数和多项式

【数系】数是现代数学的基本概念之一。历史上它随着人类生产生活的实际需要逐步形成。若数集 B 是 A 的扩集,数的扩张原则是:(1)在 B 中, A 的元素间的基本运算及关系与在 A 中的一致;(2)某些在 A 中不可行的运算在 B 中可行;(3)同构意义下, B 是满足

上述条件的 A 的最小扩集。

自然数系的建立是在一些原始概念和公理的基础上进行的,根据所选概念与公理的不同,有不同的方法。可用基数理论及序数理论建立自然数。基数理论以原始概念“集合”为基础,称一切非空等价集合的共同特征的标志为基数,有限集合的基数称自然数。序数理论是由“集合”、“后继”两个概念及皮亚诺公理定义的:

设 $N \neq \emptyset$, 存在 $N \rightarrow N$ 的一个“后继”映射 σ , $\forall a \in N, a'$ 表示 a 的象, 称 a' 为 a 的后继, 满足以下公理: (1) N 中存在 $1, \forall a \in N, a' \neq 1$; (2) N 中每个元有且仅有一个后继; (3) $\forall a' \neq 1 \in N, \exists ! a, a \rightarrow a'$; (4) 归纳公理: $M \subseteq N$, 若 ① $1 \in M$, ② 当 $a \in M$ 时, 有 $a' \in M$, 则 $M = N$ 。此时称 N 是自然数系, N 中元素称自然数, 归纳公理是数学归纳法的理论基础。

自然数的加法, 指 $N \times N \rightarrow N$ 的映射 f , 满足 (1) $f((a, 1)) = a'$, 即 $a + 1 = a'$, (2) $f((a, b')) = f((a, b))'$, 即 $a + (b + 1) = (a + b) + 1$ 。加法满足结合律, 交换律; 自然数的乘法, 指 $N \times N \rightarrow N$ 的映射 φ , 满足 (1) $\varphi((a, 1)) = a$, 即 $a \times 1 = a$, (2) $\varphi((a, b')) = \varphi((a, b)) + a$, 即 $a(b + 1) = ab + a$, 乘法满足结合律, 交换律, 并且乘法对加法满足分配律。自然数的运算是—切数系运算的基础。

自然数可以定义大小: 对 a, b 两个自然数, 若存在自然数 $k, a = b + k$, 称 $a > b$ ($b < a$); 若存在自然数 $e, b = a + e$, 称 $b > a$ ($a < b$)。任意两个自然数 $a, b, a = b, a < b$ 和 $b < a$ 有且仅有一个成立。此谓自然数的三歧性, 并且若 $a < (>) b, b < (>) c$, 则 $a < (>) c$, 称“ $<$ ” (“ $>$ ”) 有传递性。

利用自然数及其相反数、0, 可得整数集, 它是含有自然数集的最小数环, 我们也可以由自然数集 N , 利用前面的扩张原则得整数集。

有理数集由全体整数、分数组成,它是最小的数域,记 Q ,也可以利用扩张原则得到。

无限不循环十进小数称无理数,有理数、无理数统称实数。另一种定义是称含 Q 为子域的最小的连续域为实数域,在同构意义下是唯一的,记 R 。无理数产生于度量的需要,19 世纪 70 年代,随着数学分析的发展,实数成了主要研究对象。迪德金、康托尔等人分别建立了实数理论,尽管形式不同,但彼此同构。

复数系是数系的最后一次扩充,这次扩充产生于要求 $X^2+1=0$ 有解。16 世纪,卡当在解一元三次方程时,遇到了负数开方问题,人们认为没有实际意义,不予承认,取名虚数,有虚幻之意。18 世纪后期,欧拉引入 i ,规定 $i^2=-1$ 。复数的表示法很多,常用 $a+bi$ 表示, $a, b \in R$;也利用复数与实平面上点的对应,记 $a+bi$ 为 (a, b) ;复数的三角表示形式给出了复数在平面上的位置,且乘、除法有简单公式,计算简便。若 $\alpha = r(\cos\varphi + i\sin\varphi)$, $\beta = r_1(\cos\varphi_1 + i\sin\varphi_1)$, 则 $\alpha\beta = rr_1(\cos(\varphi + \varphi_1) + i\sin(\varphi + \varphi_1))$, $\alpha/\beta = r/r_1(\cos(\varphi - \varphi_1) + i\sin(\varphi - \varphi_1))$;复数还有指数表示 $r(\cos\varphi + i\sin\varphi) = re^{i\varphi}$ 。复数集记 C ,较之以前数集之一大不同是有序性,即在 C 中无法定义“ $<$ ”,使满足(1) $\forall a, b \in C, a=b, a<b, b<a$ 有且仅有一个成立;(2)若 $a<b, b<c$ 则 $a<c$;(3)若 $a<b, \forall c$, 则 $a+c<b+c$;(4)若 $a<b, 0<c$, 则 $ac<bc$ 。

1861 年,威尔斯特拉斯证明了不存在比复数集更大的数集仍满足数的加乘运算的五算律。

【数环与数域】 S 是一个非空数集,若 $\forall a, b \in S$, 都有 $a \pm b, ab \in S$, 则称 S 是一个数环。若 F 是一个非零数环,且 $\forall a, b \in F, b \neq 0$ 有 $a/b \in F$, 称 F 是一个数域。最小的数域为 Q , 且任何数域均包含

Q, C 是最大数域。数域有无穷多个。

【四元数】 $a, b, c, d \in R$, 形如 $a + bi + cj + dk$ 的数。设 $\alpha = a + bi + cj + dk, \beta = a' + b'i + c'j + d'k$, 则 $\alpha = \beta \Leftrightarrow a = a', b = b', c = c', d = d'$; $\alpha + \beta = (a + a') + (b + b')i + (c + c')j + (d + d')k$; α 与 β 相乘时, 利用分配律及 $i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j$, 这表明乘法不满足交换律。全体四元数记作 H , 对上述加、乘作成除环(参见“环与代数”)。四元数由哈密顿发现, 对代数学具有不可估量的重要性, 在此之前扩充数的过程中, 运算都满足 5 个算律, 现在发现了乘法不可交换的数系, 因此人们思想大为解放, n 元超复数相继出现。四元数的发现, 推动了线性代数对向量的研究。四元数除环还可以表为

$M_2(C)$ 的子环 H_1 ,

$$H_1 = \left\{ \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} \mid \alpha, \beta \in C \right\},$$

也可表为四元实数组集 H_2 ,

$$H_2 = \{(a_1 \ a_2 \ a_3 \ a_4) \mid a_i \in R\},$$

此时若 $\alpha = (a_1 \ a_2 \ a_3 \ a_4), \beta = (b_1 \ b_2 \ b_3 \ b_4) \in H_2, \alpha + \beta = (a_1 + b_1 \ a_2 + b_2 \ a_3 + b_3 \ a_4 + b_4), \alpha\beta = (a_1b_1 - a_2b_2 - a_3b_3 + a_4b_4 \ a_2b_1 + a_1b_2 - a_4b_3 - a_3b_4 \ a_1b_3 + a_2b_4 + a_3b_1 + a_4b_2 \ a_2b_3 + a_1b_4 - a_3b_2 - a_4b_1)$, 且 H, H_1, H_2 是同构的。

【进位制与换算】通常使用十进制, 逢十进一, 但也有其他进制, 如 1 小时 = 60 分, 1 分 = 60 秒, 即六十进制, 在计算机上则多用二进制, 这是由于二进制在电气元件中容易实现, 且运算也简单。

一个正数都可表为有限或无限小数, 各数字的值与所在位置

有关。若 a 是一个正数, $a = a_n a_{n-1} \cdots a_1 a_0 \cdot a_{-1} a_{-2} \cdots, a_i, a_{-j} \in \{0, 1, \cdots, 9\}$, a 可改写为 $a = a_n \times 10^n + a_{n-1} \times 10^{n-1} + \cdots + a_0 \times 10^0 + a_{-1} \times 10^{-1} + \cdots$ 这是按十进制写的, 记为 $a_{(10)}$ 。

不同的进位制之间可以互相换算, 对于十进制与二进制, 若 $a_{(10)} = a_0 \times 2^n + a_1 \times 2^{n-1} + \cdots + a_n \times 2^0$, 则 $a_{(10)} = (a_0 a_1 \cdots a_n)_{(2)}$ 。于是对于 $a_{(10)}$, 用 2 累次去除, 将所得余数从后往前书写, 即为 $a_{(2)}$ 。例 $a_{(10)} = 19, b_{(10)} = 14$, 可得 $19_{(10)} = (10111)_{(2)}, 14_{(10)} = (1110)_{(2)}$ 。若 $\alpha = (a_0 a_1 \cdots a_n)_{(2)}$, 则 $\alpha_{(10)} = a_0 \times 2^n + a_1 \times 2^{n-1} + \cdots + a_n$ 。

一般可引入 q 进制, 常用的 q 为 10, 2, 8, 16 等等, 若 $a_{(q)} = a_n a_{n-1} \cdots a_0$, 则 $a_{(10)} = a_n \times q^n + \cdots + a_0$ 。

【一元多项式】 F 是数域, X 是文字, 形式表达式 $a_0 + a_1 X + \cdots +$

$a_n X^n = \sum_{i=0}^n a_i X^i$ 称为 F 上 X 的多项式, 常记

$f(X), g(X), \cdots$, 其中 n 是非负整数, $a_i \in F$ 。

若 $f(X) = \sum_{i=0}^n a_i X^i$,

且 $a_n \neq 0$, 称 $f(X)$ 为 n 次多项式, 记 $\partial(f(X)) = n, a_i X^i$ 称 i 次项, 两个多项式仅当它们至多差一些 oX^i 项时被认为是相等的, 若 $f(X)$

$= \sum_{i=0}^n a_i X^i, g(X) = \sum_{i=0}^m b_i X^i (n \leq m)$, 称

$\sum_{i=0}^n (a_i + b_i) X^i + \sum_{j=n+1}^m b_j X^j$ 为 $f(X)$ 与 $g(X)$ 的和, 记 $f(X) + g(X)$ 。令

$c_K = \sum_{i+j=K} a_i b_j, 0 \leq K \leq n+m$, 称 $\sum_{K=0}^{n+m} c_K X^K$ 为 $f(X)$ 与 $g(X)$ 的积, 记 f

$(X)g(X)$ 。 $F[X]$ 表示 F 上 X 的全体多项式的集合。 $F[X]$ 对上面的加法、乘法做成环(参见“环与代数”), 称为 F 上 X 的一元多项

式环。作为代数系, $F[X]$ 与 Z 都是唯一分解整环(参见“环与代数”)。对 $F[X]$ 可以讨论多项式的整除, 得到与算术基本定理平行的唯一分解定理: F 上 X 的任一个次数大于零的多项式可以分解成不可约因式之积, 且分法本质唯一, 不可约多项式在 $F[X]$ 中作用与素数在 Z 中作用相当(参见“初等数论”)。

若 $f(X) = \sum_{i=0}^n a_i X^i \in F[X], c \in F$, 称 $\sum_{i=0}^n a_i c^i$ 为 $X=c$ 时 $f(X)$ 的值, 记

$f(c)$, 若 $f(c)=0$, 称 c 是 $f(X)$ 的一个根, 也称 c 是方程 $f(X)=0$ 的一个解。求多项式的根可用综合除法: 设 $f(X)=a_0 X^n + a_1 X^{n-1} + \cdots + a_{n-1} X + a_n$, $f(X)=(X-c)q(X)+r$, 其中 $q(X)=b_0 X^{n-1} + b_1 X^{n-2} + \cdots + b_{n-1}$, 比较等式两端可得:

$b_0=a_0, b_1=cb_0+a_1, b_2=cb_1+a_2, \cdots, b_{n-1}=cb_{n-2}+a_{n-1}, r=cb_{n-1}+a_n$, 通常利用如下格式求 b_i :

$$\begin{array}{r|rrrrrr}
 c & a_0 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\
 + & & cb_0 & cb_1 & \cdots & cb_{n-2} & cb_{n-1} \\
 \hline
 & a_0=b_0 & b_1 & b_2 & \cdots & b_{n-1} & r
 \end{array}$$

c 是 $f(X)$ 的根 $\Leftrightarrow r=0$ 。

代数基本定理指出, 任一个数系数的 $n(>0)$ 次多项式在复数域 C 中至少有一个根, 从而任一个数系数的 n 次多项式在 C 中有 n 个根(k 重根按 k 个计), 在寻求 5 次及 5 次以上方程解的过程中, 这个定理首先被高斯证明(1799 年), 他的证明方法对数学中有关存在性的证明给了新的思路, 意义深远。利用代数基本定理不难得出, C 上不可约多项式均为一次的, 因而每一个复系数多项式的典型分解式为 $a(X-c_1)^{k_1} \cdots (X-c_t)^{k_t}$ 形式。

实系数多项式非实复根成对出现, 即若 $\alpha=a+bi$ 是实 $f(X)$ 的

根,则 $\bar{\alpha}=a-bi$ 也是 $f(X)$ 的根。换言之, R 上不可约多项式只有一次或有一对共轭复根的二次的,一个 2 次实多项式不可约 $\Leftrightarrow$ 判别式 $D < 0$ 。因而实多项式的典型分解式为

$a(X-c_1)^{k_1} \cdots (X-c_s)^{k_s} (X^2+a_1X+b_1)^{l_1} \cdots (X^2+a_tX+b_t)^{l_t}$, 其中 $X^2+a_iX+b_i$ 在 R 上不可约。

有理数域 Q 上多项式的情况要复杂,但 Q 上多项式的可约问题可归结为 Z 上多项式的可约问题,且 Q 上多项式的求根问题可归结为 Z 上多项式的求根问题。艾森施坦因判别法指出:对于整系数多项式

$f(X)=a_0+a_1X+\cdots+a_nX^n$, 若存在素数 p , 使得 (1) $p \nmid a_n$, (2) $p \mid a_i$, $i \neq n$; (3) $p^2 \nmid a_0$, 则 $f(X)$ 在 Q 上不可约。这样,对任意 n , X^n+2 均为 Q 上不可约多项式,但这个条件不是必要的。若整系数多项式 $f(X)=a_0X^n+\cdots+a_n$ 有有理根 $\alpha=\frac{v}{u}$, 则 $u \mid a_n$, $v \mid a_0$, 特别,首 1 的整系数多项式的有理根均为整数,有理多项式的无理根成对出现,即若

$a+b\sqrt{p}$ 是有理多项式 $f(X)$ 的根,则 $a-b\sqrt{p}$ 也是。

若 $f(X)=X^n+a_1X^{n-1}+\cdots+a_n$ 是 F 上多项式,在 C 中有根 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则根与系数间有如下关系:

$$\begin{aligned}\alpha_1 + \cdots + \alpha_n &= -a_1 \\ \alpha_1\alpha_2 + \cdots + \alpha_{n-1}\alpha_n &= a_2 \\ \alpha_1 \cdots \alpha_n &= (-1)^n a_n\end{aligned}$$

也称根与系数的韦达定理。

实际问题中常遇到实系数多项式实根的近似计算,笛卡儿符号法则指出,若 $f(X)=\sum_{i=0}^n a_i X^i$ 中系数 $a_0, a_1, \cdots, a_n$ 的变号数为 s , 则 $f(X)$ 正实根个数或为 s 或比 s 少一个正偶数,但无法判定 $f(X)$

是否有正实根。另一个行之有效的方法由施图姆给出:令 $f_0 = f(X)$, $f_1 = f'(X)$, 作序列 $f_0, f_1, \dots, f_s$ 使得

$$\begin{aligned} f_0 &= f_1 q_1 - f_2 \\ f_1 &= f_2 q_2 - f_3 \quad (\text{其中 } -f_i \text{ 为 } f_{i-1} \text{ 除 } f_{i-2} \text{ 所得余式}) \\ &\vdots \\ f_{s-2} &= f_{s-1} q_{s-1} - f_s \\ f_{s-1} &= f_s q_s \end{aligned}$$

称 $f_0, f_1, \dots, f_s$ 为 $f(X)$ 的施图姆组, $a < b$ 且 $f(a) \neq 0, f(b) \neq 0$ 时, $f(X)$ 在 $[a, b]$ 根的个数等于施图姆组的变号数 $v(a) - v(b)$, 这样, 当 $f(X)$ 有实根时, 可求得若干区间, 在每一区间, $f(X)$ 恰有一个根, 利用秦九韶法、牛顿法、二分法、线性插值法等可求得满足一定精确度的根的近似值。

对于给定数域 F , n 次多项式 $f(X)$ 在 F 中至多有 n 个根 (k 重根按 k 个计), 由此, 若 $f(X), g(X) \in F[X]$, 次数不超过 n , 且存在多于 n 个 $c_i \in F$, 使 $f(c_i) = g(c_i)$, 那么 $f(X)$ 与 $g(X)$ 是两个相等的多项式, 由此, 对 F 中任意 $n+1$ 个不同的数 a_i 及任意 $n+1$ 个数 $b_i (i=1, \dots, n+1)$, 拉格朗日给出了插值公式:

$f(X) =$

$$\sum_{i=1}^{n+1} \frac{b_i (X-a_1) \cdots (X-a_{i-1})(X-a_{i+1}) \cdots (X-a_{n+1})}{(a_i-a_1) \cdots (a_i-a_{i-1})(a_i-a_{i+1}) \cdots (a_i-a_{n+1})}$$

满足 $f(a_i) = b_i, 1 \leq i \leq n+1$ 。利用这个公式及已知数据, 可求得一个近似的多项式, 在实际问题中有一定用途。

在初等代数中, 多项式常看成是函数, 因为 $f(X) = g(X) \Leftrightarrow \forall c \in R, f(c) = g(c)$, 但当考虑一般域上的多项式时, 多项式与多项式函数未必一致, 所以在一般域上讨论时, 多项式定义为形式表达式。

【多元多项式】 F 上 $X_1, \dots, X_n$ 的形式表达式称 F 上一个 n 元多项式, 常记 $\sum_{i_1, \dots, i_n} a_{i_1, i_2, \dots, i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$, 简记 $f(X_1, \dots, X_n), g(X_1, \dots, X_n)$ 或 $f, g, \dots$ 。 $a_{i_1, i_2, \dots, i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ 称 $f(X_1, X_2, \dots, X_n)$ 的一项, $a_{i_1, i_2, \dots, i_n}$ 称该项的系数, $i_1 + i_2 + \dots + i_n$ 称该项的次数, 若 $f(X_1, X_2, \dots, X_n) \neq 0$, 它的各项次数的最大者称这个多项式的次数, 记为

$$\partial(f(X_1, X_2, \dots, X_n)),$$

若 $f(X_1, X_2, \dots, X_n)$ 的每项次数均为 k , 称 $f(X_1, X_2, \dots, X_n)$ 为 k 次齐次多项式或 k 次齐式。两个 k 次齐式之和或为零或为 k 次齐式, 两个齐式之积仍为齐式, 其次数为两个齐式次数之和。

对于项

$$aX_1^{k_1}X_2^{k_2}\cdots X_n^{k_n} \quad (1),$$

$$bX_1^{l_1}X_2^{l_2}\cdots X_r^{l_r} \quad (2),$$

若

$$k_1=l_1, k_2=l_2, \dots, k_{r-1}=l_{r-1}, k_r>l_r$$

称项(1)先于项(2), 这种排列多项式的方法称多项式的字典排法, 是多元多项式常用的表示法, 每个多项式按字典排法有唯一的表达式, 排在第一项的称多项式的首项, 两个多项式乘积的首项等于各多项式的首项之乘积, 对于一元多项式, 字典排法即降幂排法, 首项就是最高次项, 对于 $n>1$ 时的 n 元多项式, 首项未必是最高次项, 如 $f(X_1, X_2)=X_1+X_2^3$, 首项是 X_1 , 最高次项是 X_2^3 。

多元多项式中最重要的是对称多项式。对任一个 n 元排列 $i_1, i_2, \dots, i_n, f(X_1, X_2, \dots, X_n)=f(X_{i_1}, X_{i_2}, \dots, X_{i_n})$, 称 $f(X_1, X_2, \dots, X_n)$ 为 n 元对称多项式。

$$\sigma_1=X_1+X_2+\cdots+X_n$$

$$\sigma_2=X_1X_2+X_1X_3+\cdots+X_{n-1}X_n$$

$$\vdots$$

$$\sigma_n = X_1 X_2 \cdots X_n$$

称作初等对称多项式, 每一个 n 元对称多项式均可以唯一地表为初等对称多项式的多项式。这个事实称对称多项式的基本定理。

若 $f(X) = a_0 X^n + a_1 X^{n-1} + \cdots + a_n$ 在 C 中的 n 个根为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 称

$$a_0^{2n-2} \prod_{n \geq j > i \geq 1} (\alpha_j - \alpha_i)^2$$

为 $f(X)$ 的判别式, 记作 D_f , 它是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的对称多项式。

$f(X)$ 有重根当且仅当 $D_f = 0$ 。

利用结式与判别式的关系可以由 $f(X)$ 的系数表示 D_f , 若 $f(X) = a_0 X^n + a_1 X^{n-1} + \cdots + a_n$, $g(X) = b_0 X^m + b_1 X^{m-1} + \cdots + b_m$, 称 $m+n$ 阶行列式

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & \cdots & a_n & & & & \\ & a_0 & a_1 & a_2 & \cdots & \cdots & a_n & & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \\ & & & a_0 & a_1 & a_2 & \cdots & \cdots & a_n & \\ b_0 & b_1 & b_2 & \cdots & b_m & & & & & \\ & b_0 & b_1 & b_2 & \cdots & b_m & & & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & & & \\ & & & b_0 & b_1 & b_2 & \cdots & & & b_m \end{vmatrix}$$

为 $f(X), \nabla g(X)$ 的结式, 记 $R(f, g)$, 有

$$R(f, g) = (-1)^{mn} R(g, f)$$

$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_0 D$, 从而

$$D = (-1)^{\frac{n(n-1)}{2}} \frac{1}{a_0} R(f, f')$$

一元二次多项式 $f(X) = aX^2 + bX + c$ 的判别式 $D = b^2 - 4ac$ 即由此而得。

线 性 代 数

【线性代数】代数学的基础分支。以行列式,矩阵,线性方程组,向量空间,线性变换,二次型等为主要研究内容。对线性方程组,线性变换,二次型的讨论,均借助它们的矩阵表示,转化为矩阵问题讨论。线性代数也称一次代数,在数学各分支及众多学科都有广泛的应用。

【行列式】重要的数学概念和工具之一。用符号

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |a_{ij}|_n$$

表示的一个数,它的值为

$\sum_{j_1 j_2 \cdots j_n} (-1)^s a_{1j_1} a_{2j_2} \cdots a_{nj_n}$, 其中 $j_1 j_2 \cdots j_n$ 是 n 元排列, s 是排列 $j_1 j_2 \cdots j_n$ 的

反序数, $\sum_{j_1 j_2 \cdots j_n}$ 表示对一切 n 元排列求和,共 $n!$ 项。

行列式的主要性质有:(1)行列式等于它的转置行列式,它保证凡对行成立的事实对列均成立;(2)交换两行(列),行列式改变符号;(3)某一行(列)乘以一个非零常数 k ,其值为原行列式的 k 倍;(4)某一行(列)乘一个数后加到别一行(列)上,行列式值不变;(5)每一个行列式都可化简为上(下)三角形行列式,三角形行列式

之值为对角元之积。

在 n 阶行列式 $|a_{ij}|_n$ 中,划掉 a_{ij} 所在的行、列所剩 $n-1$ 阶行列式称 a_{ij} 的余子式,冠以符号 $(-1)^{i+j}$ 后称 a_{ij} 的代数余子式,记 A_{ij} 。有以下重要结果:(1)行列式依行(列)展开规则:

$$D = \sum_{k=1}^n a_{ik} A_{ik}, D = \sum_{k=1}^n a_{kj} A_{kj};$$

(2) $\sum_{k=1}^n a_{ik} A_{jk} = 0, \sum_{k=1}^n a_{ki} A_{kj} = 0 (i \neq j)$ 。行列式还有依若干行(列)展开的拉普拉斯定理。

日本数学家关孝和和德国数学家莱布尼兹最早引入行列式概念,法国数学家范德蒙是第一个对行列式理论做出连贯阐述的人,柯西首先将行列式元素采用双足标记法并排成方阵。

【矩阵】重要的数学概念和工具之一。由 $m \times n$ 个数排成的数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

叫 m 行 n 列矩阵,也称 $m \times n$ 矩阵,简记为 $A = (a_{ij})_{mn}$,当 $m = n$ 时,称 A 为 n 阶方阵。这里 a_{ij} 取自给定数域。

矩阵的秩 矩阵中非零子式的最大阶数。零矩阵的秩为 0, A 的秩记

$r(A), r(A_{mn}) \leq \min(m, n)$,若 P 是可逆方阵, $r(PA) = r(A)$ 。

矩阵的初等变换: (1)交换矩阵的两行(列); (2)用一个非零数乘矩阵的某一行(列); (3)用一个数乘矩阵的某一行(列)后加到另一行(列)上。初等变换不改变矩阵的秩。

矩阵的运算 若 $A = (a_{ij})_{mn}, B = (b_{ij})_{mn}$, 称 $(a_{ij} + b_{ij})_{mn}$ 为 A 与

B 的和,记 $A+B$,这种运算称矩阵的加法;称 $(ka_{ij})_{mn}$ 为 k 与 A 之积,记 kA ,这种运算称数与矩阵的乘法;若 $A=(a_{ij})_{mt}$, $B=(b_{ij})_{tn}$,

$$\text{令 } c_{ij} = \sum_{k=1}^t a_{ik} b_{kj},$$

$1 \leq i \leq m, 1 \leq j \leq n$,称 $C=(c_{ij})_{mn}$ 为 A 与 B 之积,记 AB ,这种运算称矩阵的乘法,矩阵运算满足以下算律:加法结合律、交换律;乘法结合律;乘法对加法分配律;数乘对加法分配律;数乘对乘法结合律。矩阵的乘法不满足交换律,如 $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$,两个非零矩阵之积可能为零,如 $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$,因此矩阵乘法不满足消去律。数域 F 上全体 n 阶方阵之集合 $M_n(F)$ 对于上述加法、乘法做成非交换么环。对上述三个运算做成代数(参见“环与代数”)。

一些特殊方阵 将一个矩阵的行列互换所得矩阵称原矩阵之转置, A 的转置记为 A' ,且 $(AB)' = B' A'$;若方阵 $A, A' = A$, A 称对称方阵;若一个方阵 A 对角线以外元素均为 0,称 A 为对角阵;若对角阵 A 中, $a_{11} = \cdots = a_{nn} = a$,称 A 为纯量阵; $a=1$ 的纯量阵称单位阵,记 I 。对于方阵 A ,若存在方阵 $B, AB=BA=I, A$ 称可逆方阵, B 称 A 的逆方阵,可逆方阵 A 有唯一的逆方阵,记 A^{-1} ,且 $(A^{-1})^{-1} = A$. A 可逆 $\Leftrightarrow |A| \neq 0 \Leftrightarrow A$ 经初等变换得 I 。若 A, B 可逆,则 AB 可逆,且 $(AB)^{-1} = B^{-1} A^{-1}$,若 A 是实方阵,且 $A^{-1} = A'$,称 A 为正交方阵。初等方阵是指换位阵 P_{ij} ,倍法阵 $D_i(k), k \neq 0$,消法阵 $T_{ij}(l)$ 。其中

$$P_{ij} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & 1 \cdots 0 & & \\ & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix},$$

$$D_{i(r)} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & k & \\ 0 & & & \ddots & \\ & & & & 1 \end{pmatrix} \quad k \neq 0,$$

$$T_{ij(l)} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & 1 \\ & & & \ddots & \\ 0 & & & & 1 & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

用初等方阵左(右)乘矩阵 A , 相当对 A 施行相应的行(列)初等变换。若 $A = (a_{ij})_n$, $A^* = (A_{ij})'$ 叫 A 的伴随方阵, 其中 A_{ij} 为 $|A|$ 中 a_{ij} 的代数余子式, 若 A 可逆, 则 $A^{-1} = \frac{1}{|A|} A^*$ 。也可以利用初等变换求可逆方阵的逆: 设 A 可逆, 记 $B = (A | I)$, 对 B 作行初等变换, 使 A 化为 I , 此时 I 就化为 A^{-1} , 即 $B = (A | I) \xrightarrow{\text{行}} (I | A^{-1})$ 。

矩阵的等价 若 A 通过初等变换为 B , 称 A, B 等价。 A 与 B 等价 $\Leftrightarrow$ 存在可逆阵 P, Q 使 $B = PAQ$ 。 $\Leftrightarrow r(A) = r(B)$ 。

矩阵的相似 F 上 n 阶方阵 A, B , 若存在可逆阵 T , 使 $T^{-1}AT = B$, 称 B 与 A 相似。若 A 与对角阵相似, 称 A 可对角化。并且 $XI - A$ 为 A 的特征方阵。 $f_A(X) = |XI - A|$ 称 A 的特征多项式, $f_A(X)$ 在 F 中的根称 A 的特征根, 相似方阵有相同的特征多项式。若 λ_0 是 A 的特征根, 齐次方程组 $(\lambda_0 I - A)X = 0$ 的非零解称 A 的属于 λ_0 的特征向量。 A 可对角化 $\Leftrightarrow A$ 有 n 个无关的特征向量。此时 T 的列由特征向量组成, 对角阵的对角元为对应特征根。若

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_t \end{pmatrix},$$

$$J_i = \begin{pmatrix} \lambda_i & & 0 \\ 1 & \ddots & \\ & \ddots & \ddots & \\ & & \ddots & \lambda_i \\ & & & 1 \end{pmatrix} \quad i = 1, \dots, t$$

称 J_i 为一个若当块, J 称若当阵, 每一个复方阵与一个若当阵相似。

矩阵的合同 F 上 n 阶方阵 A, B , 若存在可逆阵 $P, P'AP=B$, 称 B 与 A 合同, 每一个对称阵与对角阵合同, 可以通过初等变换实现。秩为 r 的 n 阶复对称方阵与 $\begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$ 合同, 因此, 复对称方阵 A, B 合同 $\Leftrightarrow r(A)=r(B)$ 。每一个秩为 r 的 n 阶实对称方阵 A 与

$\begin{pmatrix} I_p & & \\ & -I_{r-p} & \\ & & 0 \end{pmatrix}$ 合同, p 由 A 唯一确定, 并且对实对称方阵 A , 存在正

交方阵 $U: U^{-1}AU=U'AU=\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, λ_i 是 A 的特征根。

矩阵的等价, 相似, 合同都是等价关系, 具有反身性, 对称性, 传递性。

若 $h(X)=c_0+c_1X+\cdots+c_tX^t$, 对方阵 A , 称 $h(A)=c_0I+c_1A+\cdots+c_tA^t$ 为 $X=A$ 时 $h(X)$ 之值。若 $h(A)=0$, 称 A 为 $h(X)$ 的根。哈密顿-凯莱定理说, A 是它的特征多项式的根, 即 $f_A(A)=0$ 。

矩阵最早出现在《九章算术·方程章》中, 1855 年凯莱提出矩阵概念, 1858 年, 他在《矩阵论的研究报告》中, 系统地阐述了矩阵的理论, 这些内容是今天线性代数的基础内容。近几十年来, 矩阵论被广泛应用于现代科学的众多领域, 内容更加丰富, 更加广泛, 不限于有限阶矩阵。矩阵分析、矩阵函数等是当今讨论的新课题。

【线性方程组】 n 个未知量 m 个方程的线性方程组的一般形式为

$$\begin{cases} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n = b_1 \\ \quad \quad \quad \cdots \quad \quad \quad \cdots \\ a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mn}X_n = b_m \end{cases} \quad (1)$$

其中 a_{ij} 称系数, b_i 称常数项, a_{ij}, b_i 取自一个数域 F 。若 $\sum_{j=1}^n a_{ij}c_j = b_i$ $1 \leq i \leq m$, 称 $(c_1, \cdots, c_n)$, 为(1)的一个解, 有解的方程组称相容的。线性方程组研究的主要问题是: 1. (1)相容的充要条件; 2. (1)相容时, 解的个数; 3. (1)相容时, 求出(1)的所有解

令

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

称 A 为(1)的系数矩阵, $\bar{A}$ 为(1)的增广矩阵, 结论是: 1. (1)相容 $\Leftrightarrow r(A) = r(\bar{A})$; 2. 若 $r(A) = r(\bar{A}) = r$, 则 $r = n$ 时, (1)有唯一解, $r < n$ 时, (1)有无穷多解。3. 利用消元法可求(1)的解。

消元法实际是对方程组进行同解变换, 通过消去某些未知量化简方程而求解。这些变换主要有: 用一个非零数去乘某一个方程; 用一个数乘某一个方程后加到另一方程上; 交换两个方程的次序。中国早在《九章算术·方程章》中已记载了与现在基本一致的消元法, 这个方法是中国古代数学的杰出创造之一, 它比欧洲大约早 1500 年。

在(1)中, 若 $b_1 = \cdots = b_m = 0$, 称(1)为齐次(线性)方程组, 它永远有零解 $(0, 0, \cdots, 0)$, 一个齐次方程组有非零解 $\Leftrightarrow r(A) < n$;

一个齐次方程组全体解的集合做成 F^n 的一个子空间,称为这个齐次方程组的解空间,它的基常称为该方程组的基础解系。

称齐次方程组

$$\begin{cases} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n = 0 \\ \cdots \quad \cdots \quad \cdots \\ a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mn}X_n = 0 \end{cases} \quad (2)$$

为(1)的导出方程组,若 $\alpha = (c_1, c_2, \cdots, c_n)$ 是(1)的一个解, W 为(2)的解空间,则(1)的解集为 $\alpha + W = \{\alpha + \beta \mid \forall \beta \in W\}$ 。

1750 年瑞士的克莱姆给出一类特殊线性方程组解的公式,称为克莱姆规则:若线性方程组

$$\begin{cases} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n = b_1 \\ \cdots \quad \cdots \quad \cdots \\ a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nn}X_n = b_n \end{cases} \quad (3)$$

的系数行列式 $D \neq 0$, 则(3)有且仅有一组解: $X_1 = D_1/D, X_2 = D_2/D, \cdots, X_n = D_n/D$. 其中 D_j 为以常数项 $b_1 \cdots b_n$ 代替 D 的第 j 列所得行列式。

有着广泛应用的线性规划实际是解满足一定约束条件的线性方程组。

【向量空间】解析几何中普通空间 V_3 , 普通平面 V_2 的推广, 利用向量的坐标, V_2, V_3 可以看成二元实数组及三元实数组的集合 R^2, R^3 , 它们的加法、数乘法即矩阵的加法和数乘法。可以自然地推广到数域 F 的 n 元序组的集合 F^n , 称 F^n 对于矩阵的加法及数乘法为一个 n 维空间。一般, 若 F 是数域, V 是一个非空集合, V 有一个加法及 F 对 V 的数乘法, 满足: (1) V 对加法是一个加群; (2) $(a+b)\alpha = a\alpha + b\alpha$; (3) $a(\alpha + \beta) = a\alpha + a\beta$; (4) $(ab)\alpha = a(b\alpha)$; (5) $1\alpha =$

$\alpha; \forall a, b \in F, \alpha, \beta \in V$, 称 V 是 F 上向量空间。 V 中元素统称向量。 $F[X], M_n(F)$ 对各自的运算都是向量空间, 定义在 $[a, b]$ 上的实连续函数的集合 $C_{[a, b]}$ 对函数的运算做成向量空间。

向量空间的基与维数 若 V 是 F 上向量空间, $\alpha_1, \dots, \alpha_s \in V$, 存在 F 中不全为 0 的 $a_1, \dots, a_s$ 使 $\sum_{i=1}^s a_i \alpha_i = 0$, 称 $\alpha_1, \dots, \alpha_s$ 线性相关, 否则称 $\alpha_1, \dots, \alpha_s$ 线性无关。若 $\beta = \sum_{i=1}^s b_i \alpha_i$, 称 β 可由 $\alpha_1, \dots, \alpha_s$ 线性表示。矩阵经过行初等变换不改变列向量的线性关系。即若 $A = (\alpha_1 \cdots \alpha_n) \xrightarrow{\text{行}} B = (\beta_1 \cdots \beta_n)$, 则 $\sum_{i=1}^s k_i \alpha_i = 0 \Leftrightarrow \sum_{i=1}^n k_i \beta_i = 0$ 。若 V 中 $\alpha_1, \dots, \alpha_n$ 线性无关, 且 V 中每一向量都可由 $\alpha_1, \dots, \alpha_n$ 线性表示, 称 $\alpha_1, \dots, \alpha_n$ 为 V 的一个基, n 称 V 的维数, 记 $\dim V = n$, 规定零空间维数为 0, V 中每个向量可由基唯一地表示, 若 $\alpha = \sum_{i=1}^n a_i \alpha_i$, 称 $(a_1, \dots, a_n)$ 为 α 关于基 $\alpha_1, \dots, \alpha_n$ 的坐标。

子空间 子空间是向量空间 V 的非空子集, 且对 V 的加法及 F 对 V 的数乘做成向量空间。若 W_1, W_2 是 V 的子空间, 则 $W_1 \cap W_2, W_1 + W_2 = \{W_1 + W_2 \mid W_i \in W_i\}$ 是 V 的子空间, 称 W_1 与 W_2 的交, 和, 它们仍是 V 的子空间, 若 $\dim V = n$, 则 $\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2$ 。若 $V = W_1 + W_2, W_1 \cap W_2 = \{0\}$, 称 V 是 W_1, W_2 的直和, 记 $V = W_1 \oplus W_2$, 称 W_1, W_2 互为余子空间。若 $V = W_1 \oplus W_2$, 则 V 中每一向量 α 可唯一地表为 $\alpha_1 + \alpha_2$, 其中 $\alpha_i \in W_i$, 若 $\dim V = n$, 则 $\dim V = \dim W_1 + \dim W_2$ 。直和的概念可以推广到多个子空间, 若 $\alpha_1, \dots, \alpha_n$ 是 V 的基, 则 $V = W_1 \oplus \cdots \oplus W_i, W_i = L(\alpha_i) = \{a\alpha_i \mid a \in F\}$

向量空间的同构 φ 是 V_1 到 V_2 的双射, 且 $\varphi(\alpha + \beta) = \varphi(\alpha)$

$+\varphi(\beta)$,

$\varphi(a\alpha)=a\varphi(\alpha), \forall \alpha, \beta \in V_1, a \in F$, 称 φ 是 V_1 到 V_2 的同构映射。记

$V_1 \stackrel{\varphi}{\simeq} V_2$, 两个同构的空间有相同维数, 且若 $\dim V = n$, 则 $V \simeq F^n$ 。

【欧氏空间】在 R^n 中, 任意 $\alpha = (a_1, \dots, a_n), \beta = (b_1, \dots, b_n)$, 定义 $(\alpha,$

$\beta) = \sum_{i=1}^n a_i b_i$ 为 α 与 β 的内积, 是 $R^n \times R^n \rightarrow R$ 的映射。对于加法, 数

乘, 内积三种运算, R^n 叫一个欧氏空间, 可以类似于普通平面, 普

通空间, 定义向量的长及向量的夹角。在 R 上向量空间 V 中, 定义

内积 $(\cdot, \cdot)$, 使满足: (1) $(\alpha, \beta) = (\beta, \alpha)$, (2) $(\alpha + \beta, r) = (\alpha, r) + (\beta,$

$r)$, (3) $(a\alpha, \beta) = a(\alpha, \beta)$, (4) $\alpha \neq 0$ 时, $(\alpha, \alpha) > 0$, 称 V 是一个欧氏

空间, $C_{[a,b]}$ 对于

$(f(X), g(X)) = \int_a^b f(X)g(X)dX$ 是一个欧氏空间。称 $\sqrt{(\alpha, \alpha)}$ 为 α 之长,

记 $|\alpha|$, α, β 间夹角 $\angle \alpha \beta = \cos^{-1} \frac{(\alpha, \beta)}{|\alpha||\beta|} \quad \alpha \neq 0, \beta \neq 0$, 当 $(\alpha, \beta) = 0$ 时,

称 α, β 垂直(正交), 若 $\alpha_1, \dots, \alpha_n$ 是 V 的基, 且

$(\alpha_i, \alpha_j) = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$, 称 $\alpha_1, \dots, \alpha_n$ 为标准正交基, 每个 n 维欧氏空间

V 存在标准正交基, 可由施密特方法得到: 若 $\alpha_1, \dots, \alpha_n$ 是 V 的一

个基, 令 $\beta_1 = \alpha_1, \beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1, \dots, \beta_n = \alpha_n - \frac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \dots -$

$\frac{(\alpha_n, \beta_{n-1})}{(\beta_{n-1}, \beta_{n-1})}\beta_{n-1}$, 再令 $\gamma_j = \frac{\beta_j}{|\beta_j|}$, 则 $\gamma_1, \dots, \gamma_n$ 是 V 的标准正交基, 若

φ 是欧氏空间 V_1 到 V_2 的双射, 且保持加法, 数乘及内积运算:

$\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta), \varphi(a\alpha) = a\varphi(\alpha), (\alpha, \beta) = (\varphi(\alpha), \varphi(\beta))$, 称 φ

是欧氏空间 V_1 到 V_2 的同构映射,两个欧氏空间 V_1 与 V_2 同构 $\Leftrightarrow \dim V_1 = \dim V_2$ 。

正交补 若 W 是欧氏空间 V 的子空间,令 $W^\perp = \{\alpha \in V \mid (\alpha, W) = 0\}$,则 $W^\perp$ 是 V 的子空间,称 W 之正交补,且 $(W^\perp)^\perp = W$ 。若 V 有限维,则 $V = W \oplus W^\perp$,若 $W_1 \subseteq W_2$,则 $W_1^\perp \supseteq W_2^\perp$ 。普通空间 V_2, V_3 中关于距离的性质在一般欧氏空间仍成立。

欧氏空间的定义可以推广,若基础域换成复数域, C 上向量空间 V 对于适当的内积做成的空间称酉空间;若基础域换成一般数域 F , F 上向量空间 V 对于适当的内积做成的空间称内积空间。

【线性变换】向量空间到自身的一种特殊映射。有限维空间的线性变换可通过矩阵讨论。若 V 是 F 上向量空间, σ 是 V 到 V 的映射,且保向量空间运算,即 $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$, $\sigma(a\alpha) = a\sigma(\alpha)$; $\forall \alpha, \beta \in V, a \in F$,就称 σ 是 V 的一个线性变换。

核与象 若 σ 是 V 的线性变换,记 $\ker \sigma = \{\alpha \in V \mid \sigma(\alpha) = \theta\}$, $\text{Im} \sigma = \{\sigma(\alpha) \mid \alpha \in V\}$,称 $\ker \sigma$ 为 σ 的核, $\text{Im} \sigma$ 为 σ 的象,它们都是 V 的子空间, σ 是单射 $\Leftrightarrow \ker \sigma = \{\theta\}$, σ 是满射 $\Leftrightarrow \text{Im} \sigma = V$,特别,当 $\dim V = n$ 时, $\dim \ker \sigma + \dim \text{Im} \sigma = n$,若 σ 是双射,称 σ 是 V 的自同构。

线性变换的矩阵 若 $\alpha_1, \dots, \alpha_n$ 是 V 的基, σ 是 V 的线性变换,且

$\sigma(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \dots + a_{ni}\alpha_n, 1 \leq i \leq n$,记 $(\sigma(\alpha_1), \dots, \sigma(\alpha_n)) = (\alpha_1 \cdots \alpha_n)A$, $A = (a_{ij})_n$,称 A 为 σ 关于基 $\alpha_1, \dots, \alpha_n$ 的矩阵, σ 关于不同基的矩阵相似:若 σ 关于基 $\beta_1, \dots, \beta_n$ 的阵为 B ,且

$(\beta_1, \dots, \beta_n) = (\alpha_1 \cdots \alpha_n)T$,则

$B = T^{-1}AT$,对于 σ ,若存在基 $\gamma_1, \dots, \gamma_n$,使 σ 的矩阵为对角阵,就

称 σ 可对角化。

线性变换的主要性质有: (1) 若 $\alpha_1 \cdots \alpha_n$ 相关, 则 $\sigma(\alpha_1), \dots, \sigma(\alpha_n)$ 相关; (2) $\alpha_1 \cdots \alpha_n$ 是 V 的基, $\beta_1, \dots, \beta_n$ 是 V 中任意 n 个向量, 则存在唯一的 σ , 使 $\sigma(\alpha_i) = \beta_i, 1 \leq i \leq n$, (3) 若 V 的线性变换 σ 关于基 $\alpha_1 \cdots \alpha_n$ 的阵为 A , 则 σ 可逆 $\Leftrightarrow A$ 可逆, 且 σ^{-1} 关于基 $\alpha_1 \cdots \alpha_n$ 的阵为 A^{-1} 。

欧氏空间的线性变换 最重要的是正交变换和对称变换。前者保持图形性质不变, 后者可以对角化, 若 σ 是 n 维欧氏空间 V 的线性变换, 且 $\forall \alpha, \beta \in V, (\alpha, \beta) = (\sigma(\alpha), \sigma(\beta))$, 称 σ 是正交变换, 正交变换保长、角、内积不变, 若 $\dim V = n$, 正交变换在标准正交基下的矩阵是正交方阵。对称变换是欧氏空间中又一类有内积性质的线性变换, 若 σ 是欧氏空间的线性变换, 且 $\forall \alpha, \beta \in V, (\alpha, \sigma(\beta)) = (\sigma(\alpha), \beta)$, 称 σ 是对称变换, 若 $\dim V = n$, 对称变换在标准正交基下的矩阵是实对称方阵, 因而对称变换均可对角化。

【二次型(二次齐式)】起源于解析几何中二次曲线, 二次曲面标准形的研究, 是线性代数主要内容之一。

数域 F 上 $X_1, \dots, X_n$ 的二次齐次多项式称 F 上二次型或二次齐式, 若

$$f(X_1, \dots, X_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} X_i X_j; \text{ 其中 } a_{ij} = a_{ji}, \text{ 则 } f(X_1, \dots, X_n) = X'AX,$$
 其中 $X' = (X_1, \dots, X_n)$, $A = (a_{ij})_n = A'$, 且 A 由 f 唯一确定, 称 A 为 f 之矩阵, $r(A)$ 称 f 之秩。如果 $f(X_1 \cdots X_n)$ 经变量的可逆变换 $X = PY, X = (X_1, \dots, X_n)'$, $Y = (y_1, \dots, y_n)'$ 得到 $g(y_1, \dots, y_n)$, 就说 $f(X_1, \dots, X_n)$ 与 $g(y_1, \dots, y_n)$ 等价。此时 g 的矩阵为 $P'AP$ 。一个只含平方项的二次型称标准二次型, 每个二次型在变量的可逆

变换下等价于一个标准二次型。

复数域上二次型的分类 由于复对称方阵合同于对角阵 $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, 这表示, 秩为 r 的复二次型等价于 $y_1^2 + \cdots + y_r^2$, n 个变量的复二次型共 $n+1$ 类, 它们有标准形 $0, y_1^2, y_1^2 + y_2^2, \cdots, y_1^2 + \cdots + y_n^2$ 。

实数域上二次型的分类 由于实对称方阵合同于对角阵 $\begin{bmatrix} I_p & & 0 \\ & -I_{r-p} & \\ 0 & & 0 \end{bmatrix}$, 且 r, p 由所给实对称阵唯一确定, 这表示, 每个实二次型的标准形是唯一的, 因此秩为 r 的不等价的标准形是 $y_1^2 + \cdots + y_r^2, y_1^2 + \cdots + y_{r-1}^2 - y_r^2, \cdots, -y_1^2 - \cdots - y_r^2$ 共 $r+1$ 类, 因此 n 个变量的实二次型互不等价的共 $1+2+\cdots+(n+1) = (n+2)(n+1)/2$ 类。正定二次型是实二次型中最重要的一类, 它的标准形为 $y_1^2 + \cdots + y_n^2$, 若二次型 $f(X_1 \cdots X_n) = X'AX$ 正定, 则 $\forall (c_1 \cdots c_n) \neq 0, f(c_1, \cdots, c_n) > 0$, 正定二次型的等价定义不下几十种。完全类似的可定义负定, 半正(负)定二次型。

实二次型 $f(X_1 \cdots X_n) = X'AX$ 还可以通过变量的正交变换 $X = UY$, (U 正交阵) 化为 $\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$, 其中 λ_i 为 A 的特征根。这种变换不改变图形的性质, 称这种标准形为主轴形式。

【多重线性代数】代数学的一个分支。近年来已成为许多数学分支常用的工具。它以建立在若干向量空间的卡氏积上的多重线性变换, 多重线性函数的代数结构为主要研究对象。

$V_1, V_2, \cdots, V_r, W$ 都是域 F 上的向量空间, f 是 $V_1 \times V_2 \times \cdots \times V_r$ 到 W 的映射, 且适合 $f(v_1, \cdots, v_i + av'_i, \cdots, v_r) = f(v_1, \cdots, v_i, \cdots,$

$v_r) + af(v_1, \dots, v'_i, \dots, v_r) \quad 1 \leq i \leq r$ 则称 f 是 r 重线性映射 (多重线性映射)。特别, $W=F$, 则称 f 是 r 重线性函数。

A 为 $m \times n$ 矩阵, $X \in F^m, y \in F^n, f(X, y) = X' Ay$, 则 f 是 $F^m \times F^n$ 到 F 的 2 重线性函数, 通称双线性函数。

V 为 F 上向量空间, V 到 F 的线性映射称为线性函数, 则 V 上全体线性函数集 V^* 作成 F 上的向量空间, 称为 V 的对偶空间。

$V_1, V_2, \dots, V_r$ 上全体 r 重线性函数集 $p(V_1, V_2, \dots, V_r)$ 作成 F 上的向量空间。 $r=1, p(V_1)$ 就是 V_1^* 。

设 σ 是 $V_1 \times V_2$ 到 W 的双线性映射, 如果对任意向量空间 U , $V_1 \times V_2$ 到 U 的任意双线性映射 φ , 总存在唯一的 W 到 U 的线性映射 ψ , 使得 $\varphi = \psi\sigma$, 就称 W 为 V_1, V_2 的一个张量积, 当 V_1, V_2 均为有限维时, V_1, V_2 的张量积存在, 且在同构意义下唯一, 记作 $V_1 \otimes V_2$ 。若 $e_1, \dots, e_n; f_1, \dots, f_m$ 分别为 V_1, V_2 的基, 则 $e_i \otimes f_j = \sigma(e_i, f_j)$ 构成 $V_1 \otimes V_2$ 的基, 因此 $\dim(V_1 \otimes V_2) = \dim V_1 \cdot \dim V_2$ 。

设 σ, τ 分别是 V_1, V_2 的线性变换, 则存在 $V_1 \otimes V_2$ 的唯一线性变换 φ , 使 $\varphi(\alpha \otimes \beta) = \sigma(\alpha) \otimes \tau(\beta), \forall \alpha \in V_1, \beta \in V_2$, φ 称 σ, τ 的张量积, 记 $\sigma \otimes \tau$ 。若 σ, τ 对于 V_1, V_2 的基 $e_1, \dots, e_n; f_1, \dots, f_m$ 的矩阵分别为 A, B , 则 $\sigma \otimes \tau$ 关于基 $\{e_i \otimes f_j, i=1, \dots, n, j=1, \dots, m\}$ 的阵为 $A \otimes B = (a_{ij}B)$, 这是一个 $n \times m$ 阶方阵, 也称 A, B 的张量积。

张量代数利用张量积建立了与向量空间相联系的代数结构。 V 是 F 上 n 维向量空间, 记

$$T_0(V) = F, T_r(V) = V \otimes \dots \otimes V \\ r=1, 2, \dots$$

$T_r(V)$ 的维数是 $n^r, r=0, 1, \dots$

令 $T(V)$ 表直和 $T_0(V) + T_1(V) + T_2(V) + \dots$, 则 $T(V)$ 是一个代数, 称为 V 的张量代数。

抽 象 代 数

【抽象代数】数学的一个重要分支。萌芽于 18 世纪末,19 世纪初,20 世纪 20 年代建立。伽罗瓦关于方程可根号解的研究引入了群、域的概念,哈密顿对四元数的发现,打破了“数”的传统观念,随着数系的扩充,四元数、向量、矩阵等对象的出现,代数的面貌发生了一系列深刻变化,作为研究数的运算性质的数学分支发展为以研究有运算的一般集合(称代数结构)为对象的数学分支。抽象代数的兴起是 20 世纪数学最突出的进展之一。今天,抽象代数、拓扑与泛函分析成为现代数学理论的三大支柱。

一个代数结构是指用公理化方法定义的这样一个非空集合,它有一个或几个代数运算,并且满足一定的算律。根据代数运算及所满足的算律而有各种不同的代数结构。若 V 是一个非空集合, V 的一个代数运算是指 $V \times V \rightarrow V$ 的一个映射 $\circ$,对于 V 中任意 $a, b, a \circ b \in V$ 。若 $a \circ b = b \circ a$,称 $\circ$ 满足交换律。群是一个重要的代数结构(参见“群论”),此外还有环、域、格、模等,对每个代数结构性质的讨论形成了抽象代数的各个专门学科,如群论、环论等。

等价关系与集合的分类是抽象代数的两个基本概念。集合 A 的一个等价关系是指 A 的一个关系 $\sim$,满足:(1)自反性:对于 a 属于 $A, a \sim a$;(2)对称性:对于 a, b 属于 A ,若 $a \sim b$,则 $b \sim a$;(3)传递性:对于 a, b, c 属于 A ,若 $a \sim b, b \sim c$,则 $a \sim c$,集合 A 的一个分类是指对集合 A 的一个划分,使得 A 中每一元素属于且仅属于

某一子集,这样的子集称类。 A 的一个等价关系可以确定 A 的一个分类;反之, A 的一个分类决定 A 的一个等价关系。 A 的由 $\sim$ 决定的各类的集合称 A 关于 $\sim$ 的商集,记作 $A/\sim$ 。

同态、同构是抽象代数的重要概念。若 V_1, V_2 是两个代数结构,分别有运算 $\circ, \bar{o}, V_1$ 到 V_2 的一个映射 φ ,若对于 a, b 属于 V_1 ,有 $\varphi(a \circ b) = \varphi(a) \bar{o} \varphi(b)$,称 φ 是 V_1 到 V_2 的同态映射。特别,满足上述条件的 φ 是双射时,称 φ 是 V_1 到 V_2 的同构映射。两个存在同态(构)映射的代数结构称为同态(构)的,两个同构的代数结构有相同的代数性质。在抽象代数中,关心的是代数结构的代数性质,同构的代数系看成一个。

女数学家诺特和阿廷对抽象代数的建立和发展起了巨大作用。她们的学生范·德·瓦尔登是抽象代数的传播者。1930 年出版的《近世代数学》传播了诺特、阿廷的思想,于是抽象代数蓬勃发展起来。代数结构中最常见、最基本的是群、环、域,也是形成较早的三个分支,继之格、模、交换代数,同调代数等相继发展起来。如今抽象代数不仅为现代数学各学科提供了有力工具,而且它的观点和方法渗透到了许多学科。

中国在抽象代数方面最早有成就的是曾炯之,他是诺特的学生。著名数学家华罗庚在代数方面做了不少杰出工作。此外中国数学家在代数的不少分支如有限群、典型群、环论、李代数等都有不少好的结果。

【群论】群是数学中一个重要概念,有关群的性质及结构的理论称群论,是抽象代数重要组成部分之一。群论的产生,根源之一是高次方程的根号解问题。拉格朗日关于方程根的对称函数的工作,引起人们注意根的置换的性质,导致置换群理论的产生。群论的创立

者被认为是伽罗瓦。如今,群论已发展成一个内容丰富、分支众多的学科。被认为是 19 世纪意义最重大的数学创造之一,它深刻地改变了人们对数学性质的认识,开创了全新的研究领域。

基本概念和性质 若 G 是一个有乘法 $\cdot$ 的非空集合,且满足以下条件,称 G 对 $\cdot$ 做成一个群,这些条件是:(1)乘法结合律;(2) G 中有单位元,即存在 e ,使对 G 中每个 a , $a \cdot e = e \cdot a = a$;(3) G 中每个元有逆元,即对 a ,存在 a^{-1} ,使 $a \cdot a^{-1} = a^{-1} \cdot a = e$,条件(2),(3)也可用(2)':对 G 中任意 a, b ,方程 $a \cdot X = b, y \cdot a = b$ 在 G 中有解代替。例如整数集 Z 对数的加法,非零有理数集 Q^* 对数的乘法,全体模 n 的剩余类 Z_n 对剩余类加法都是群。又如保持平面上正 n 边形不变的变换对于变换的合成也是一个群,记作 D_n ,称二面体群。以正方形为例, D_4 是指绕中心旋转 $0^\circ, 90^\circ, 180^\circ, 270^\circ$ 以及关于过中心的对称轴的反射,一共 8 个变换组成。

一个非空集合 A 的若干一一变换对于变换的合成做成的群,称 A 的一个变换群。每一个群都同某一变换群同构,这是著名的凯莱定理。若一个群对乘法满足交换律,称交换群或阿贝尔群。有时称交换群为加群,运算称加法,单位元改称零元,逆元改称负元。

若 H 是群 G 的非空子集, H 对 G 的运算做成群,称 H 是 G 的子群,记 $H \leq G, \{e\}, G$ 是 G 的两个当然子群,其它子群称 G 的真子群,若干个子群的交仍是子群。

若 $N \leq G$,且对 G 中每个元 $a, aN = Na$,称 N 是 G 的不变子群或正规子群,记作 $N \trianglelefteq G$ 。一个没有真不变子群的群称单群。若 $H \leq G, K \trianglelefteq G$,则 $HK \leq G$;若 $H, K \trianglelefteq G$,则 $HK \trianglelefteq G$ 。

若 $N \trianglelefteq G$,令 $G/N = \{aN | a \in G\}$,规定 $aN \cdot bN = (ab)N$,则 G/N 做成群,称 G 模 N 的商群。 $aN = \{an | n \in N\}$ 。

一个群所含元素的个数称群的阶,群 G 的阶记为 $|G|$,当 $|G|$

有限时,称 G 是有限群, $|G|$ 无限时,称 G 为无限群, $\mathbb{Z}, \mathbb{Q}^*$ 是无限群, $\mathbb{Z}_p$ 是有限群。若 $|G|$ 有限, $H \leq G$, 则 $|H| \mid |G|$, 且 $|G| = |H| \cdot (G:H)$, 这是著名的拉格朗日定理。若 $|G|$ 有限, $m \mid |G|$, G 中未必有 m 阶子群。对素数 p , 若 $p \mid |G|$, 则 G 中存在 p 阶子群。若 $a \in G$, 使 $a^k = e$ 的最小正整数 k 称 a 的周期, 记 $o(a)$ 或 $|a|$, 周期也称阶。若上述 k 不存在, 称 a 的周期无限或 0 , $\forall a \in G, a, a^{-1}$ 有相同的周期。

G, G' 是两个群, 运算分别为 $\circ, \bar{\circ}$, φ 是 G 到 G' 的一个映射, 若 $\forall a, b \in G$, 有 $\varphi(a \circ b) = \varphi(a) \bar{\circ} \varphi(b)$, 称 φ 是 G 到 G' 的同态映射, 当 φ 是满射时, 记 $G \xrightarrow{\varphi} G'$, 当 φ 是双射时, 称 φ 是 G 到 G' 的同构映射, 记 $G \xrightarrow{\varphi} G'$ 。若 $G' = G$, 上述 φ 称自同态或自同构。若 φ 是 G 到 G' 的同态映射, 称 $\ker \varphi = \{g \in G \mid \varphi(g) = e'\}$ 为 φ 之同态核, 称 $\text{Im} \varphi = \{\varphi(g) \mid g \in G\}$ 为 φ 之同态象, 且 $\ker \varphi \trianglelefteq G, \text{Im} \varphi \leq G'$ 。群的同态, 同构有以下基本定理:

同态基本定理 若 $N \trianglelefteq G$, 则 $G \xrightarrow{\varphi} G/N$; 若 $G \xrightarrow{\varphi} G'$, 则 $G' \simeq G / \ker \varphi$ 。这样, 同构意义下, G 只与商群同态, 只要知道 G 的所有不变子群, 就完全掌握了 G 的同态。

同构定理 若 $G \xrightarrow{\varphi} G', H \leq G$, $H \supseteq \ker \varphi, H' = \varphi(H)$, 则 $G/N \simeq G'/H'$;

若 $H, K \trianglelefteq G$, 则 $HK/K \simeq H/(H \cap K)$; 若 $H, K \trianglelefteq G$, 且 $H \supseteq K$, 则 $G/H \simeq G/K \mid H/K$ 。

群 G 的全体自同构集 $\text{Aut} G$ 对变换的合成做成群, 称 G 的自同构群。 $\varphi_a: X \rightarrow aXa^{-1} \quad \varphi_a \in \text{Aut} G$, 称 G 的内自同构, 全体内自同构的集合 $\text{Inn} G$ 是 $\text{Aut} G$ 的不变子群。

若 a 属于 G , G 中每个元有 a^i 形式, 称 G 为循环群, 记 $G =$

(a) , a 称 G 的生成元, 若 $o(a)$ 无限, G 称无限循环群, 否则 G 称有限循环群。循环群是一类结构完全清楚的交换群。若 $G=(a)$ 是无限循环群, 生成元只有 a, a^{-1} , 且 $G \simeq Z$ 。 n 阶循环群 $G=(a)$ 与 Z_n 同构, 且 a^k 是生成元 $\Leftrightarrow (n, k)=1$ 。这表示 n 阶循环群有 $\varphi(n)$ 个生成元。 $\varphi(n)$ 为欧拉 φ 函数。

若 G_1, G_2 是两个群, 那么 $G_1 \times G_2 = \{(a, b) | a \in G_1, b \in G_2\}$, 对于乘法 $(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 b_2)$ 做成群, 称 $G_1 \times G_2$ 为 G_1 与 G_2 的外直和, 它是由已知群构造的新群。若 $H_1, H_2 \leq G$, 且 $G = H_1 H_2, H_1 \cap H_2 = \{e\}$, 则 $G \simeq H_1 \times H_2$, 称 G 是子群 H_1, H_2 的内直和。此时 G 的结构完全由 H_1, H_2 决定, 从而对 G 的讨论可归结为对 H_i 的讨论。内、外直和的定义不难推广。一个群 G 若能表为真不变子群的直和, 称 G 可分解, 否则称 G 不可分解。

设 G 有一个有限递降子群列

$$(1) G = G_0 > G_1 > G_2 > \cdots > G_s = \{e\}$$

其中 $G_i \trianglelefteq G_{i-1} (i=1, \cdots, s)$, 称 (1) 为 G 的一个次正规列, $G_0/G_1, G_1/G_2, \cdots, G_{s-1}/G_s$ 为 (1) 的因子组; 若每个 G_i/G_{i-1} 是单群, 称次正规列为 G 的一个合成列。每个有限群的任意两个无重复项的合成群列有相同的长度, 而且它们的因子组在同构意义下不计次序对应相等。

若 $a, b \in G$, 称 $aba^{-1}b^{-1}$ 为一个换位元, 记 $[a, b]$, 此时 $[a, b]^{-1} = [b, a]$ 。 G 的一切有限个换位元之积的集合是 G 的不变子群, 称为 G 的换位子群, 记作 $G^{(1)}$ 。 $G^{(1)}$ 的换位子群记 $G^{(2)}, \cdots$, G 可换时, $G^{(1)} = \{e\}$ 。对每个群 $G, G/G^{(1)}$ 可换, 且若 $N \leq G, G/N$ 可换, 必有 $N \supseteq G^{(1)}$ 。对群 G , 若存在 $k, G^{(k)} = \{e\}$, 称 G 是可解群, 因此交换群都是可解群。 $n > 4, S_n$ 不是可解群。

群表示是研究群的有力工具之一。令 V 是 C 上向量空间, GL

(V)表 V 的全体可逆线性变换做成的群。对群 G , G 到 $GL(V)$ 的同态 ρ 称作 G 在 V 内的一个线性表示。即 $\rho: G \rightarrow GL(V)$, 满足对于任意 $s, t \in G$, $\rho(st) = \rho(s)\rho(t)$ 。称 V 是 G 的表示空间, 也称 V 是 G 的一个表示。若 $\dim V = n$, 称 n 为表示 ρ 的级。若 W 是 V 的子空间, ρ 在 W 上的限制即 G 到 W 内一个线性表示, 称 W 为 V 的一个子表示。如果 $V \neq \{0\}$, ρ 是 G 在 V 内的一个线性表示, 且不存在 V 的关于 ρ 的真不变子空间, 则称 ρ 为不可约表示或单表示。有限群表示的一个重要结果是: 每一个表示都是不可约表示的直和。

特征标是线性表示的函数, 它完全刻划了线性表示 ρ 。若 $\dim V = n$, 对于 V 的一个线性变换 σ , 若 σ 的矩阵为 $A = (a_{ij})$, 定义 $\text{Tr}\sigma = \sum_{i=1}^n a_{ii}$, 称为 σ 之迹, 若 ρ 是 G 在 V 内的线性表示, $\forall s \in G$, 令 $\chi_\rho(s) = \text{Tr}(\rho_s)$, 称 χ_ρ 为 ρ 之特征标。具有相同特征标的两个表示是同构的。同构意义下, G 的不可约表示的个数等于 G 的共轭类的个数。利用群表示理论可以证明 $p^m q^n$ 阶群是可解群, 利用特征标理论可以证明 $p^m (m > 1)$ 阶群不是单群, 有限群表示论在有限群的结构理论中有重要作用。

群表示是用具体的线性群(矩阵群)描述群的理论, 19 世纪末 20 世纪初由弗罗贝尼乌斯及伯恩赛德独立开创, 20 世纪三四十年代, 布饶尔更把群表示进一步深化, 引入模表示, 模特征标, 应用于有限单群分类的研究。

有限群集合 $A = \{1, 2, \dots, n\}$ 的一个一一变换 $\varphi: k \rightarrow i_k$ 称为一个 n 次置换, 常用 $\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ 表示。A 的若干置换所成的群称为一个 n 次置换群, 全体 n 次置换作成的群称为 n 次对称群, 记作

S_n , 且 $|S_n| = n!$ 。当 $i_1, \dots, i_n$ 是偶排列时, 称 $\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ 为偶置换, 全体偶置换做成一个群, 记作 A_n , 称为 n 次交代(错)群, $|A_n| = n! / 2$, 因此 $A_n \leq S_n$ 。置换群是一类特殊的有限群, 由凯莱定理可知, 每个有限群与一个置换群同构。

决定有限群的结构是有限群的重要课题之一。有限交换群结构已完全清楚。若 G 是交换群, $|G| = n = d_1 d_2 \cdots d_m$, 其中 $d_i > 0$, 且 $d_i | d_{i+1}$, 称 d_i 为 G 的不变因子, $\{d_1, d_2, \dots, d_m\}$ 为 G 的不变因子组。且 $G \simeq Z_{d_1} \oplus Z_{d_2} \oplus \cdots \oplus Z_{d_m}$ 。称 d_i 的各个素数方幂的因数为 G 的初等因子, 全体初等因子的集合为 G 的初等因子组。若 $\{n_1, n_2, \dots, n_t\}$ 是 G 的初等因子组, n_i 均为素数的方幂, 则 $G \simeq Z_{n_1} \oplus Z_{n_2} \oplus \cdots \oplus Z_{n_t}$ 。同构意义下, 有限交换群的个数取决于初等因子组或不变因子组的个数。

有限非交换群较之有限交换群要复杂很多。

关于有限群的基本结果, 除拉格朗日定理、凯莱定理外, 还有西罗定理及 p 群的两个结果。一个阶为 p 的幂的群称为一个 p 群。 p 群的中心仍为 p 群, 且非循环的交换 p 群必为一些循环 p 群的直和。若 $|G| = p^l m$, $(p, m) = 1$, 称 G 的阶为 p^e 的子群为 G 的西罗 p 子群。西罗第一定理指出, 一个有限群对每个素数 p 均存在西罗 p 子群, 西罗第二定理指出, G 的每个 p 子群均含于某个西罗 p 子群中, 且 G 的任意两个西罗 p 子群共轭。(若 H_1, H_2 是 G 的子群, 存在 $a \in G$, $H_2 = aH_1a^{-1}$, 则 H_1, H_2 共轭) 西罗第三定理指出, 若 G 的西罗 p 子群有 t 个, 则 $t \equiv 1 \pmod{p}$, 利用西罗定理可以决定某些有限群的结构或判断是不是单群。有限单群的分类定理: 每个有限单群必同构于下述情形之一: (1) 素数阶循环群; (2) $n \geq 5$, n 次交代群; (3) 16 个李型群无限族中的某一个成员; (4) 26 个散在群

之一。

无限群、典型群、无限李群、拓扑群等都是无限群。20 世纪 30 年代以来,无限群的研究有了迅速发展,人们期望把有限群的结果尽可能多地搬到无限群中,同时也研究无限群特有的问题。无限群的研究在其他代数系的研究中,也起着示范的作用。典型群的研究在中国有较长的历史和独特的方法,著名数学家华罗庚等人用矩阵方法讨论典型群的结构及自同构,被国际上公认为独特的方法。典型群是线性群、正交群、辛群、酉群的统称。若一个群 G 又是一个拓扑空间,而且此拓扑对群的运算为连续的,即(1) $\forall X, y \in G$ 及 Xy 的一个邻域 W ,必有 X 与 y 的邻域 U 与 V 使 $UV \subset W$; (2) $\forall X \in G, X^{-1}$ 的邻域 W ,必有 X 的邻域 U 使 $U^{-1} \subset W$,此时称 G 为一个拓扑群,如实数加群与非零实数乘法群对普通度量的拓扑,均为拓扑群。此外,线性代数群、算术群也是无限群。

【环与代数】环、代数是抽象代数的两个重要概念,环论的发展可追溯到 19 世纪关于实数域的扩张及其分类的研究,直到 20 世纪才形成抽象理论。诺特 1921 年的《环中的理想论》是环论的真正奠基。数十年来,环论的研究不论在理论方面还是在与其他学科相联系的应用方面都取得了很大进展。

基本概念与性质 设 R 是一个非空集合,有代数运算加法和乘法,满足以下条件时,称 R 为一个环:(1) R 对加法是交换群;(2)乘法满足结合律;(3)乘法对加法满足左、右分配律。有时也称这样的环为结合环,而称满足(1),(3)的为非结合环。例如整数环 $\mathbb{Z}$,偶数环 $2\mathbb{Z}$,数域 F 上全体 n 阶方阵环 $M_n(F)$, F 上一元多项式集 $F[X]$,模 n 的剩余类环 $\mathbb{Z}_n$ 都是(结合)环。一个乘法满足交换律的环称为交换环;一个非结合的交换环 R ,若满足 $[(aa)b]a = (aa)$

$(ba), (\forall a, b \in R)$, 称 R 为若当环; 一个非结合环 R , 若 $\forall a \in R, a^2 = 0$, 对于 $a, b, c \in R, (ab)c + (bc)a + (ca)b = 0$ 称 R 为李环。若环 R 中存在 $e, \forall a \in R$, 有 $ae = ea = a$, 称 e 为单位元, 一个有单位元的环称么环。在一个环 R 中, 若 $a \neq 0, b \neq 0, ab = 0$, 称 a, b 为 R 的零因子。若环 R 无零因子, $ab = 0$, 必有 $a = 0$ 或 $b = 0$ 。一个无零因子的交换么环称为整环。 $\mathbb{Z}, F[X]$ 是整环, $M_n(F)$ 有零因子, 因而不是整环。整环 $R = \{a + bi \mid a, b \in \mathbb{Z}\}$ 称高斯整环。在么环 R 中, 若 $a \neq 0, ab = ba = e$, a 称可逆元, b 称 a 的逆元, 记为 a^{-1} 。一个么环 R , 若非零元皆可逆, 则 R 为除环, 哈密顿 1843 年给出了四元数除环 (参见“数与多项式”)。一个交换除环称为域。

一个环所含元素个数为该环的阶, 阶有限的环称有限环, 否则称无限环。一个无零因子的环 R 中的非零元对加法有共同的阶, 或为 0, 或为素数, 称为环 R 的特征, 记作 $\text{ch}R$ 。子环、子整环、子除环概念类似子群概念。若 S 是 R 的子环, 且 $\forall r \in R, a \in S, ra (ar) \in S$, 称 S 是 R 的一个左(右)理想。一个既左又右的理想为(双边)理想。一个不含真理想的环为单环。环 R 的由一个元素 a 生成的理想 (a) 称 R 的主理想。若 R 是交换么环, 则 $(a) = \{ar \mid \forall r \in R\}$ 。理想在环中作用类似于不变子群在群中的作用。若 I 是环 R 的理想, 记 $R/I = \{a + I \mid \forall a \in R\}$, 对于加法: $(a + I) + (b + I) = (a + b) + I$, 乘法: $(a + I)(b + I) = ab + I$ 做成一个环, 称为 R 模 I 所得剩余类环或商环。

设 R, R' 是两个环, φ 是 R 到 R' 的一个映射, 且 $\forall a, b \in R, \varphi(a + b) = \varphi(a) + \varphi(b), \varphi(ab) = \varphi(a)\varphi(b)$, 称 φ 是 R 到 R' 的一个环同态(映射)。当

$\varphi(R) = R'$ 时, 称 R 与 R' 同态, 记为 $R \overset{\varphi}{\sim} R'$; 当 φ 是一一映射时,

称 R 与 R' 同构, 记为 $R \stackrel{\varphi}{\simeq} R'$; $\ker \varphi = \{a \in R \mid \varphi(a) = 0'\}$, 称为 φ 的同态核, $\ker \varphi$ 是 R 的理想。与群的同态、同构类似, 环有以下同态基本定理及同构定理:

环的同态基本定理: 若 I 是 R 的理想, 则 $R \sim R/I$; 若 $R \stackrel{\varphi}{\simeq} R'$, 则 $R' \simeq R/\ker \varphi$ 。

环的同构定理 若 $R \stackrel{\varphi}{\simeq} R'$, I' 是 R' 的理想, $I = \varphi^{-1}(I') = \{r \in R \mid \varphi(r) \in I'\}$, 则 $R/I \simeq R'/I'$; 若 I_1, I_2 是 R 的理想, 则 $I_1 + I_2/I_1 \simeq I_2/I_1 \cap I_2$; 若 I, J 是 R 的理想, 且 $J \supseteq I$, 则 $R/I \mid J/I \simeq R/J$ 。

若 R_1, R_2 是两个环, 做加群的外直和 $R = R_1 \oplus R_2$, 在 R 中定义 $(a_1, a_2) \cdot (b_1, b_2) = (a_1 b_1, a_2 b_2)$, 则 R 做成环, 为 R_1 与 R_2 的外直和。 R 中存在理想 R_1', R_2' , 使 $R_1' \simeq R_1, R_2' \simeq R_2$, 且 $R_1' \cap R_2' = (0)$ 。若 I_1, I_2 是环 R 的理想, 且 $R = I_1 + I_2, I_1 \cap I_2 = (0)$, 则有 $R \simeq I_1 \oplus I_2$, 称 R 为 I_1, I_2 的内直和。

若 H, K 是么环 R 的理想, 且 $R = H + K$, 称 H, K 互素; 若 R 为交换么环, H, K 是 R 的理想, 且 H, K 互素, 则 $HK = H \cap K$; 若 H, K 与 N 互素, 则 HK 与 N 互素。

若么环 R 的理想 $N_1, N_2, \dots, N_s$ 两两互素, 则 $R/N_1 \cap \dots \cap R/N_s \simeq R/N_1 \oplus \dots \oplus R/N_s$ 。

整数的同余概念可以推广到环上, 若 N 是么环 R 的理想, $a, b \in R, a - b \in N$, 称 a, b 模 N 同余, 记为 $a \equiv b \pmod{N}$, 上述事实可用同余式组表示, 即著名的中国剩余定理。

中国剩余定理 R 是么环, $N_i (i=1, \dots, r)$ 是 R 的两两互素的理想, 则对于任意 $b_1, b_2, \dots, b_r$ 属于 R ,

$$\begin{cases} X \equiv b_1 \pmod{N_1} \\ \vdots \\ X \equiv b_r \pmod{N_r} \end{cases}$$

在 R 中有解,且模 $(N_1 \cap \cdots \cap N_r)$ 唯一。

交换么环有两类重要的理想:素理想及极大理想。在一个交换么环 R 中,若 P 是 R 的理想,且 $P \neq R$,满足 $ab \in P$,则 $a \in P$ 或 $b \in P$,这样的 P 称为 R 的素理想。若 M 是 R 的理想, $M \neq R$,若不存在 R 的任意含 M 的真理想,则 M 是 R 的极大理想。一个交换么环一定有极大理想,且一个交换么环 R 的理想 $P, M, P \neq R$ 是素理想 $\Leftrightarrow R/P$ 是整环; $M \neq R$ 是 R 的极大理想 $\Leftrightarrow R/M$ 是域。

唯一分解整环 一类重要的整环。在整环 I 中,若 $a \in I, a$ 可逆,称 a 为 I 的可逆元素(或称单位)。 I 中非可逆元素的非零元 p ,若 $p=ab$ 则 a, b 之一为可逆元素,就称 p 为素元,其它非可逆元素的非零元称可约元。若 $a \neq 0, a$ 为非可逆元素,且 a 可表为有限个素元之积,称 a 可分解;若 $a=p_1 \cdots p_t=q_1 \cdots q_s$ (p_i, q_j 为素元),就有 $t=s$,且适当调换元素次序后, p_i 与 q_j 至多差一个可逆元素的因子,称 a 可唯一分解;若一个整环 I 中的每个非零非可逆元素均可唯一分解,称 I 是唯一分解整环, $\mathbb{Z}, F[X]$ 都是唯一分解整环,但整环 $R=\{a+b\sqrt{3}i \mid a, b \in \mathbb{Z}\}$ 不是唯一分解整环。一类重要的唯一分解整环是主理想整环,即一个整环,它的每个理想都由一个元素生成;另一类重要的唯一分解整环是欧氏环,即 R 是一个整环,存在 R 的非零元素集 R^* 到非负整数集 N 的一个映射 φ ,取定 R 的一个非零元 a ,对于 R 的任意元 b 都可以写成 $b=aq+r$ ($q, r \in R$) 的形式,这里 $r=0$ 或 $\varphi(r) < \varphi(a)$, $\mathbb{Z}, F[X]$ 都是欧氏环。

【(结合)代数及其表示】 A 是一个非空集合, F 是域,若满足以下条件,称 A 是一个代数:(1) A 有加法,乘法及 F 对 A 的数乘法运算;(2) A 对加法及乘法做成环, A 对加法及数乘法作成 F 上向量空间。比如 F 上全体 n 阶方阵的集合 $M_n(F)$ 对矩阵的加法,乘法

及数乘法作成代数,称为全阵代数。特别若 A 对加法,乘法作成交换么环,称 A 为交换代数;若 A 对加法,数乘法作成有限维向量空间,称 A 为有限维代数,与环类似,代数也有子代数、商代数、直和、同态等概念。

结合代数的研究始于 19 世纪中期,哈密顿对四元数代数的考察以及凯莱等人对矩阵代数的讨论。其目标是刻划各种类型的结合代数的结构及表示。

20 世纪初期提出的韦德伯恩定理,对结合代数的研究具有深远的影响。这一理论的主要内容是:(1)任意有限维结合代数 A 含有一个极大的幂零理想 N (所谓 N 是幂零的,指存在一个自然数 n ,使 N 中任意 n 个元素之积均为 0)。它包含 A 的一切幂零理想,称为 A 的幂零根。商代数 A/N 的幂零根为零,幂零根为零的代数称为半单代数;(2)半单代数是有限个单代数的直和;(3)域 F 上的单代数形如全阵代数 $M_n(D)$,其中 D 是 F 上有限维可除代数,且 D 和 n 是唯一确定的。根据这一定理,对有限维代数的研究,基本上可以归结为对幂零代数与可除代数的研究。

所谓有限维结合代数的表示,是指代数到域 F 上的全阵代数内的同态映射,或等价地说,是代数 A 上的一个模。近 20 年来蓬勃发展起来的代数表示论,就是研究一个有限维代数的模范畴的理论,这一研究广泛应用了同调代数、代数拓扑等工具,对多种类型的非半单代数进行了深入的刻划。代数 A 的一个表示(或一个 A 模)就象 A 的一张照片。称代数 A 是有限表示型的,若在同构的意义下,仅有有限多个不可分解的有限维 A 模(所谓模 M 不可分解,是指如果 $M=M_1\oplus M_2$,则 $M_1=0$ 或 $M_2=0$),否则称 A 为无限表示型的。因而一个有限表示型代数是有限张照片就可以表达清楚的一种代数。而无限表示型代数需要无限多张照片进行刻划,

相应地复杂一些。代数表示论的任务是对一个有限表示型代数,给出全体有限维不可分解模的同构类。否则,若代数是无限型的,给出模的分布情况。目前,有限表示型代数已经得到了详尽的刻划,人们正在向无限型迈进。

有限维代数的绝大多数结果,都可以平行地推广到阿廷环上去。阿廷环是对理想有降链条件的环,即若 $I_1 \supset I_2 \supset I_3 \supset \cdots$ 是环中理想的一个降链,则存在正整数 n , 使 $I_n = I_{n+1} = \cdots$ 。

【域论】域是代数学中最基本的概念之一,历史悠久。19 世纪 30 年代,伽罗瓦在他的著作中已有了域概念的萌芽,如今,域论已成为抽象代数的一个重要分支,是不少学科的基础。一个交换除环称域。即若 F 至少含有两个元素,且有两个代数运算 $+$, $\cdot$, F 对 $+$ 及 F 非零元集 F^* 对 $\cdot$ 均为交换群,且 $\cdot$ 对 $+$ 满足分配律。研究域的代数分支称域论。一个含无穷多元素的域称无限域;一个含有限个元素的域称有限域或伽罗瓦域。由于域是交换单环,无真理想,因而域不能象群,环那样,通过对不变子群,商群或理想、商环来讨论。常用的方法是从已知域出发,研究它的扩域。

一个不含真子域的域称素域,有理数域 Q 是特征 0 的素域, Z_p 是特征 p 的素域。并且若 F 是域, $ch F = 0$ 时, F 含与 Q 同构的素子域; $ch F = p$ 时, F 含与 Z_p 同构的素子域。

若 F 是 E 的子域,即 E 是 F 的扩域,记为 E/F 。若 E/F , 且 $\alpha \in E$, E 的含 F, α 的最小子域记为 $F(\alpha)$ 。为添加 α 于 F 所得的单扩域,且 $F(\alpha) = \{f(\alpha)/g(\alpha) \mid g(\alpha) \neq 0, f(\alpha), g(\alpha) \text{ 均为 } F \text{ 上 } \alpha \text{ 的多项式}\}$; 若 $\alpha_1, \cdots, \alpha_r \in E$, E 的添加 $\alpha_1, \cdots, \alpha_r$ 于 F 的最小子域记为 $F(\alpha_1, \cdots, \alpha_r) = F(\alpha_1) \cdots F(\alpha_r)$ 。

若 $E/F, \delta \in E$, 且存在多项式 $f(X) = \sum_{i=0}^n a_i X^i (a_i \in F)$ 使 $f(\delta) = \sum_{i=0}^n a_i \delta^i = 0$, 称 δ 为 F 上代数元, 否则称 δ 为 F 上超越元; 若 E/F , 且 E 中元均为 F 上代数元, 则称 E 是 F 的代数扩域; 若 α 是 F 上代数元, 则 $F(\alpha)$ 是 F 的代数扩域, 称为 F 的单代数扩域, 添加一个超越元所得扩域即为单超越扩域。单代数扩域与单超越扩域有不同的结构: 若 α 为 F 上超越元, 则 $F[\alpha] \simeq F[X]$ 的商域; 若 α 为 F 上代数元, 则 $F(\alpha) \simeq H[X]/(p(X))$, 其中 $p(X)$ 是 F 上首 1 的不可约多项式, 且 $p(\alpha) = 0$ 时。称 $p(X)$ 为 α 在 F 上的极小多项式。并且若 $\partial(p(X)) = n$, 则 $F(\alpha) = \left\{ \sum_{i=0}^{n-1} a_i \alpha^i \mid a_i \in F \right\}$ 。

给定扩域 E/F , E 作为 F 上向量空间, 若 $\dim_F E = n$, 称 n 为 E 在 F 上的次数, 记作 $(E:F)$ 。若 $(E:F)$ 有限, 称 E 为 F 上有限扩域, 否则称为无限扩域。若有域列 $E \supseteq I \supseteq F$, 且 $(E:I), (I:F)$ 皆有限, 则 $(E:F)$ 有限, 且 $(E:F) = (E:I)(I:F)$ 。一般, 若 $E = E_0 \supseteq E_1 \supseteq E_2 \supseteq \cdots \supseteq E_{s-1} \supseteq E_s$, 且 $(E_{i-1}:E_i)$ 有限 ($i = 1, \dots, s$), 则 $(E:E_s)$ 有限, 且 $(E:E_s) = (E_0:E_1)(E_1:E_2)\cdots(E_{s-1}:E_s)$ 。单代数扩域 $F(\alpha)$ 是 F 的有限扩域, 若 α 在 F 上的极小多项式为 $p(X)$, 则 $(F(\alpha):F) = \partial(p(X))$ 。并且 F 的每个有限扩域一定是 F 的代数扩域; 若 $\alpha_1, \dots, \alpha_t$ 均为 F 上代数元, 则 $F(\alpha_1, \dots, \alpha_t)$ 是 F 的有限扩域, 因而是代数扩域。

每一个次数大于 0 的数系数多项式在给定的数域上未必能分解成一次因式之积, 但代数基本定理保证了每一个次数大于 0 的多项式在复数范围都可分解为一次因式之积。如果 F 上每一个次数大于 0 的多项式均可分解为 F 上一次因式之积, 称 F 为代数闭域, 此时 F 不再有真的代数扩域。每个域 F 均存在代数扩域 E , 使

E 为代数闭域;复数域是代数闭域。

若 E 是 F 的扩域,对于 F 上多项式 $f(X)$,在 E 上 $f(X)$ 可以分解为一次因式之积,并且对 F 的任一较小扩域 I , $f(X)$ 在 I 上不能分解为一次因式之积,称 E 为 $f(X)$ 在 F 上的分裂域或根域。对于 F 上每个多项式,同构的意义下均存在唯一的分裂域。一个多项式的分裂域依赖于这个多项式系数所在的域,如 X^2-2 作为有理系数多项式,其分裂域为 $Q(\sqrt{2})$,但作为实系数多项式,其分裂域仍为 R 。若 $E \supseteq F$, $f(X) \in F[X]$, 在 E 上 $f(X) = a(X - \alpha_1)^{k_1} \cdots (X - \alpha_t)^{k_t}$, 则 $f(X)$ 在 F 上分裂域恰为 $F(\alpha_1, \dots, \alpha_t)$ 。

有限域是一类重要的域,有限域的特征为素数。设 F 为特征 p 的有限域, Δ 为 F 的素子域, 且 $(F : \Delta) = n$, 则 $|F| = p^n$, 且 F 是 $X^p - X$ 在 Δ 上分裂域。对任何素数 p 及正整数 n , 均存在 p^n 个元的有限域, 且同构意义下唯一。有限域结构简单: 若 F 是有限域, Δ 为 F 的素域, 则存在 F 中 α , $F = \Delta(\alpha)$ 。有限域的非零元构成的乘群皆为循环群。若 $F^* = (a)$, 则 $F = \Delta(\alpha)$ 。利用有限域的性质可以构造出各种对称性质的组合结构, 如正交拉丁方, 平衡区组设计等, 这些组合结构有效地应用于试验设计, 通信系统等许多实际领域中, 特别随着计算机技术蓬勃发展, 有限域理论成了广大工程技术人员不可缺少的数学工具。

【伽罗瓦理论】历史上, 这个理论起源于方程根号解的问题。16 世纪前, 5 次以下代数方程已经了解公式, 5 次及 5 次以上的代数方程是否也有解的公式, 是当时数学家们所关注的问题。直到 1770 年, 法国数学家拉格朗日首先认识到用代数运算解一般的高次 ($n > 4$) 方程看来是不可能的。1801 年高斯也感到这个问题也许是不能解决的。1824 年, 挪威年仅 22 岁的阿贝尔证明了一般 5 次

方程的根式求解的不可能性,不幸他过早去世。法国另一位天才数学家伽罗瓦继续了阿贝尔的工作,深刻阐明了代数方程可以用根号解的充要条件。伽罗瓦理论主要是在一个特定的扩域 K/F 的中间域所组成的集合与 K 的 F 自同构群的子群所组成的集合之间建立了一个映射,从而将域论的研究归为群论的研究,得到域论中一些较为深刻的结果,以下只假定域的特征为 0 介绍这一结果。

伽罗瓦理论主要研究的是正规扩域,若 N 是特征零的域 F 的有限扩域,且对 F 上任一不可约多项式 $f(X)$,或者 $f(X)$ 在 N 中无根,或者 $f(X)$ 的全部根均在 N 中,称 N 为 F 的正规扩域,或 F 的伽罗瓦扩域。若 $N \supseteq F, (N:F)=n$, 则 N 是 F 的正规扩域 $\Leftrightarrow N$ 是 F 上某多项式的分裂域 $\Leftrightarrow$ 若 $\alpha \in N$, 则 F 上与 α 有相同极小多项式的 β (称与 α 共轭) 也在 N 中。

伽罗瓦理论的主要研究对象是伽罗瓦群。若 N 是 F 的正规扩域,称由 N 的保持 F 的元不动的自同构集组成的 $\text{Aut} N$ 的子群为 N 在 F 上的伽罗瓦群,记为 $\text{Gal} N/F, \text{Gal} N/F = \{\sigma \in \text{Aut} N \mid \sigma|_F = I_F\}$, 且 $|\text{Gal} N/F| = (N:F)$; 若 $f(X) \in F[X], N$ 是 $f(X)$ 在 F 上的分裂域,则 $\text{Gal} N/F$ 为方程 $f(X)=0$ 在 F 上的伽罗瓦群。 F 上无重根的 n 次方程 $f(X)=0$ 的伽罗瓦群与 S_n 的一个子群同构。若 $u_1, \dots, u_n$ 为域 K 上无关未定元, K 上 n 次一般方程 $f(X) = X^n - u_1 X^{n-1} + \dots + (-1)^n u_n = 0$ 在 $F = K(u_1, \dots, u_n)$ 上的伽罗瓦群与 S_n 同构。

若 N 是 F 的正规扩域,令

$$\tilde{L} = \{L \mid L \text{ 是 } N \text{ 的含 } F \text{ 的子域}\},$$

$$\tilde{H} = \{H \mid H \text{ 是 } \text{Gal} N/F \text{ 的子群}\}, \text{ 令 } \varphi: \tilde{H} \rightarrow \tilde{L}$$

$$H \rightarrow I_{nv} H \quad I_{nv} H = \{a \in N \mid a^\tau = a, \forall \tau \in H\} \text{ 称 } H \text{ 的不变子域。}$$

$$\phi: \tilde{L} \rightarrow \tilde{H}$$

$$L \rightarrow \text{Gal} N/L$$

称 φ, ψ 为伽罗瓦映射, φ, ψ 有以下性质:

$$G_{a_i}(N/\text{Inv}H) = H,$$

$$\text{Inv}(\text{Gal} N/L) = L$$

伽罗瓦基本定理 I: 设 N 是 F 的正规扩域, 上述 φ, ψ 满足 $\varphi_0\psi = l_{\tilde{H}}, \varphi_0\psi = l_{\tilde{L}}$ 。

伽罗瓦基本定理 II: 对上面定义的 φ , 若 H 是 $\text{Gal} N/F$ 的正规子群, 则 $\varphi(H) = \text{Inv}H$ 是 F 的正规扩域, 反之亦然, 并且 $\text{Gal} N/F/\text{Gal} N/L \simeq \text{Gal} L/F$ 。

伽罗瓦理论的应用:

(1) 方程可根号解的充要条件 域 F 上方程 $f(X)=0$ 在 F 上可根号解 $\Leftrightarrow f(X)=0$ 在 F 上的伽罗瓦群是可解群。

阿贝尔, 鲁菲尼并且证明了特征零的域上 $n(>4)$ 次一般方程不能用根号解。

(2) 尺规作图问题 指用无刻度直尺及圆规经有限步做几何图形。利用尺规可做如下工作: 1. 过两点作一直线; 2. 用给定的圆心及半径作圆; 3. 若二直线相交, 可求其交点; 4. 若一直线与一圆相交, 可求其交点; 5. 若两圆相交, 可求其交点。用代数的语言叙述, 即在取定直角坐标系后, 每个点有确定的坐标 (a, b) , 每条直线、每个圆有确定的方程 $y=kX+d, X^2+Y^2+eX+fY+g=0$, 称直线或圆方程的系数为直线或圆的坐标。这样, 取定直角坐标系后, 平面上每一初等几何图形有图形的坐标 (一组确定实数) 与之对应。一个几何图形可用尺规作图 $\Leftrightarrow$ 任意取定平面上的一个直角坐标系, 对于给定图形的任意取定了的坐标 $a, b, c, \dots$ 来说, 只要能选取所做图形的坐标, 使它们属于 $F=Q(a, b, c, \dots)$ 的一个正规

扩域 K , 而 $(K : F) = 2^m$, m 为非负整数。

利用上述尺规作图的充要条件, 可以解决正 n ($n > 2$) 边形的作图问题。正 n 边形可作 $\Leftrightarrow n = 2^s$ 或 $n = 2^l p_1 \cdots p_t$, $s > 1, l \geq 0, p_i$ 为互不相同的费马素数。

现 代 代 数

【格】一种代数系统, 具有两个代数运算 $\vee, \wedge$, 适合以下对称的算律, 记作 $(L, \vee, \wedge)$

$$a \vee a = a \quad a \wedge a = a \quad \text{幂等律}$$

$$a \vee b = b \vee a \quad a \wedge b = b \wedge a \quad \text{交换律}$$

$$(a \vee b) \vee c = a \vee (b \vee c) \quad \text{结合律}$$

$$(a \wedge b) \wedge c = a \wedge (b \wedge c)$$

$$a \vee (a \wedge b) = a \quad \text{吸收律}$$

$$a \wedge (a \vee b) = a$$

格 $(L, \vee, \wedge)$ 中, 可定义偏序关系 (适合自反性、反对称性、传递性) $\leq$: $a \leq b$ 当且仅当 $a \wedge b = a$ 。于是, 集 L 与 $\leq$ 构成偏序集 $(L, \leq)$, 并且任意二元 a, b 都有最大下界 $a \wedge b$ 与最小上界 $a \vee b$ 。

格有等价定义: 任二元 a, b 都有最大下界与最小上界的偏序集。

集合 S 的幂集 $p(S)$ 关于集合运算 $\cup, \cap$ 作成格 $(p(S), \cup, \cap)$ 。此时 S_1, S_2 的最大下界, 最小上界分别为 $S_1 \cap S_2$ 与 $S_1 \cup S_2$ 。自然数集 N 对整除关系 $|$ 作成格, a, b 的最大下界与最小上界分别是

a, b 的最大公因数 (a, b) 与最小公倍数 $[a, b] = \frac{ab}{(a, b)}$ 。

格中分配不等式成立, 即 $\forall a, b, c$

$$a \vee (b \wedge c) \leq (a \vee b) \wedge (a \vee c),$$

$$(a \wedge b) \vee (a \wedge c) \leq a \wedge (b \vee c)$$

对任意元, 分配律成立的格称为分配格。即 $\forall a, b, c$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

模格是指 $\forall a, b, c$, 以下关系式成立的格:

$$b \leq a, a \wedge (b \vee c) = b \vee (a \wedge c)$$

模格也称为迪德金格。

分配格都是模格, 但模格未必是分配格。

若格 $(L, \leq)$ 中有元素 1 , 对任何 $X \in L$ 都有 $X \leq 1$, 则称 1 是格 $(L, \leq)$ 的全上界。若格有全上界 1 , 则必唯一。同样可定义格的全下界 0 , 即对任何 $X \in L$, 都有 $0 \leq X$ 。若有全下界, 也必唯一。

格 $(p(S), \subseteq)$ 是分配格, 且有全上界 1 , 即 S ; 也有全下界 0 , 即空集 $\emptyset$ 。格 $(N, |)$ 有全下界 0 , 即自然数 1 ; 但没有全上界。 $N \cup \{0\}$ 对整除关系 $|$ 也作成格, $(N \cup \{0\}, |)$ 既有全下界 0 , (即数 1), 又有全上界 1 (即数 0)。实数集 R 对普通小于等于关系的格 $(R, \leq)$ 既没有全上界也没有全下界。

既有全上界 1 又有全下界 0 的格称为有界格。有界格中, 对元 a , 若存在元 b 适合: $a \vee b = 1$, 且 $a \wedge b = 0$, 则称 b 是 a 的一个补元。若有界格的每个元都有补元, 则称为有补格。有界分配格中, 若 a 有补元, 则必唯一, 记作 $\bar{a}$ 。

有补分配格称为布尔格, 记为 $(B, \vee, \wedge)$ 。由于每个元 a 都有唯一的补元 $\bar{a}$, 求 a 的补元可认为是 B 的一元代数运算。这样, 布

尔格有三个代数运算, $\vee, \wedge, -$, 记做 $(B, \vee, \wedge, -)$ 。也称为布尔代数。布尔代数 $(B, \vee, \wedge, -)$ 还有以下算律:

$$\overline{\overline{a}} = (\overline{\overline{a}}) = a \quad \text{对合律}$$

$$\overline{a \vee b} = \overline{a} \wedge \overline{b} \quad \text{德·摩根律}$$

$$\overline{a \wedge b} = \overline{a} \vee \overline{b}$$

布尔代数 $(B, \vee, \wedge, -)$ 中, 对每个元 a , 都有 $0 \leq a$ 。若不存在 $b \neq 0, b \neq a$, 使 $0 \leq b \leq a$, 则称 a 是原子。对任一元 b , 必有原子 $a, a \leq b$ 。原子 a 的基本性质是: 对任一非零元 b 都有: $a \leq b, a \leq \overline{b}$ 有且仅有一个成立。

有限布尔代数中, 任一非零元 a 都可唯一地写成

$$a = c_1 \vee c_2 \vee \cdots \vee c_k, c_1, c_2, \cdots, c_k \text{ 是所有 } \leq a \text{ 的原子。}$$

设有限布尔代数 $(B, \vee, \wedge, -)$ 的所有原子集为 S , 则 $(B, \vee, \wedge, -)$ 与幂集布尔代数 $(p(S) \cup, \cap, \sim)$ 是同构的 (这里 $\tilde{A}$ 表示 A 的补 $S - A$)。即存在 B 到 $p(S)$ 的 1-1 映射 f , 适合

$$f(a \vee b) = f(a) \cup f(b), f(a \wedge b) = f(a) \cap f(b), f(\overline{a}) = \widetilde{f(a)}.$$

由于 $|S| = n, |p(S)| = 2^n$, 任何有限布尔代数恰有 2^n 个元, n 为其原子数。

布尔代数 $(B = \{0, 1\}, \vee, \wedge, -)$ 特别重要。它的三个运算表如下:

$\vee$	0	1	$\wedge$	0	1	$-$	
0	0	1	0	0	0	0	1
1	1	1	1	0	1	1	0

布尔代数的元、变元, 以及由有限运算 $\vee, \wedge, -$ 构成的符号串, 称为布尔表达式。若布尔代数 B 上的 n 元函数 (即 B^n 到 B 的

函数),能用布尔表达式表示,则称该函数是布尔函数。

布尔代数 $B = (\{0, 1\}, \vee, \wedge, -)$ 上的任何 n 元函数均为布尔函数。

命题逻辑可用布尔代数 $B = (\{0=F, 1=T\}, \vee, \wedge, -)$ 来描述。命题作为变元,任何复合命题可用布尔表达式表示出来。开关代数也可用布尔代数 $B = (\{0, 1\}, \vee, \wedge, -)$ 来表述,0 表示断开,1 表示闭合, $\vee$ 表示并联, $\wedge$ 表示串联, $-$ 表示反向。任何仅用并联、串联、反向构成的电路,都可用布尔表达式表示出来。对该表达式的化简,相应地得到电路的简化。这是有重要意义的工作。

【模】一个代数系,抽象代数研究的内容之一。模是域上向量空间的推广,是将向量空间的系数域改为环得到的。模范畴是同调代数研究的主要对象,也是现代群和代数的表示理论的主要工具。

令 R 是一个环, M 是一个加法交换群,定义环 R 对 M 的模运算如下: $\forall a \in R, m \in M, a \cdot m$ 是模 M 中的一个元素,并且这一运算满足下述公理: $\forall a, b \in R, m, n \in M$. (1) $a \cdot (m + n) = a \cdot m + a \cdot n$; (2) $(a + b) \cdot m = a \cdot m + b \cdot m$; (3) $(ab) \cdot m = a \cdot (b \cdot m)$ 。这时称 M 是环 R 的一个左模。类似地,可以定义环 R 上的右模,若 R 中有单位元,且模运算满足 $1 \cdot m = m$,则 M 称为么模。若 M, N 是 R 模, f 是 M 到 N 的一个交换群同态,并且保持模运算: $\forall a \in R, m \in M, f(a \cdot m) = a \cdot f(m)$,称 f 是模 M 到 N 的一个(模)同态,平行于群和环,模也有子模、商模,模的直和等概念,有模同态基本定理及同构定理。

一类最简单的模是循环模,若 M 是一个 R 模,存在 M 中的元素 X ,对于 M 中每个元素 y ,均存在 R 中的 a_y ,使 $y = a_y X$,称 M 是一个循环 R 模,每一个循环群 (a) 是一个循环 $\mathbb{Z}$ 模,么环 R 是一个

循环 R 模。

最常用到的模有如下几类:

(1) 单模 若 R 模 M 除 $\{0\}$ 和 M 外, 无其它真子模, 称 M 为单模或不可约模。例如, 若 I 是环 R 的一个极大左理想, 则 R/I 是单模, 又如, 若 p 是素数, Z 是整数环, 则 $Z/(p)$ 是单模。 R 模 M 是单模, 当且仅当 M 是一个非零循环模, 而且每个非零元都是它的生成元。一个单模的自同态环是除环, 这就是有广泛应用的舒尔引理。

(2) 自由模 设 M 是一个 R 模, 若有 M 的子集 $S = \{s_i | i \in I\}$, 使 M 中的每一个元素 m 都可以表成如下形式: $m = \sum_{k=1}^n a_k \cdot s_{i_k}$, 其中 $a_k \in R, s_{i_k} \in S$, 则称 S 是 M 的一个生成元系。若表达式中的系数 a_k 由 m 唯一确定, 则称 M 是定义在集合 S 上的自由模, 记作 F 。自由模的重要性在于, 任意 R 模 M 都是某个自由 R 模的同态象, 事实上, 取 M 的一个生成元素 $\{m_i | i \in I\}$, 取一个抽象集合 $S = \{s_i | i \in I\}$ 与之一一对应, 令 F 是定义在 S 上的自由模, 令 $\varphi: F \rightarrow M$, $\varphi\left(\sum_{k=1}^n a_k \cdot s_{i_k}\right) = \sum_{k=1}^n a_k m_{i_k}$, 则 φ 是满同态。令 $K = \ker \varphi$, 得短正合序列 $O \rightarrow K \rightarrow F \rightarrow M \rightarrow O$ 。(参见“同调代数”)。

(3) 投射模 如果对于任意 R 模的满同态 $\pi: C \rightarrow B$ 以及任意模同态 $f: P \rightarrow B$, 恒有 $g: P \rightarrow C$, 使 $\pi \circ g = f$, 即图形

$$\begin{array}{ccc} & P & \\ g \swarrow & & \searrow f \\ C & \xrightarrow{\quad \pi \quad} & B \end{array}$$

可交换, 则称 P 是一个投射模。以下四个命题等价: (1) P 是投射模; (2) 对任何满同态 $M \rightarrow P$, 必有 Q , 使 $M \simeq P \oplus Q$; (3) $P = P_1 \oplus P_2$

时, P_1 与 P_2 均为投射模; (4) 存在模 P' 及自由模 F , 使 $F = P \oplus P'$, 这些等价叙述刻画了投射模的基本性质。由 (4) 可知自由模是投射模。因此任取模 M , 有投射模 P , 使序列 $0 \rightarrow K \rightarrow P \rightarrow M \rightarrow 0$ 正合, 此列称为 M 的投射表现。若 M 有另一个投射表现, $0 \rightarrow K' \rightarrow P' \rightarrow M \rightarrow 0$, 则由沙鲁尔定理, $K \oplus P' \simeq K' \oplus P$, K 与 K' 或者均为投射模, 或者均非投射模。正是由于这个性质, 投射模在模论与同调代数中占有特殊重要的地位。

(4) 入射模 这是一个与投射模相对偶的概念, 其定义可由投射模定义中将所有的箭头调转方向得到。任何一个模 M 都可以嵌入到一个入射模 I 中, 令 $\psi: M \rightarrow I$ 是嵌入映射, $W = \text{Coker} \psi$, 则有正合序列 $0 \rightarrow M \rightarrow I \rightarrow W \rightarrow 0$ 。

模的分解是模论的一个基本问题, 主理想整环上有限生成模的分解问题已彻底解决。

若 D 是整环, M 是一个 D 模, $a \in M, \exists r \in D, r \neq 0, ra = 0$, 称 a 为 M 的扭元素, 否则称 a 为 M 的自由元素。若 M 为 D 上有限生成模, 则 M 中全体扭元素做成 M 的子模, 记为

$\text{Tor}(M)$, 若 $X \in M$, 称 $\text{ann}(X) = \{b \in R \mid bX = 0\}$ 为 D 模 M 中 X 的阶理想, 它是 D 的理想。若 $\text{ann}(z) = (p^e)$, p 为 D 的素元, 称 Dz 为 p 准素循环模。主理想整环上有限生成模的基本结构定理是说: 主理想整环 D 上任一非零有限生成模 M 是一些循环模的直和:

$$M = Dz_1 \oplus Dz_2 \oplus \cdots \oplus Dz_s$$

其中 $\text{ann}(z_1) \supseteq \text{ann}(z_2) \supseteq \cdots \supseteq \text{ann}(z_s), \text{ann}(z_i) \neq 0, D, 1 \leq i \leq r; \text{ann}(z_j) = 0, r+1 \leq j \leq s$ 。此时 $\text{Tor}(M) = Dz_1 \oplus \cdots \oplus Dz_r, K = Dz_{r+1} \oplus \cdots \oplus Dz_s$ 是自由子模, $M = \text{Tor}(M) \oplus K$, 并且称在基本结构定理中 $\text{ann}(z_1), \text{ann}(z_2), \cdots$ 为 M 的不变因子理想, 它由 M 唯一确定, 有限生成扭模 $\text{Tor}(M)$ 可唯一地表为准素循环模 Dy_i 之直和, 称

$\text{ann}(y_i)$ 为 M 的初等因子理想, 它由 M 唯一确定。

由于有限生成阿贝尔群 G 是一个有限生成 Z 模, 利用主理想整环上有限生成模的结构定理可以得到有限生成阿贝尔群的分解定理; 若 σ 是 F 上 n 维空间 V 的线性变换, 求 V 的基使 σ 的矩阵为某个标准形状的问题与 V 做为 $F[\lambda]$ 模, (此时仍是主理想整环上的有限生成模。) 讨论 $F[\lambda]$ 模 V 的分解有密切关系。

【交换代数】20 世纪初期以代数几何与代数数论为背景产生和发展起来, 并为这两个古老的学科提供了新的, 有力的数学工具。交换代数的研究对象是有单位元的交换环 R (参见“环与代数”)。事实上, 在代数几何中, 需要考察基域 F 上的多元多项式环 $F[X_1, X_2, \dots, X_n]$ 。而在代数数论中, 需要讨论整数环 Z , 两者即为交换环的典型例子。

交换代数的核心概念是素理想 (参见“环与代数”)。它是素数概念的一种推广。事实上, 在整数环 Z 中, 一个理想是素理想, 当且仅当它由一个素数生成或等于零。一个交换么环 R 中全体素理想构成的集合叫做环 R 的谱。根据上述讨论, 整数环 Z 的谱是集合 $\{(p) \mid p \text{ 是素数}\} \cup \{0\}$ 。

交换代数的主要方法是所谓分式环和局部化。令 S 是交换么环 R 中的一个子集, 若 $1 \in S$, 且, $s_1, s_2 \in S$ 意味着 $s_1 \cdot s_2 \in S$, 则称 S 是一个乘闭子集。在 $R \times S$ 上定义关系 $\sim: (a, s) \sim (a', s')$ 当且仅当存在 $t \in S$, 使 $t(as' - a's) = 0$, $\sim$ 是一个等价关系。将 (a, s) 代表的等价类记作 $\frac{a}{s}$, 等价类的全体记作 $S^{-1}R$, 则 $S^{-1}R$ 在加法: $\frac{a}{s} + \frac{b}{t} = \frac{ta + sb}{st}$ 及乘法 $\frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}$ 下做成一个环, 称 R 对 S 的分式环。

从交换环 R 得到 $S^{-1}R$ 的过程, 可以用从整数环 Z 得到有理数域

Q 为例。事实上,令 $S = \mathbb{Z} \setminus \{0\}$, $S^{-1}\mathbb{Z}$ 就是 Q 。最重要的分式环是取 $S = R \setminus P$, 其中 P 是 R 的一个素理想。这时 $S^{-1}R$ 记作 R_P , 称为 R 在 P 处的局部化。 R_P 仅有唯一的极大理想 $S^{-1}P$, 因而是一个局部环。在多项式环

$F[X_1, X_2, \dots, X_n]$ 中, 理想 $P = (X_1 - \alpha_1, X_2 - \alpha_2, \dots, X_n - \alpha_n)$ 就是一个素理想, 其中 $\alpha_1, \alpha_2, \dots, \alpha_n$ 取自基域 F 。显而易见, 这个素理想对应于 F^n 中的一个点 $(\alpha_1, \alpha_2, \dots, \alpha_n)$, 因而研究这点邻近的几何性质, 可以用研究 R_P 的代数性质来实现。这就是代数簇的局部性质(在一点邻域的性质)可以用交换代数的方法进行研究的原理。

交换代数的奠基性工作是由诺特和克鲁尔完成的。称一个环是诺特环, 若这个环满足对理想的升链条件, 也就是说, 任取 $0 \subset I_1 \subset I_2 \subset \dots$ 是一个理想的升链, 则存在正整数 n , 使 $I_n = I_{n+1} = \dots$ 。在一个交换诺特环中, 任取一个理想都可以分解为若干准素理想之交, 且这一分解在某种意义下是唯一的, 而克鲁尔建立了赋值论, 局部化理论和克鲁尔维数, 为古典几何提供了全新的代数工具。20 世纪五十年代以后, 交换代数得到了很大的发展, 模论、同调代数, 特别是格罗腾迪克的概型理论, 对交换代数的发展起到了巨大的推动作用。

【同调代数】将代数拓扑学中的同调论略去其几何意义, 平移至环上的模范畴。这门学科是美国和欧洲的数学家于 20 世纪 40 年代彼此独立提出并发展起来的。现在, 同调代数已成为研究群、环、一般代数、李代数、代数几何的不可缺少的有力工具, 同调一词出自希腊文, 意指和谐一致。

平行于代数拓扑, 我们可以定义复形及其同调模。令 R 是有单位元的环, $n \in \mathbb{Z}$, M_n 是 R 模, 则模与同态的序列

$$\cdots \rightarrow M_{n+1} \xrightarrow{d_{n+1}} M_n \xrightarrow{d_n} M_{n-1} \rightarrow \cdots$$

称为一个复形,若 $d_n d_{n+1} = 0$,这一复形记作 (M, d) 。 $Z_n = \text{Ker} d_n$ 叫作 n 循环, $B_n = \text{Im} d_{n+1}$ 叫作 n 边缘。由复形的定义条件 $B_n \subset Z_n$, $H_n(M, d) = Z_n / B_n$ 叫作第 n 个同调模。如果对任意整数 n , $B_n = Z_n$,换言之, $H_n(M, d) = 0$,则称 (M, d) 是一个正合列。例如令 M 是任意 R 模, $\epsilon: P_0 \rightarrow M$ 是投射模 P_0 到 M 的一个满同态。以 $K_0 = \text{Ker} \epsilon$ 代替 M , 作投射模 P_1 到 K_0 的满同态 $d_1, \dots$, 由归纳法, 得到一个正合序列:

$$\cdots \rightarrow P_2 \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} M \rightarrow 0$$

使得 M 右边的项皆为 0, 称为 M 的投射分解。类似地, 运用一个模到入射模的嵌入, 可以得到模的入射分解。最后, 若正合序列 (M, d) 中除三个互相连接的项外皆为 0, 即

$$0 \rightarrow M_2 \xrightarrow{d_2} M_1 \xrightarrow{d_1} M_0 \rightarrow 0$$

即称为短正合序列。这时 d_2 是单射, d_1 是满射, 且 $\text{Ker} d_1 = \text{Im} d_2$ 。

同调代数的主要内容是研究 Hom 函子, $\otimes$ 函子, 及其导出函子 Ext 和 Tor 。现举例说明如下: 固定一个左 R 模 A 。 $\text{Hom}(A, -): {}_R\mathbf{M} \rightarrow \mathbf{Ab}$ 是从左 R 模范畴到 $\mathbf{Abel}$ 群范畴的一个函子, 任取左 R 模 B , 令

$$\cdots \rightarrow P_2 \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow[\epsilon]{\epsilon} B \rightarrow 0$$

是 B 的一个投射分解。

$$\cdots \rightarrow \text{Hom}(A, P_2) \xrightarrow{\text{Hom}(A, d_2)} \text{Hom}(A, P_1) \xrightarrow{\text{Hom}(A, d_1)} \text{Hom}(A, P_0) \rightarrow 0$$

是一个复形, 这个复形的第 n 个同调模记作 $\text{Ext}^n(A, B) = \text{Ker} \text{Hom}(A, d_n) / \text{Im} \text{Hom}(A, d_{n+1})$, 且 $\text{Ext}^0(A, B) = \text{Hom}(A,$

B)。称模 E 是模 B 借助 A 的扩张,若有短正合列

$$0 \rightarrow B \rightarrow E \rightarrow A \rightarrow 0。$$

也就是说, B 是 E 的一个子模,而 E 对这一子模的商模恰好是 A 。两个扩张 E 与 E' 称为等价,若有交换图

$$\begin{array}{ccccc} & & & E & \\ & & & \nearrow & \searrow \\ 0 \longrightarrow B & \downarrow \varphi & A \longrightarrow & 0 & \\ & & & \searrow & \nearrow \\ & & & E' & \end{array}$$

此处 φ 必为模同构。这样定义的等价是一个等价关系,令 $e(A, B)$ 表所有等价类的集合,可以证明 $e(A, B)$ 与 $Ext^1(A, B)$ 一一对应。

计算各种群、环或代数的上同调模与同调模是这门学科的重要研究课题。

【范畴论】范畴和函子的理论是由艾伦伯格与麦克莱恩于 1945 年正式提出来的。这个理论一经提出就受到数学家们的普遍欢迎并迅速深入到数学的各个领域,成为一种统一的行之有效的数学语言。

一个范畴 C 由下述三要素组成:(1)一类象元 $A, B, C, \dots$;(2) C 中的任意两个象元的有序对 (A, B) 对应着一个集合 $C(A, B)$, 集合中的元素称为射元;(3) $\forall f \in C(A, B), g \in C(B, C)$, 乘积 $g \circ f \in C(A, C)$, 且乘法符合以下三条公理:①若有序对 $(A, B) \neq (A', B')$, 则 $C(A, B) \cap C(A', B') = \emptyset$;②若 $f \in C(A, B), g \in C(B, C), h \in C(C, D)$, 则 $h \circ (g \circ f) = (h \circ g) \circ f$ 有结合律;③对任意象元 A , 有恒等射元 $1_A \in C(A, A)$, 使得 $\forall f \in C(A, B)$, 有 $f \cdot 1_A = f = 1_B \cdot f$ 。按照这一定义,集合论研究的是集合范畴 S , 其象元是集合,而射元是集合之间的映射;群论研究的是群范畴 G , 其象元是

任意群,而射元是群同态;拓扑学研究的是拓扑空间范畴 T_{op} ,其象元是拓扑空间,而射元是拓扑空间的连续映射。在上述三个特例中,射元的合成是自然的,且满足范畴所要求的三条公理。

从范畴 C 到 C 的一个函子 $F:C \rightarrow C'$ 是一种双重意义的对应。一方面,它将 C 中的每一个象元 A 取 C' 中的一个象元 $F(A)$ 与之对应。另一方面, $\forall f \in C(A, B)$, 有 $F(f) \in C'(F(A), F(B))$ 与之对应,且满足 $F(1_A) = 1_{F(A)}$, $F(g \circ f) = F(g) \cdot F(f)$, 例如令 C 是群范畴, C' 是全体交换群作成的范畴。对任意群 G , $G^{(1)}$ 记 G 的换位子群, 令 $F(G) = G/G^{(1)}$, 若 $f: G \rightarrow H$ 是一个群同态, 令 $F(f): G/G^{(1)} \rightarrow H/H^{(1)}$ 是 f 在商群上的诱导同态, 则 F 是一个函子。

【李群和李代数】李于 19 世纪创立,并由基灵、嘉当、外尔等人于 19 世纪末 20 世纪初加以完善和发展的重要代数学分支。

李群是兼有代数和解析结构的一种数学系统。 G 是一个李群意味着: (1) G 是一个群; (2) G 是一个实或复的解析流形; (3) 对于 G 的任意元素 X, y , $(X, y) \rightarrow Xy^{-1}$ 所定义的 $G \times G \rightarrow G$ 的映射是一个解析映射。例如 $GL(n, R)$ (或 $GL(n, C)$), n 维实 (或复) 向量空间的所有自同构组成的群, 是一个李群。

考虑李群 G 在单位元 e 处的切向量空间 $\mathfrak{Y}$, 它是有限维的。在 $\mathfrak{Y}$ 内引入运算

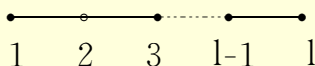
$$[X, y] = Xy - yX, \quad \forall X, y \in \mathfrak{Y}$$

就赋予了 $\mathfrak{Y}$ 一个代数结构, 称为对应于 G 的李代数, 李代数的运算适合反交换律: $[X, y] = -[y, X]$ 及雅可比恒等式 $[X, [y, z]] + [y, [z, X]] + [z, [X, y]] = 0$ 。进一步, 任一有限维向量空间 $\mathfrak{Y}$, 若有 $\mathfrak{Y}$ 中满足上述条件的双线性二元运算, 则称 $\mathfrak{Y}$ 是一个李代数。

李的基本定理: 两个单连通李群同构的充分必要条件是它们

有同构的李代数;任一抽象的李代数必为一单连通李群的李代数。

令 $\mathfrak{g}$ 是域 $\mathfrak{F}$ 上的一个李代数, $\mathfrak{M}, \mathfrak{N}$ 是 $\mathfrak{g}$ 的子空间, 记 $[\mathfrak{M}, \mathfrak{N}] = \{\sum [X, y] \text{ (有限和)} \mid X \in \mathfrak{M}, y \in \mathfrak{N}\}$, 则 $[\mathfrak{M}, \mathfrak{N}]$ 是 $\mathfrak{g}$ 的一个子空间。如果 $[\mathfrak{M}, \mathfrak{g}] \subset \mathfrak{M}$, 称 $\mathfrak{M}$ 是 $\mathfrak{g}$ 的子代数, 如果 $[\mathfrak{M}, \mathfrak{g}] \subset \mathfrak{N}$, 称 $\mathfrak{N}$ 是 $\mathfrak{g}$ 的一个理想。由于 $[\mathfrak{M}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{M}]$, 李代数的理想都是双边的。称一个李代数 $\mathfrak{g}$ 是单的, 如果 $\mathfrak{g}$ 除了本身和 $\{0\}$ 以外, 不含其它的理想。复数域上单李代数的结构和分类, 早在 19 世纪末就已经得到, 在同构的意义下, 它们只有 $A_l (l \geq 1), B_l (l \geq 1), C_l (l \geq 1), D_l (l \geq 3)$ 这四类和五个例外李代数, 分别记作 G_2, F_4, E_6, E_7, E_8 , 它们的维数分别是 14, 52, 78, 133 和 248。例如 A_l 由全体迹为 0 的 $(l+1)$ 阶复方阵组成, 这些单李代数都可以用一种非常简单的图形, 即所谓邓肯图刻划出来, 例如 A_l 对应的邓肯图是



这个漂亮的分类定理是运用基灵型, 嘉当子代数, 外尔群等概念证明的。

1968 年, 卡茨和穆迪彼此独立地提出了一类新的李代数, 可看作复单李代数在无限维情况的自然类比, 称之为卡茨-穆迪代数。目前对这种代数及特征 $p > 0$ 的域上的李代数的研究正方兴未艾。

【代数几何】现代数学的一个重要分支, 同时也是一个历史悠久的古老课题, 代数几何起源于对平面上的代数曲线的研究。如我们在解析几何中熟知的直线和二次曲线(包括椭圆、双曲线、抛物线)。由于次数为一或二的曲线都是有理曲线, 它们在代数几何的意义下同构于直线。故而人们一般认为, 代数几何始于 19 世纪上半叶

关于三次或更高次的平面曲线的研究,例如阿贝尔在 1827—1829 年提出的椭圆曲线,方程 $y^2 = X^3 - X$ 是复平面上的一条椭圆曲线。再如著名的费马猜想可以归结为下述问题:在实平面中,由方程 $X_1^n + X_2^n = 1$ 定义的曲线(费马曲线),当 $n \geq 3$ 时没有坐标都是非零有理数的点。对于这个经典问题的研究,近年来运用代数几何的方法获得了重大突破。1983 年,法尔廷斯证明了,费马方程在 $n \geq 4$ 时最多只有有限多个非零有理解。

在现代数学中,代数几何的基本研究对象为:令 F 是一个固定的域,在任意维 F 的仿射或射影空间中,由若干个代数方程的公共零点所构成的集合叫作一个代数簇,这些方程叫作这个代数簇的定义方程组。一个代数簇通常记作 V ,定义方程组中方程的系数及 V 中点的坐标取自 F , F 叫作 V 的基域。代数几何就是用代数的方法研究代数簇的几何特性。实或复平面上的代数曲线当然是代数簇的一个特例。如果一个代数簇 V 不能分解成两个比它小的代数簇的并,则称 V 是不可约的。这时,定义在 V 上的有理函数的全体构成一个域,叫作 V 的有理函数域。它是基域 F 上的一个有限生成的扩域(见域论)。通过这样的对应,将一个几何实体与有限生成的扩域这样一个代数问题紧密地联系起来。

人们对代数簇的研究通常分为局部和整体两个方面。局部方面的研究主要是用交换代数的方法讨论代数簇中的奇异点以及代数簇在奇异点周围的性质。作为奇异点的例子,可以考察由方程 $X^2 - y^3 = 0$ 。所定义的平面曲线中的原点 $(0, 0)$,这是一个歧点,当前代数几何的研究重点是整体问题。主要是代数簇的分类以及给定的代数簇的子簇的性质。同调代数的方法在这类研究中起着关键的作用。

黎曼是对代数几何的发展作出过杰出贡献的数学家。他于

1857 年引入的代数函数论,使代数曲线的研究获得了决定性的突破,运用代数函数域的概念,黎曼定义了代数曲线的一个最重要的数值不变量:方格。这也是代数几何历史上的第一个绝对不变量,既不依赖于代数簇对空间的嵌入的不变量,黎曼还提出了黎曼—罗赫定理的原始形式。这一著名的定理,是由希莱布鲁赫和格罗腾迪克完善并证明的。

20 世纪 50 年代以后,代数几何的研究对象又从代数簇拓广到了概型。它允许点的坐标在任意有单位元的交换环中选取,并允许结构层中存在幂零元。概型理论是由格罗腾迪克建立的。它完成了代数几何与代数数域算术的完美的统一,使得人们可以用几何方法来研究代数数论的一些经典问题,例如费马猜想。

微 积 分 学

【微积分学】分析数学是在微积分学基础上形成的,其研究内容包括极限理论、微积分学、微分方程、函数论、泛函分析与傅里叶分析等分支。它的产生与发展是由于力学、天文学、几何学等领域中所提出的问题的推动。例如,求曲线的切线等导致微分学的产生,而求平面曲边图形的面积等则是积分学的最早发源地之一。

作为微积分学的基础的极限的思想,可以追溯到古代。我国《庄子》(公元前 389—286)有“一尺之棰,日取其半,万世不竭”的论述。公元前 5 世纪希腊人已有了无限小的概念。但微积分是在

17 世纪才产生,其标志是牛顿的“流数法”和莱布尼茨的“逆切法”。微积分的产生与运用,解决了许多长期使人困惑的问题。经过达朗贝尔、拉格朗日、阿贝尔、波尔查诺等人的探索,直到 19 世纪,柯西建立了严格的极限理论,才奠定了微积分的基础,又经外尔斯特拉斯的完善,形成了现在的数学分析。而且由此发展出函数论、微分方程、泛函分析等,使分析数学成了数学的一大分支。

极限理论

【函数】设 X 和 Y 是两个非空的实数集合, f 是某一规则,若对 X 中的每个 x ,由 f 有 Y 中唯一的 y 与之对应,就称 f 是定义于 X 上的一个函数,记作 $f:X \rightarrow Y$,或 $y=f(x)$,称 X 为函数 $f(x)$ 的定义域, Y 为其值域, x 为自变量, y 为因变量。研究函数一般从下列简单性质入手:有界性、单调性、奇偶性、周期性、渐近性等。

【复合函数】设函数 $y=\varphi(u)$ 的定义域为 U , $u=f(x)$ 的定义域为 X ,值域为 $U^* \subseteq U$,则 $y=\varphi(f(x))$ 定义了一个函数,称为复合函数,例如 $y=\lg \sin x$,此时 $f(x)=\sin x$, $\varphi(u)=\lg u$,这个复合函数的定义域是 $(2k\pi, (2k+1)\pi)$ 。

【反函数】设 $y=f(x)$ 为已知的函数,若对每个 $y \in Y$,有唯一的 $x \in$

X 与之对应,使 $f(X)=y$, X 成了 y 的函数,记为 $X=f^{-1}(y)$, f^{-1} 称为 f 的反函数。习惯上以 X 表示自变量, y 表示因变量,故 $y=f(X)$ 的反函数为 $y=f^{-1}(X)$ 。例如 $y=\sin X$ 与 $y=\arcsin X$ 互为反函数,在同一坐标系中, $y=f(X)$ 的图象和 $y=f^{-1}(X)$ 的图象关于直线 $y=X$ 对称。(图 1)。

图 1

【隐函数】若能由函数方程 $F(X, y)=0$ 确定 y 为 X 的函数 $y=f(X)$, 则称 y 是 X 的隐函数。

【多元函数】设点 $(X_1, \dots, X_n) \in G \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^1$, 若对每一点 $(X_1, \dots, X_n) \in G$, 由某规则 f 有唯一的 $u \in U$ 与之对应, 则称 f 为一个 n 元函数, 记作 $u=f(X_1, \dots, X_n)$, G 为其定义域, U 为值域。

【基本初等函数及其图象】幂函数 $y=X^\mu$ ($\mu \neq 0$ 为任意实常数) 定义域: ① μ 为正整数时为 $(-\infty, +\infty)$, μ 为负整数时是 $(-\infty, 0)$

$\cup (0, +\infty)$; ② $\mu = \frac{1}{\alpha}$ (α 为正整数): 当 α 是奇数时为 $(-\infty, +\infty)$, 当 α 是偶数时为 $(0, +\infty)$; ③ $\mu = p/q$ (p, q 互素), 作为 $y = (X^p)^{1/q}$ 的复合函数进行讨论。(图 2、3)。

图 2

图 3

【指数函数】 $y = a^x$ ($a > 0, a \neq 1$)。定义域为 $(-\infty, +\infty)$, 值域为 $(0, +\infty)$ 。 $a > 0$ 时是严格单调增加的函数; $0 < a < 1$ 时是严格单减函

数。对任何 a , 图形均过 $(0, 1)$ 点。 $y=a^x$ 与 $y=(\frac{1}{a})^x$ 的图形关于 y 轴对称(图 4)。

图 4

【双曲函数】 $y=shX=\frac{e^X-e^{-X}}{2}$, $y=chX=\frac{e^X+e^{-X}}{2}$, $y=thX=\frac{shX}{chX}=\frac{e^X-e^{-X}}{e^X+e^{-X}}$ 依次叫做双曲正弦函数、双曲余弦函数及双曲正切函数, 统称双曲函数。它们的反函数分别记作 $y=arishX$, $y=archX$, $y=arthX$, 统称为反双曲函数。 $y=arshX=\ln(X+\sqrt{X^2+1})$, $y=archX=\ln(X+\sqrt{X^2-1})$, $y=arthX=\frac{1}{2}\ln\frac{1+X}{1-X}$ 。

双曲函数列表于下:

双曲正弦	双曲余弦	双曲正切	双曲余切
$\frac{e^X-e^{-X}}{2}$	$\frac{e^X+e^{-X}}{2}$	$\frac{e^X-e^{-X}}{e^X+e^{-X}}$	$\frac{e^X+e^{-X}}{e^X-e^{-X}}$

【对数函数】 $y=\log_a X (a>0)$, 定义域为 $(0, +\infty)$, 值域为 $(-\infty, +\infty)$, $a>1$ 时是严格单增的, $0<a<1$ 时是严格单减的。不论 a 为何值, 对数函数的图形均过 $(1, 0)$ 点。对数函数与指数函数互为反函

数。以 10 为底的对数称为常用对数,简记为 lgX 。在科学技术中普遍使用的是以 e 为底的对数,即自然对数,记作 lnX (图 5)。

图 5

【三角函数】列表于下(“正弦”即正弦函数…)

	记法	定义域	值域	周期
正弦	$\sin X$	$(-\infty, +\infty)$	$[-1, 1]$	2π
余弦	$\cos X$	$(-\infty, +\infty)$	$[-1, 1]$	2π

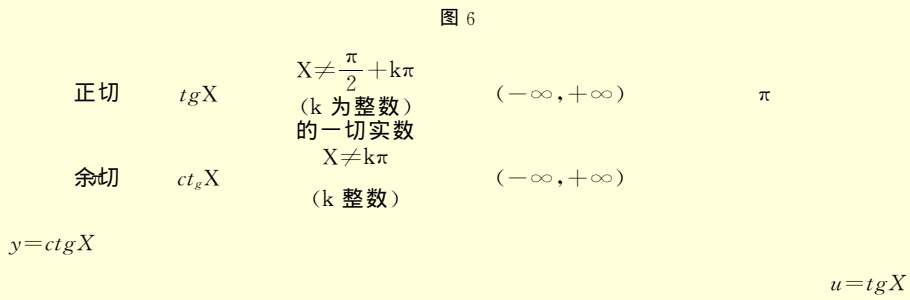


图 7

【反三角函数】列表于下：

反正弦	$arcsinX$	$[-1, 1]$	$[-\frac{\pi}{2}, \frac{\pi}{2}]$
反余弦	$arccosX$	$[-1, 1]$	$[0, \pi]$
反正切	$arctgX$	$(-\infty, +\infty)$	$[-\frac{\pi}{2}, \frac{\pi}{2}]$
反余切	$arcctgX$	$(-\infty, +\infty)$	$(0, \pi)$

图 8

【 R^n 空间】 n 维欧氏空间 $\{X = (X_1, X_2, \dots, X_n) \mid -\infty < X_i < +\infty\}$, 其中区间定义为 $I = \{X \mid a_i < X_i < b_i, i = 1, 2, \dots, n\}$ 。这里 $a_i, b_i \in R$, 且 $a_i < b_i$, a_i 和 b_i 可分别取为 $-\infty$ 与 $+\infty$, 点 $X = (X_1, X_1, \dots, X_n)$,

$y = (y_1, y_2, \dots, y_n)$ 间的距离 $d(X, y) = |X - y| = \sqrt{\sum_{i=1}^n (X_i - y_i)^2}$,

距离 $= d(X, y)$ 具下列三性质: $d(X, y) \geq 0$, $d(X, y) = 0 \Leftrightarrow X = y$; $d(X, y) = d(y, X)$; $d(X, y) \leq d(X, z) + d(z, y)$ 。

【邻域】设 $a \in R^n, r > 0$, 称集 $E = \{X \in R^n: d(X, a) < r\}$ 为点 a 的 r 邻域, 记作 $N(a, r)$, 称 r 为这个邻域的半径, 也称 $N(a, r)$ 为开球。例如, 平面 R^2 上点 $P_0(X_0, y_0)$ 的 ϵ 邻域是满足 $d(P, P_0) = \sqrt{(X - X_0)^2 + (y - y_0)^2} < \epsilon$ 的点集, 亦即以 P_0 为中心, 以 ϵ 为半径

的圆的内部。

【内点】设集合 $E \subset \mathbb{R}^n$, 点 $a \in \mathbb{R}^n$ 。若存在 $r > 0$, 使 $N(a, r) \subset E$, 则称 a 为 E 的内点。如果 E 的每一点都是 E 的内点, 就称 E 为开集。例如直线上的开区间 $(a, b) (a < b)$ 就是开集。

【外点】设 $P_1 \notin E$, 如果存在 P_1 的 δ 邻域 $N(P_1, \delta)$, 它不含 E 的点, 就称 P_1 是 E 的外点。

【边界点】设 $P^* \in \mathbb{R}^n$, P^* 既可属于 E , 也可不属于 E , 若对 P^* 的任何邻域 $N(P^*, \epsilon)$, 其中既含 E 的点, 又含有不是 E 中的点, 就称 P^* 是 E 的边界点。例: $\mathbb{R}^2$ 中的以 (X^0, y^0) 为中心的圆: $(X - X_0)^2 + (y - y_0)^2 \leq r^2$, 当不等式成立时, $p(X, y)$ 为内点, 使等式成立的点, 为边界点。

【极限点(零点)】若对 a 的任一邻域 $N(a, r)$, 均有 $E \cap (N(a, r) - \{a\}) \neq \emptyset$, 就称 a 是集合 E 的一个极限点。下面几个命题是等价的: a 是 E 的极限点 $\Leftrightarrow a$ 的任一邻域中均含有 E 的无限多个点 $\Leftrightarrow E$ 中有点列 $\{a_n\}$, $a_i \neq a_j (j \neq i)$, $a_n \neq a, (n = 1, 2, \dots)$ 而 $a_n \rightarrow a$ 。

【孤立点】设 $X_0 \in E$, 若存在 X_0 的邻域 $N(X_0, r)$ 使 $(N(X_0, r) - \{X_0\}) \cap E = \emptyset$, 则称 X_0 为 E 的孤立点。 E 的全体极限点所组成的集合称为 E 的导集, 记作 E' 。若 $E' \subset E$, 则称 E 为闭集。 $E \cup E'$ 称为闭包, 记作 $\bar{E}$, 当 $E' = E$ 时称 E 为完全集。任意个开集的并是开集, 有限个开集的交为开集; 任意个闭集的交为闭集; 有限个闭集的并为闭集。

【连续统】微积分学是建立在极限理论基础上的,而极限论又基于实数的完备性(连续性)。从有理数集 Q 出发建立实数系,有多种方法,戴德金分割理论是其中的一种。

设 $\zeta \subset Q$, ζ 关于 Q 的余集 $\bar{\zeta} = Q \setminus \zeta$, 当且仅当满足: (1) $\zeta \neq \emptyset$ 且 $\bar{\zeta} \neq \emptyset$, (即 $\zeta \neq Q$); (2) $\forall a \in \zeta, \forall \bar{a} \in \bar{\zeta}$, 有 $a < \bar{a}$, (在 Q 的序的意义下); (3) ζ 没有最大的(有理)数, 就称 Q 的分类 $(\zeta, \bar{\zeta})$ 是一个戴德金分割。并称 ζ 为这个分割的下类, $\bar{\zeta}$ 为上类。当且仅当 $(\zeta, \bar{\zeta})$ 的上类有最小的数时, 说 ζ 是有理实数; 当且仅当 $\bar{\zeta}$ 没有最小的数时, ζ 是一个无理实数。“割口”的数仍记为 ζ 。这就是说, 在上下类的分界处, 若有“空隙”, 便填充了一个无理数。将实数域记为 R 。例: $\zeta = Q^- \cup \{0\} \cup \{a \in Q^+ \mid a^2 < 2\}$, ζ 是实数 $\sqrt{2}$ 。 Q^-, Q^+ 分别为负、正有理数集。

【序】 $\forall \zeta, \eta \in R$, 当且仅当 $\bar{\zeta} \cap \eta \neq \emptyset$ 时, 称 ζ 小于 η , 记作 $\zeta < \eta$, 或 η 大于 ζ , 记作 $\eta > \zeta$ 。“序”具有三歧性, $\forall \zeta, \eta \in R$, 下面三种情况恰有一种成立: $\zeta = \eta, \zeta < \eta, \eta < \zeta$; 传递性, $\zeta < \eta, \eta < \zeta \Rightarrow \zeta < \zeta$, 关于正实数、负实数, 零, 四则运算(做除法时须除数不等于零), 遵守通常的算律, 对这些运算构成了实数域。实数域 R 也叫数的连续统。实数域具有(1)稠密性: $\forall \zeta, < \eta$, 存在有理数 b 使 $\zeta < b < \eta$; (2) E 实数集的阿基米德性: $\forall \zeta \in R^+$, 总存在 $n^* \in Z^+$ (正整数集), 使 $\zeta < n^*$; (3) R 没有最大数, 也没有最小数; (4) 完全(备)性: 设 $\{A, \bar{A}\}$ 是一个再分割, 则上类 $\bar{A}$ 必有最小数。也就是说, R 不再有“空隙”。

与实数集完备性有关的基本定理: (1) 闭区间套定理, 以自然数为自变量的函数 $f: N \rightarrow X$ 叫做序列; 设 $X_1, X_2, \dots, X_n, \dots$ 是集合

的序列若 $X_1 \supset X_2 \supset \cdots \supset X_n \supset \cdots$, 则称它是区间套; 对于任何闭区间套 $I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots$, 且 $\forall \epsilon > 0, \exists I_k$, 使其长 $|I_k| < \epsilon$, 则存在唯一的一点 $C \in \bigcap_{k=1}^{\infty} I_k$ 。(2)有限覆盖定理, 设 $S = \{X\}$ 是由一些集合 X 构成的集族, 如果 $Y \subset \bigcup_{X \in S} X$, 就说 S 覆盖集合 Y 若一个闭区间 $[a, b]$ 被一个开区间族 $\{X\}$ 所覆盖, 则存在 $\{X\}$ 中有限个开区间仍覆盖 $[a, b]$ 。(3)极限点定理: 设点 $P \in \mathbb{R}$ 的任意小的邻域包含 $X \subset \mathbb{R}$ 的一个无穷子集, 就称点 $P \in \mathbb{R}$ 为集合 X 的极限点。引理: 每个有界无穷集至少有一个极限点。(4)非空有上界函数集必有上确界。设集合 $E \subset \mathbb{R}$, 若 M 满足: (i) $\forall X \in E \Rightarrow X \leq M$; (ii) 若 $M^{2'}$ 是 E 的上界, 则 $M \leq M'$, M 为 E 的上确界, 记为 $M = \sup E$ 。下确界仿此。极限 $\lim$ 是微积分最基本的概念之一, 极限理论是微积分学的基础。

数列(序列)的极限 (1)给定数列 $\{X_n\}$ 及实数 a , 若 $\forall \epsilon > 0, \exists N \in \mathbb{N}$, 使当 $n > N$ 时, 有 $|X_n - a| < \epsilon$, 就称 a 为数列 $\{X_n\}$ 的极限, 此时称数列 $\{X_n\}$ 为收敛的, 或 $\{X_n\}$ 收敛于 a , 记为 $\lim_{n \rightarrow \infty} X_n = a$, 或 $X_n \rightarrow a (n \rightarrow \infty \text{ 时})$ 。(2)若对任意给定的 $M > 0, \exists N \in \mathbb{N}$, 使当 $n > N$ 时有 $X_n > M$, 就称 X_n 趋于正无穷大, 记为 $\lim_{n \rightarrow \infty} X_n = +\infty$ 。同理可定义 $\lim_{n \rightarrow \infty} X_n = -\infty, \lim_{n \rightarrow \infty} X_m = \infty$ 。

收敛数列的性质与运算法则: (1) $X_n \rightarrow a (n \rightarrow \infty)$ 则 a 是唯一的; (2) $\{X_n\}$ 收敛则必有界; (3) 设有两个数列 $\{X_n\}, \{y_n\}, X_n \rightarrow A, y_n \rightarrow B$, 则数列 $\{X_n \pm y_n\}, \{X_n y_n\}, \{X_n / y_n\}$ 分别收敛于 $A \pm B, AB, A/B, (B \neq 0)$ 。

数列收敛性的判别法则: (1)(柯西收敛原理): $\{X_n\}$ 收敛的充分必要条件是: $\forall \epsilon > 0, \exists N > 0$, 当 $n, n' > N$ 时, $|X_n - X_{n'}| < \epsilon$; (2)单调有界数列必有极限(对任意 $k \in \mathbb{N}, X_k \leq X_{k+1}$, 就称 $\{X_n\}$ 为单调不减数列; 若为严格不等式 $X_k < X_{k+1}$, 称 $\{X_n\}$ 严格单增数列。单

减数列仿此)。(3) $\{X_n\}$ 为单增数列, $\{y_n\}$ 为单减数列, 且 $X_n < y_n$, 又 $y_n - X_n \rightarrow 0 (n \rightarrow \infty)$, 则两个数列的极限存在且相等。(4) 若 $\{X_n\}$ 有极限 a , 则其任何子列 $X_{n_1}, X_{n_2}, \dots, X_{n_k}, \dots$ 有相同的极限 a , 子列是保持 $\{X_n\}$ 的原有次序的部分数列。(5) 任何有界数列必可选出有极限的子列。(6) 设 $X_n \leq y_n \leq z_n (n=1, 2, \dots)$, 又 $\lim_{n \rightarrow \infty} X_n = \lim_{n \rightarrow \infty} z_n = a$, 则 $\lim_{n \rightarrow \infty} y_n = a$ 。(7) (施笃兹定理) 对数列 $\{X_n/y_n\}$ 若 (i) $n \geq N_0 (N_0 \text{ 为某一自然数})$ 时, $y_{n+1} > y_n$; (ii) $\lim_{n \rightarrow \infty} y_n = +\infty$; (iii) $\lim_{n \rightarrow \infty} [X_{n+1} - X_n] / [y_{n+1} - y_n] = A$ (有限或 $\pm\infty$), 则 $\lim_{n \rightarrow \infty} X_n/y_n = A$ 。

【函数的极限】若对于 $\forall \varepsilon > 0, \exists \delta > 0$, 使当 $0 < |X - a| < \delta$ 时, 有 $|f(X) - A| < \varepsilon$, 则称函数 $f(X)$ 在点 a 的极限为 A , 记作 $\lim_{X \rightarrow a} f(X) = A$, 或 $f(X) \rightarrow A (X \rightarrow a)$ 。若 $\forall M > 0, \exists \delta > 0$, 使当 $0 < |X - a| < \delta$ 时, $f(X) > M$ 。则称 $f(X)$ 趋于正无穷大, 记作 $\lim_{X \rightarrow a} f(X) = +\infty$, 或 $f(X) \rightarrow +\infty (X \rightarrow a)$ 。仿此可定义 $f(X) \rightarrow -\infty (X \rightarrow a)$ 与 $f(X) \rightarrow \infty (X \rightarrow a)$ 。

单侧极限 在上述定义中, 限制自变量 $X > a (< a)$, 则称 A 为 $f(X)$ 当 $X \rightarrow a$ 时的右(左)极限, 记作 $\lim_{X \rightarrow a+} f(X) = A$, 或 $f(a+0)$ ($\lim_{X \rightarrow a-} f(X) = A$, 或 $f(a-0)$), 显然, 当且仅当 $f(a+0)$ 与 $f(a-0)$ 都存在且相等时 $\lim_{X \rightarrow a} f(X)$ 才存在。

极限的性质与运算规则: (1) 设 $f: E \rightarrow R, g: E \rightarrow R$ 均有极限 $f(X) \rightarrow A, g(X) \rightarrow B$ (当 $X \rightarrow a$ 时), 且 $A < B$, 则存在 a 的空心邻域 $U(a)$ 使当 $X \in U(a)$ 时, $f(X) < g(X)$ 。(2) 设 $f: E \rightarrow R, g: E \rightarrow R, h: E \rightarrow R, f(X) \leq g(X) \leq h(X)$, $\lim_{X \rightarrow a} f(X) = \lim_{X \rightarrow a} h(X) = c$, 则

$\lim_{X \rightarrow a} g(X) = c$ 。(3) 设 $\lim_{X \rightarrow a} f(X) = A$,

$\lim_{X \rightarrow a} g(X) = B$, 则 $\lim_{X \rightarrow a} \{f(X) \pm g(X)\}$

$= A \pm B$; $\lim_{X \rightarrow a} f(X)g(X) = AB$,

$\lim_{X \rightarrow a} f(X)/g(X) = A/B (B \neq 0)$ 。

函数极限存在的判别法则(1)(柯西准则), 当 $X \rightarrow a$ 时 $f(X)$ 存在极限的充分必要条件是: 对于 $\varepsilon > 0$, 存在 $\delta > 0$, 只要 $|X - a| < \delta$ 及 $|X' - a| < \delta$, 便有 $|f(X) - f(X')| < \varepsilon$; (2) 单调有界函数必有极限存在。(3) $X \rightarrow a$ 时 $f(X)$ 有极限 A 的充分必要条件是: 对任意收敛于 a 的数列 $\{X_n\} (n=1, 2, \dots)$ 都有 $\lim_{n \rightarrow \infty} f(X_n) = A$ 。

洛必达法则 (此法则在掌握导数计算后应用)(1) 设 $f(X)$, $g(X)$ 在 $0(a, \delta)$ 上连续, 且 $\lim_{X \rightarrow a} f(X) = \lim_{X \rightarrow a} g(X) = 0$, 在 $0(a, \delta)$ 上 $f'(X)$, $g'(X)$ 存在且 $g'(X) \neq 0$, $\lim_{X \rightarrow a} (f'(X)/g'(X))$ 存在(可以是无穷大), 则下式(*)

$\lim_{X \rightarrow a} f(X)/g(X) = \lim_{X \rightarrow a} f'(X)/g'(X)$ 成立; (2) 把 a 换成 $+\infty$ 或 $-\infty$ 时, 命题仍成立; (3) 设 $f(X)$, $g(X)$ 在 $0(a, \delta)$ 上连续且 $\lim_{X \rightarrow a} g(X) + \infty$ (或 $-\infty$), $f(X)$, $g(X)$ 在空心邻域 $0(a, \delta)$ 可微且 $g'(X) \neq 0$, $\lim_{X \rightarrow a} (f'(X)/g'(X))$ 存在(可以是无穷大), 则(*)式成立; (4) $0 \cdot \infty$ 型, $\infty - \infty$ 型等可化为 $0/0$ 型或 ∞/∞ 型来求极限。

两个重要的极限 $\lim_{X \rightarrow 0} \sin X / X = 1$, (式中的 X 要用弧度来度量);

$$\lim_{X \rightarrow \infty} (1 + \frac{1}{X})^X = e.$$

无穷小量与无穷大量: (1) 极限为零的量称为无穷小量。在同一过程中两个无穷小量之和仍为无穷小量, 无穷小量与有界变量之积为无穷小量。显然, 有极限的变量等于常数加无穷小量。(2)

极限为 ∞ 的变量称为无穷大量。(3)关系:(i) $f(X)$ 为无穷小量,则 $1/f(X)$ 为无穷大量,反之亦然。(ii)设 $g(X)$ 在空心邻域 $U_0(a)$ 上不为零, $f(X),g(X)$ 均为无穷小(大)量:当 $\lim_{X \rightarrow a} f(X)/g(X) = A \neq 0$ 时,记作 $f(X) \sim A g(X) (X \rightarrow a)$,称 $f(X)$ 与 $g(X)$ 为同阶的无穷小(大)量;当 $A=1$ 时,则称它们为等阶的无穷小(大)量;(iii)若 $f(X),g(X)$ 均为无穷小(大)量, $\lim_{X \rightarrow a} f(X)/g(X) = 0$,称 $f(X)$ 是比 $g(X)$ 高(低)阶的无穷小(大)量,(iv)若 $f(X),g(X)$ 均为无穷小(大)量, $\lim_{X \rightarrow a} f(X)(g(X))^k = A$,则称 $f(X)$ 为 $g(X)$ 的 k 阶无穷小(大)量。

【上极限和下极限】考虑数列 $\{S_n\} (n=1,2,\cdots)$, (1) 设 $\{S_n\}$ 有上界, 即 $S_n \leq M$, 于是 $\{S_{n+1}, S_{n+2}, \cdots\}$ 有上界, 从而上确界存在: 令 $M_n = \sup \{S_{n+1}, \cdots\}$ 。(a) 若 $\{M_n\}$ 收敛, 定义 $\{S_n\}$ 的上极限为 $\limsup_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} M_n$; (b) 若 $\{M_n\}$ 发散于负无穷大, 记 $\limsup_{n \rightarrow \infty} S_n = -\infty$ 。(2) 若 $\{S_n\}$ 没有上界, 记 $\limsup_{n \rightarrow \infty} S_n = \infty$ 。(3) 当 $\{S_n\}$ 收敛时, 必有 $\limsup_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} S_n$ 。(4) 设 $\{S_n\}$ 是有下界的实数列, 从而下确界存在: $m_n = \inf \{S_n, S_{n+1}, \cdots\}$, (a) 若 $\{m_n\}$ 收敛, 定义 $\{S_n\}$ 的下极限为 $\liminf_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} m_n$ 。(b) 若 $\{m_n\}$ 发散于无穷大, 记 $\liminf_{n \rightarrow \infty} S_n = \infty$ 。(5) 若 $\{S_n\}$ 没有下界, 记 $\liminf_{n \rightarrow \infty} S_n = -\infty$ 。(6) 当 $\{S_n\}$ 收敛时, 有 $\liminf_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} S_n$ 。(7) $\liminf_{n \rightarrow \infty} S_n \leq \limsup_{n \rightarrow \infty} S_n$ 。(8) 设 $\{S_n\}$ 是有界的实数列, 且 $S_n \leq t_n$ 则 $\liminf_{n \rightarrow \infty} S_n \leq \limsup_{n \rightarrow \infty} t_n$; $\liminf_{n \rightarrow \infty} S_n \leq \liminf_{n \rightarrow \infty} t_n$ 。(9) $\limsup_{n \rightarrow \infty} (S_n + t_n) \leq \limsup_{n \rightarrow \infty} S_n + \limsup_{n \rightarrow \infty} t_n$ 。

$$\liminf_{n \rightarrow \infty} (S_n + t_n) \geq \liminf_{n \rightarrow \infty} S_n + \liminf_{n \rightarrow \infty} t_n。$$

【函数的连续性】设函数 $f(X)$ 在点 a 的某邻域内有定义。若 $\lim_{X \rightarrow a} f(X) = f(a)$, 就称 $f(X)$ 在点 a 连续, 称 a 为 $f(X)$ 的连续点。当 $f(X)$ 在区间 X 内的每点都连续时, 就称 $f(X)$ 在 X 内连续。记作 $f(X) \in C(X)$ 。 X 可以是开区间、半开区间或闭区间。 $f(X)$ 在点 a 的连续蕴含下面三个式子成立:

$$\lim_{X \rightarrow a+0} f(X) = f(a+0) \quad (1)$$

$$\lim_{X \rightarrow a-0} f(X) = f(a-0)$$

$$f(a+0) = f(a-0) \quad (2)$$

$$f(a+0) = f(a) = f(a-0) \quad (3)$$

当只有(1),(2)成立时, 称 a 为可去间断点; 当只有(1)式成立时, 称 a 为第一类间断点, (1)式不成立, 称为第二类间断点。若只有 $\lim_{X \rightarrow a-0} f(X) = f(a)$ 成立, 就称 $f(X)$ 在点 a 为左连续的, 若 $\lim_{X \rightarrow a+0} f(X) = f(a)$ 成立, 则为右连续的。若 $f(X), g(X)$ 都在 X 内连续, 则 $f(X) \pm g(X), f(X) \cdot g(X)$ 与 $f(X)/g(X), (g(X) \neq 0)$ 都在 X 内连续。连续的复合函数仍为连续函数, 若 $u = g(X)$ 在点 X_0 连续, $g(X_0) = u_0, y = f(u)$ 在 u_0 连续, 则复合函数 $y = f(g(X))$ 在 X_0 点连续。基本初等函数在其定义域内都连续。

一致连续性: 设 $f(X)$ 在 X 上满足: 对于任意的 $\varepsilon > 0$, 存在 $\delta = \delta(\varepsilon) > 0$, 使对于一切 $X_1, X_2 \in X$, 当 $|X_1 - X_2| \leq \delta$ 时, 总有 $|f(X_1) - f(X_2)| < \varepsilon$, 就称 $f(X)$ 在 X 内一致连续。

闭区间 $[a, b]$ 上连续函数的性质: $f(X)$ 在 $[a, b]$ 上连续, 则(1) (康托尔定理): $f(X)$ 必于 $[a, b]$ 上一致连续; (2) 必有界; (3) 必能达到上确界与下确界; (4) 设 $f(a) = A, f(b) = B$, 对于 A 与 B 之间的任一数 C , 存在 $c \in [a, b]$, 使 $f(c) = C$ 。特别, 若 $f(a)f(b) < 0$, 则存在 $c \in (a, b)$ 使 $f(c) = 0$; (5) 严格单调函数的反函数存在且为

严格单调的连续函数。

微 分 学

【微分学】曲线的切线。设函数 $y=f(X)$ 的图象为 l , $P(X_0, y_0)$ 是 l 上的一个点。欲求点 P 处 l 的切线, 作割线 PQ , Q 的坐标是 $(X_0 + \Delta X, y_0 + \Delta y)$, PQ 与 X 轴夹角为 ψ , $\operatorname{tg}\psi = RQ/PR = [f(X_0 + \Delta X) - f(X_0)]/\Delta X$, 令 $\Delta X \rightarrow 0$, 得(在极限存在的条件下) $\operatorname{tg}\theta = \lim_{\Delta X \rightarrow 0} \Delta y/\Delta X$, $\operatorname{tg}\theta$ 就定义为切线 PT 在 P 点处的斜率。

导数 设给定函数 $y=f(X)$, 在点 X_0 处给 X 以改变量 ΔX , 相应地 $f(X)$ 的改变量为 $\Delta y = f(X + \Delta X) - f(X)$, 若极限 $\lim_{\Delta X \rightarrow 0} \Delta y/\Delta X$ 存在, 就称此极限为 $f(X)$ 在点 X 的导数, 记作

$f'(X)$ 或 $\frac{dy(X)}{dX}$ 。若在区间 X 内的每点 X , 导数都存在, 就称 $f(X)$

在 X 内可导, 并称 $y' = f'(X) = \lim_{\Delta X \rightarrow 0} [f(X + \Delta X) - f(X)] / \Delta X$ 为函数 $y = f(X)$ 的导函数。它的几何意义是曲线切线的斜率。其物理意义可以是运动物体的瞬时速度: 设运动方程为 $s = s(t)$, 其中 s 表示距离, t 代表时间, 在 $[t, t + \Delta t]$ 时间段内的平均速度是 $[s(t + \Delta t) - s(t)] / \Delta t$, 令 $\Delta t \rightarrow 0$, 若极限 $v = \lim_{\Delta t \rightarrow 0} [s(t + \Delta t) - s(t)] / \Delta t$ 存在。就定义为 t 时的瞬时速度。

分别称 $\lim_{\Delta t \rightarrow 0-} [f(X + \Delta X) - f(t)] / \Delta X$ 与 $\lim_{\Delta t \rightarrow 0+} [f(X + \Delta X) - f(t)] / \Delta X$ 为 X 处的左、右导数。显然, $y = f(X)$ 在 X 可导的充分必要条件是 X 处的左、右导数都存在且相等。

基本初等函数的导数

函数 $f(X)$	导数 $f'(X)$
C (常数)	0
X^α	$\alpha X^{\alpha-1}$
a^X	$a^X \ln a$
$\log_a X $	$1/(X \ln a)$
$\sin X$	$\cos X$
$\cos X$	$-\sin X$
$\operatorname{tg} X$	$1/\cos^2 X$
$\operatorname{ctg} X$	$-1/\sin^2 X$
$\arcsin X$	$1/\sqrt{1-X^2}$
$\arccos X$	$-1/\sqrt{1-X^2}$
$\operatorname{arctg} X$	$\frac{1}{1+X^2}$
$\operatorname{sh} X$	ch^X
$\operatorname{ch} X$	$\operatorname{sh} X$

$\operatorname{th} X$	$1/\operatorname{ch}^2 X$
cth	$-1/\operatorname{sh}^2 X$
$\operatorname{arch} X = \ln(s + \sqrt{1+X^2})$	$1/\sqrt{1+X^2}$
$\operatorname{arch} X = \ln(X \pm \sqrt{X^2-1})$	$\pm 1/\sqrt{X^2-1}$
$\operatorname{arth} X = \frac{1}{2} \ln \frac{1+X}{1-X}$	$\frac{1}{1-X^2}$
$\operatorname{arcth} X = \frac{1}{2} \ln \frac{X+1}{X-1}$	$\frac{1}{1-X^2}$
$\operatorname{arclg} X$	$-\frac{1}{1+X^2}$

导数的运算法则

设 $f(X)$ 与 $g(X)$ 都可导, 则

$$\begin{aligned} [f(X) \pm g(X)]' &= f'(X) \pm g'(X), \\ [f(X)g(X)]' &= f'(X)g(X) + f(X)g'(X), \\ \left[\frac{f(X)}{g(X)} \right]' &= \frac{f'(X)g(X) + f(X)g'(X)}{g^2(X)} \\ &\quad (g(X) \neq 0) \end{aligned}$$

参数式给出的函数 $X=X(t)$, $y=y(t)$, 则

$$\frac{ay}{aX} = \frac{y'(t)}{X'(t)} \quad (X'(t) \neq 0)。$$

$y=f(X)$, 其反函数 $X=\varphi(y)$, 则

$$\varphi'(y) = 1/f'(X)$$

$y=f(u)$, $u=g(X)$, 则 $y=f(g(X))$,

$$\frac{dy}{dX} = f'(u)g'(X)。$$

导数运算的链式法则 设 $u=\varphi(X)$, 在 X 点可导, 在对应点

$u, y=f(u)$ 也可导, 则复合函数 $y=f(\varphi(X))$ 在 X 点可导, 且 $\frac{dy}{dX} = \frac{dy}{du} \cdot \frac{du}{dX} = f'(u) \cdot \varphi'(X)$ 。

微分 一个变量在变化过程中的线性主要部分称为微分。给定函数 $y=f(X)$, 若自变量的改变量 ΔX , 如果 $\Delta y = A\Delta X + o(\Delta X)$ ($\Delta X \rightarrow 0$), 其中 A 是 X 的函数, 与 ΔX 无关, 此时 $A\Delta X$ 就是 $f(X)$ 在点 X 的微分, 记作 $dy=df(X)=A\Delta X=AdX$, (自变量的微分 $dX=\Delta X$)。若 $y=f(X)$ 在点 X 可导, 则 $dy=f'(X)dX$ 可导与可微互为必要条件, 由求导法则便得到相应的微分法则。

一阶微分形式不变性: 设有复合函数 $y=f(u), u=g(X)$, 则 $dy=f'(g(X))g'(X)dX=f'(u)du$, 这与 $dy=f'(X)dX$ 形式相同, 称为一阶微分的形式不变性。

高阶导数与高阶微分 将一阶导函数 $y'=f'(X)$ 再求导, 便得二阶导数。3 阶, $\dots, n$ 阶导数, 可相似定义并分别记为 $y''=f''(X), y'''=f'''(X), \dots$, 或 $y^{(2)}=f^{(2)}(X), y^{(3)}=f^{(3)}(X), \dots, y^{(n)}=f^{(n)}(X), \dots$, 从四阶开始不再使用“'”来表示高阶导数。计算法则:

$$\begin{aligned} [f(X) \pm g(X)]^{(n)} &= f^{(n)}(X) \pm g^{(n)}(X) \\ [f(X) \cdot g(X)]^{(n)} &= f^{(n)}(X) \\ &+ C_n^1 f^{(n-1)}(X) g'(X) + C_n^2 f^{(n-2)}(X) \cdot g''(X) + \dots + C_n^k f^{(n-k)}(X) \\ &g^{(k)}(X) \\ &+ \dots + f(X) g^{(n)}(X) \end{aligned} \quad (4)$$

(4) 式便是著名的莱布尼兹公式。

高阶微分: 一阶微分的微分, 称为原来函数的二阶微分, 记为 $d^2y=d(dy)$; 一般地, $d^{(n+1)}y=d(d^{(n)}y)$ 。关于乘积函数 $u(X)v(X)$ 的 n 阶微分, 有莱布尼兹公式:

$$d^{(n)}(u \cdot v) = \sum_{i=0}^n \begin{bmatrix} n \\ i \end{bmatrix} d^{n-i} u d^{(i)} v; (d^0 u = u, d^0 v = v)。$$

【微分中值定理】若 $f(X)$ 在点 X_0 的某邻域 $O(X_0, \delta)$ 内有定义, 且 $f(X) \leq$

$f(X_0)$ (或 $f(X) \geq f(X_0)$), 就称点 X_0 为 $f(X)$ 的局部极大 (或极小) 值点, 相应地, 称 X_0 处的函数值 $f(X_0)$ 为极大 (或极小) 值。

费马定理: 若 $f(X)$ 在极大值 (或极小值) 点 X_0 的某邻域内可微, 则 $f'(X_0) = 0$ 。

罗尔定理: 若 $f(X)$ 在 $[a, b]$ 上连续, 在 (a, b) 可微, 且 $f(a) = f(b)$, 则存在点 ζ 属于 (a, b) , 使 $f'(\zeta) = 0$ 。

拉格朗日定理: 若 $f(X)$ 在 $[a, b]$ 上连续, 在 (a, b) 可微, 则存在点 $\zeta \in (a, b)$, 使 $f(b) - f(a) = f'(\zeta)(b - a)$ 。

柯西定理: 若 $f(X)$ 和 $g(X)$ 在 $[a, b]$ 连续, 在 (a, b) 可微, $g'(X) \neq 0$, 则存在 $\zeta \in (a, b)$ 使 $[f(b) - f(a)] / [g(b) - g(a)] = f'(\zeta) / g'(\zeta)$ 。

达布定理: $f(X)$ 在 $[a, b]$ 上可微, 则 $f'(X)$ 的值域为一个区间。

泰勒公式: 把任意函数表为某点邻域内的近似多项式。设 $f(X)$ 在含有 X_0 的闭区间上连同其 n 阶导数都连续, 并有 $n+1$ 阶导数, 则 $f(X) = f(X_0) + f'(X_0)(X - X_0) + \frac{1}{2!} f''(X_0)(X - X_0)^2 + \cdots + \frac{1}{n!} f^{(n)}(X_0)(X - X_0)^n + r_n(X, X_0)$, 称为泰勒公式, 其中的 $r_n(X, X_0)$ 叫做 n 阶余项。余项的几种形式:

皮亚诺形式: $r_n(X, X_0) = o(|X - X_0|^n)$

柯西形式: $r_n = \frac{1}{n!} f^{n+1}(X_0 + \theta(X - X_0))(1 - \theta)^n (X - X_0)^{n+1}$
 $(0 < \theta < 1)$

拉格朗日形式: $r_n = \frac{1}{(n+1)!} f^{n+1}(\zeta) \cdot (X - X_0)^{n+1}$ (ζ 在 X_0 与 X 之间)。

利用导数来研究函数的动态, 设

$f(X)$ 在 $[a, b]$ 连续, 在 (a, b) 有有限导数 $f'(X)$, 则: (1) $f(X) \equiv$ 常数 $\Leftrightarrow f'(X) \equiv 0, X \in (a, b)$; (2) $f(X)$ 在 $[a, b]$ 内单调增大(减小) $\Leftrightarrow f'(X) \geq 0 (\leq 0)$; $f(X)$ 在 $[a, b]$ 内严格单增(减) $\Leftrightarrow f'(X) \geq 0 (\leq 0)$ 且 $f'(X)$ 在 (a, b) 的任一部分区间内不恒等于零; (3) 当 $X < X_0$ 时 $f'(X) > 0$, 而当 $X > X_0$ 时 $f'(X) < 0$, 则 $f(X_0)$ 取得局部极大值。(4) $f'(X_0) = 0, f''(X_0)$ 存在且 $f''(X_0) < 0$, 则 $f(X_0)$ 是局部极大值。极小值仿此。

平面点集的几个基本定理: (1) 矩形套定理: 记闭矩形列 $\triangle_n: a_n \leq X \leq b_n, c_n \leq y \leq d_n (n = 1, 2, \dots)$ 具: $\triangle_1 \supset \triangle_2 \supset \dots \supset \triangle_n \supset \dots, b_n - a_n \xrightarrow[n \rightarrow \infty]{d_n - c_n} 0$, 则存在唯一的一点 $P_o(X_o, y_o)$, 含于每一个矩形内, 即 $a_n \leq X_o \leq b_n, c_n \leq y_o \leq d_n (n = 1, 2, \dots)$ 。(2) 致密性定理: 有界序列 $\{P_n(X_n, y_n)\}$ 中必可选出收敛的子列。(3) 有限覆盖定理: 若有一个开矩形集 $\{\triangle\} = \{(\alpha < X < \beta, \gamma < y < \delta)\}$ 覆盖一有界闭域 $\bar{D}$, 则可从 $\{\triangle\}$ 中选出有限个开矩形也覆盖 $\bar{D}$ 。

二元函数: 设 E 为平面点集, R 为实数集, f 是一个规律, 若对 E 中的每点 (X, y) , 通过规律 f , 有 R 中唯一的实数 u 与之对应, 就称 f 是定义在 E 上的一个二元函数, $(X, y) \rightarrow u = f(X, y)$ 。

【二元函数的极限】若对于任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 使当 $|X - X_0| <$

$\delta, |y - y_0| < \delta, (X, y) \neq (X_0, y_0)$ 时, 总有 $|f(X, y) - A| < \varepsilon$, 就称数 A 为 $f(X, y)$ 在点 (X, y) 趋于 (X_0, y_0) 时的极限, 记作 $\lim_{\substack{X \rightarrow X_0 \\ y \rightarrow y_0}} f(X, y) =$

A 或 $\lim_{P \rightarrow P_0} f(P) = A$, 称之为二重极限, 即 X, y 同时趋于各自的极限时 $f(X, y)$ 的极限。这与二次极限不同: 二次极限是先 X 后 y (或先 y 后 X) 相继取的极限。 $\lim_{y \rightarrow y_0} \lim_{X \rightarrow X_0} f(X, y)$, 但有定理: 若 $f(X, y)$ 在点 (X_0, y_0) 的二重极限为

$\lim_{\substack{X \rightarrow X_0 \\ y \rightarrow y_0}} f(X, y) = A$, 且对任一靠近 y_0 的 y , 当 $X \rightarrow X_0$ 时, $f(X, y)$ 存在

有限极限 $\varphi(y) = \lim_{X \rightarrow X_0} f(X, y)$, 则二次极限

$\lim_{y \rightarrow y_0} \lim_{X \rightarrow X_0} f(X, y)$ 存在且等于二重极限 A 。

连续 称函数 $f(X, y)$ 在点

(X_0, y_0) 连续, 若 $\lim_{P \rightarrow P_0} f(X, y) = f(X_0, y_0)$ 。当 $f(X, y)$ 在区域 D 的每点 (X, y) 都连续时, 就称 $f(X, y)$ 为 D 上的连续函数。

有界闭区域 $\bar{D}$ 上连续函数的性质: (1) 存在 $\delta > 0$ 必有界。(2) 必一致连续; 即对于 $\forall \varepsilon > 0$, 使对 $\bar{D}$ 上的任意两点 $P_1(X_1, y_1)$ 与 $P_2(X_2, y_2)$, 当 $\rho(P_1, P_2) < \delta$ 时, 恒有

$|f(X_1, y_1) - f(X_2, y_2)| < \varepsilon$ 。(3) 在 $\bar{D}$ 上可取得最大值 M 与最小值 m 。(4) 零点存在定理: 设在 P_1 与 P_2 有 $f(X_1, y_1) \cdot f(X_2, y_2) < 0$, 则存在 $(X, \bar{y})$ 使 $f(X, \bar{y}) = 0$ 。

偏导数 二元函数 $u = f(X, y)$, 固定 y , 给 X 以增量 ΔX , 若极限

$\lim_{\Delta X \rightarrow 0} \frac{\Delta_X y}{\Delta X} = \lim_{\Delta X \rightarrow 0} [f(X + \Delta X, y) - f(X, y)] / \Delta X$ 存在, 就称此

极限值为 $f(X, y)$ 在点 (X, y) 处关于 X 的偏导数, 记作 $\frac{\partial y}{\partial X}$ 或 $\frac{\partial f}{\partial X} f'$

$X(X, y)$, 仿此定义 $\frac{\partial \mu}{\partial y}$ 。对偏导数再求偏导数, 就得到二阶偏导数 $f''_{xy} = \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial X} \right) = \frac{\partial^2 u}{\partial y \partial X}$, $f''_{yx} = \frac{\partial^2 u}{\partial X \partial y}$, 当 f''_{xy} 与 f''_{yx} 在点 (X, y) 都连续, 则有 $f''_{yx} \equiv f''_{xy}$ 。继续求偏导数可得更高阶的偏导数。

方向导数 多元函数沿某个特定方向的变化率(以三元函数为例), 定义为: 设 P_0 为函数 $u = f(X, y, z)$ 的定义域 E 内的一点, l 为从点 P_0 出发的射线, $P(X, y, z)$ 为 l 上且含于 E 内的任一点, 以 ρ 表示 P 与 P_0 两点间的距离, 若

$$\lim_{\rho \rightarrow 0} \frac{f(P) - f(P_0)}{\rho} = \lim_{\rho \rightarrow 0} \frac{\Delta_l f}{\rho}$$

存在, 则称此极限值为 f 在点 P_0 沿方向 l 的方向导数, 记作 $\frac{\partial f}{\partial l} \big|_{P_0}$ 或 $f_l(P_0)$ 或 $f_l(X_0, y_0, z_0)$ 。

若函数 $f(X, y, z)$ 在点 $p(X, y, z)$ 可微(参见“全微分”), 则函数在 P 点处沿任一方向 l 的方向导数都存在, 且

$$\frac{\partial f}{\partial l} = \frac{\partial f}{\partial X} \cos \alpha + \frac{\partial f}{\partial y} \cos \beta + \frac{\partial f}{\partial z} \cos \gamma, \text{ 其中 } \cos \alpha, \cos \beta, \cos \gamma \text{ 是 } l \text{ 的}$$

方向余弦。

极值 设函数 $u = f(p) = f(X_1, X_2, \dots, X_n)$ 在点 P_0 的某邻域 E 内有定义, 若对于 $\forall p \in E$ 有 $f(p) \leqslant$ (或 $\geqslant$) $f(P_0)$, 则称 $u = f(p)$ 在点 P_0 取得局部极大(或极小)值, 称 P_0 为极值点。点 $p(X_1, X_2, \dots, X_n)$ 是多元函数 $f(X_1, X_2, \dots, X_n)$ 的极值点的必要条件是 P 满足方程组 $\frac{\partial f}{\partial X_1} = \frac{\partial f}{\partial X_2} = \dots = \frac{\partial f}{\partial X_n} = 0$, 寻求充分条件一般是用泰勒展开式, 由差 Δf 所余下的最低次项来决定。这同一元函数是一样的。对于条件极值, 常用的方法是拉格朗日乘子法。如求 $f(X, y, u, v)$ 在约束 $g(X, y, u, v) = 0$ 和 $h(X, y, u, v) = 0$ 之下的极值, 以 α ,

β 为参数, 作函数 $L(X, y, u, v) = f(X, y, u, v) + \alpha g(X, y, u, v) + \beta h(X, y, u, v)$, 可得另外四个条件:

$$\frac{\partial f}{\partial u} + \alpha \frac{\partial g}{\partial u} + \beta \frac{\partial h}{\partial u} = 0,$$

$$\frac{\partial f}{\partial v} + \alpha \frac{\partial g}{\partial v} + \beta \frac{\partial h}{\partial v} = 0,$$

$$\frac{\partial f}{\partial X} + \alpha \frac{\partial g}{\partial X} + \beta \frac{\partial h}{\partial X} = 0,$$

$$\frac{\partial f}{\partial y} + \alpha \frac{\partial g}{\partial y} + \beta \frac{\partial h}{\partial y} = 0,$$

充分性条件的探求原则同上述。

【全微分】二元函数 $u = f(X, y)$, 当所有自变量改变时的变化量 $\Delta u = f(X + \Delta X, y + \Delta y) - f(X, y) = A\Delta X + B\Delta y + o(\Delta r)$ ($\Delta r = \sqrt{\Delta X^2 + \Delta y^2}$)。 A 和 B 仅与 X, y 有关(与 $\Delta X, \Delta y$ 无关), 则称函数 $f(X, y)$ 在 (X, y) 可微, 并称 Δu 的线性主部 $A\Delta X + B\Delta y$ 为 $f(X, y)$ 在点 (X, y) 的全微分, 记作 du , 经计算可知 $A = \frac{\partial f}{\partial X}, B = \frac{\partial f}{\partial y}$ 从而 $du = f_x dX + f_y dy$ 。当 f_x, f_y 在 (X, y) 都存在且连续时, 则 $u = f(X, y)$ 在 (X, y) 可微, 对 du 再求微分, 即得二阶微分 $d^2u = d(du) = d[f_x dX + f_y dy] = f_{xx} dX^2 + 2f_{xy} dX dy + f_{yy} dy^2$, 继续下去可得三阶、四阶、……的微分。用数学归纳法可证: $d^n u = \sum_{k=0}^n c_n^k \cdot \frac{\partial^k f}{\partial X^{n-k} \partial y^k} dX^{n-k} dy^k$ 。

复合函数求导的链式法则: 设 $u = f(X, y)$ 而 $X = \varphi(s, t), y = \psi(s, t)$, 则 $\frac{\partial u}{\partial t} = \frac{\partial u}{\partial X} \frac{\partial X}{\partial t} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial t}, \frac{\partial u}{\partial s} = \frac{\partial u}{\partial X} \frac{\partial X}{\partial s} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial s}$, 此时, $du = \frac{\partial u}{\partial s} ds + \frac{\partial u}{\partial t} dt = (\frac{\partial u}{\partial X} \cdot \frac{\partial X}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s}) ds + (\frac{\partial u}{\partial X} \cdot \frac{\partial X}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}) dt = \frac{\partial u}{\partial X} dX + \frac{\partial u}{\partial y} dy$ 。

【雅可比矩阵】 m 个 n 元函数 $u_i = u_i(X_1, \cdots, X_n) (i=1, 2, \cdots, n)$ 的偏导数所组成的矩阵

$$(J) = (u'_{ij}) = \begin{pmatrix} \frac{\partial u_1}{\partial X_1} & \frac{\partial u_1}{\partial X_2} & \cdots & \frac{\partial u_1}{\partial X_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u_m}{\partial X_1} & \frac{\partial u_m}{\partial X_2} & \cdots & \frac{\partial u_m}{\partial X_n} \end{pmatrix}$$

叫雅可比矩阵, 当 $m=n$ 时成为 n 阶方阵, 其行列式

$$J = \begin{vmatrix} \frac{\partial u_1}{\partial X_1} & \frac{\partial u_1}{\partial X_2} & \cdots & \frac{\partial u_1}{\partial X_n} \\ \frac{\partial u_2}{\partial X_1} & \frac{\partial u_2}{\partial X_2} & \cdots & \frac{\partial u_2}{\partial X_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u_n}{\partial X_1} & \frac{\partial u_n}{\partial X_2} & \cdots & \frac{\partial u_n}{\partial X_n} \end{vmatrix} \\ = \frac{D(u_1, u_2, \cdots, u_n)}{D(X_1, X_2, \cdots, X_n)}$$

称为雅可比行列式, 由全微分式

$$du_i = \frac{\partial u_i}{\partial X_1} dX_1 + \frac{\partial u_i}{\partial X_2} dX_2 + \cdots + \frac{\partial u_i}{\partial X_n} dX_n \quad (i=1, 2, \cdots, n) \text{ 及向量表示}$$

$$\begin{pmatrix} du_1 \\ du_2 \\ \vdots \\ du_n \end{pmatrix} = (J) \begin{pmatrix} dX_1 \\ dX_2 \\ \vdots \\ dX_n \end{pmatrix}$$

可以看出, (J) 的作用类似于一元函数的导数, 雅可比行列式具有以下性质:

$$\frac{D(u_1, \cdots, u_n)}{D(X_1, \cdots, X_n)} = \frac{D(u_1, \cdots, u_n)}{D(h, \cdots, t_n)} \cdot \frac{D(t_1, \cdots, t_n)}{D(X_1, \cdots, X_n)}$$

及

$$\frac{D(u_1, \cdots, u_n)}{D(X_1, \cdots, X_n)} \cdot \frac{D(X_1, \cdots, X_n)}{D(u_1, \cdots, u_n)} = 1$$

隐函数存在定理:(1)设函数 $F(X, y)$ 在点 $p(X_0, y_0)$ 某邻域内连续且有连续偏导数 $F_X(X, y), F_y(X, y)$,又设 $F(X_0, y_0) = 0, F_y(X_0, y_0) \neq 0$,则存在唯一的函数 $y = f(X)$,它在点 $X = X_0$ 的某一邻域内是单值连续可微的,恒满足 $F[X, f(X)] \equiv 0$,且 $y_0 = f(X_0), \frac{dy}{dX} = -F_X/F_y$ 。(2)函数 $F(X_1, X_2, \cdots, X_n, y) = 0$,确定隐函数 $y = f(X_1, X_2, \cdots, X_n)$ 的条件和有关结论与(1)类似。(3)设 $F(X, y, u, v) = 0, G(X, y, u, v) = 0$ 连续可微, $F(X_0, y_0) = 0, G(X_0, y_0) = 0, J = D(F, G)/D(u, v) \neq 0$,则在 (X_0, y_0) 的某邻域 $\triangle$ 内, $F(X, y, u, v) = 0, G(X, y, u, v) = 0$ 唯一确定一组连续可微函数 $u = u(X, y), v = v(X, y)$,即 $F(X, y, u(X, y), v(X, y)) \equiv 0, G(X, y, u(X, y), v(X, y)) \equiv 0$,且 $u_0 = u(X_0, y_0), v_0 = v(X_0, y_0)$ 。

$$\begin{aligned} \frac{\partial u}{\partial X} &= -\frac{1}{J} \frac{D(F, G)}{D(X, v)}, \\ \frac{\partial v}{\partial X} &= -\frac{1}{J} \cdot \frac{D(F, G)}{D(u, X)}, \\ \frac{\partial u}{\partial y} &= -\frac{1}{J} \cdot \frac{D(F, G)}{D(y, v)}, \\ \frac{\partial v}{\partial y} &= -\frac{1}{J} \cdot \frac{D(F, G)}{D(u, y)}. \end{aligned}$$

积 分 学

【积分学】与微分学一起构成了分析数学最早也是最基础的分支。主

要内容是一元函数的定积分与不定积分。微积分基本定理(牛顿-莱布尼茨公式)建立了微分与积分的联系。

曲边梯形的面积:要计算曲线 $y=f(X)$, 直线 $X=a$ 和 $X=b$, X 轴所围成的图形的面积, 用一组分点

$a=X_0<X_1<X_2<\cdots<X_{n-1}<X_n=b$ (1) 把区间 $[a, b]$ 分成 n 个小区间 $[X_{i-1}, X_i]$ $i=1, 2, \dots, n$, 在每个分点 X_i 处作 X 轴的垂线, 于是得到 n 个小的曲边梯形: 由两垂线 $X=X_{i-1}$ 与 $X=X_i$, $\overline{X_{i-1}X_i}$ 和 $y=f(X)$ 的在 $[X_{i-1}, X_i]$ 中的一小段曲线所围成, 如图中阴影部分所示。任取 $\zeta_i \in [X_{i-1}, X_i]$ 。当分的很细时, 矩形面积 $f(\zeta_i)(X_i - X_{i-1})$ 与该曲边梯形面积相近似, 作和数

$\sigma = \sum_{i=1}^n f(\zeta_i)(X_i - X_{i-1})$, 令 $n \rightarrow \infty$, 且各小区间长度均趋于零, 若极限 $\lim_{n \rightarrow \infty} \sigma$ 存在, 就称它为该曲边梯形的面积。

定积分定义: 设 $f(X)$ 是定义在 $[a, b]$ 上的函数, 用形如 (1) 的分点组将 $[a, b]$ 分划。记此分法为 Δ , 在每个小区间 $[X_{i-1}, X_i]$ 任取点 ζ_i , 作和式 $\sigma = \sum_{i=1}^n f(\zeta_i)(X_i - X_{i-1}) = \sum_{i=1}^n f(\zeta_i) \Delta X_i$, 设 $\lambda(\Delta) = \max_{1 \leq i \leq n} \{\Delta X_i\}$, 若极限 $\lim_{\lambda(\Delta) \rightarrow 0} \sigma$ 存在, 就称它为 $f(X)$ 在 $[a, b]$ 上的定积分, 记作 $J = \int_a^b f(X) dX$, a 和 b 分别称为积分的下限和上限, 称 σ

为黎曼和。

积分的存在性:达布和方法,令 $M_i = \sup \{f(X) \mid X \in [X_{i-1}, X_i]\}$, $m_i = \inf \{f(X) \mid X \in [X_{i-1}, X_i]\}$, $\Delta X_i = X_i - X_{i-1}$ ($i = 1, 2, \dots, n$), $\bar{S} = \sum M_i \Delta X_i$, 称为达布上和, $\underline{S} = \sum m_i \Delta X_i$, 称为达布下和, 它们具如下性质:插入新的分点, $\bar{S}$ 不增, $\underline{S}$ 不减; $\bar{S} \geq m(b-a)$, $\underline{S} \leq M(b-a)$, 其中 $M = \sup \{f(X) \mid X \in [a, b]\}$, $m = \inf \{f(X) \mid X \in [a, b]\}$; 任一分法下的 $\underline{S}$ 小于等于任一分法的 $\bar{S}$; 若 $f(X)$ 为有界函数, 则 $\lim_{X \rightarrow 0} \bar{S} = L$ 和 $\lim_{X \rightarrow 0} \underline{S} = l$ 都存在。于是有:

$f(X)$ 可积 $\Leftrightarrow L = l$; 令 $w_i = M_i - m_i$, 则

$f(X)$ 可积 $\Leftrightarrow \forall \epsilon > 0, \delta > 0$, 使 $w_i \geq \epsilon$ 的所有区间的长度之和 $\sum \Delta X_i < \delta$ 。可积函数类: $[a, b]$ 上的连续函数可积, $[a, b]$ 上的分段连续函数可积。

积分的性质与运算: 若在 $[a, b]$ 上 $f(X), g(X)$ 都可积, 则 (1) $\int_a^b [\alpha f(X) + \beta g(X)] dX = \alpha \int_a^b f(X) dX + \beta \int_a^b g(X) dX$, 其中 α 和 β 为常数; (2) $f(X) \cdot g(X)$ 可积; (3) $|f(X)|$ 可积, 且 $\int_a^b f(X) dX \leq \int_a^b |f(X)| dX$; (4) $\int_a^c f(X) dX = \int_a^b f(X) dX + \int_b^c f(X) dX$, 这里可不计 a, b, c 的顺序; (5) $F(X) = \int_a^X f(X) dX$ 是 $[a, b]$ 上的连续函数 ($X \in [a, b]$)。

积分中值定理: (1) 设在 $[a, b]$ 上,

$f(X)$ 连续, $g(X)$ 不变号且可积, 则存在点 $\zeta \in [a, b]$, 使 $\int_a^b f(X) g(X) dX = f(\zeta) \int_a^b g(X) dX$ 。(2) $f(X)$ 可积, $g(X)$ 单调, 则存在点 $\zeta \in [a, b]$ 使 $\int_a^b f(X) \cdot g(X) dX = f(a) \int_a^\zeta g(X) dX + f(b) \int_\zeta^b g(X) dX$

$$\int_{\zeta}^b g(X) dX。$$

【不定积分】若在某区间上, $F'(X) =$

$f(X)$, 就称函数 $F(X)$ 为 $f(X)$ 的一个原函数, 显然此时 $F(X) + c$ 也是其原函数 (c 为常数), 它的一般表达式称为 $f(X)$ 的不定积分, 记为 $\int f(X) dX$, 即 $\int f(X) dX = F(X) + c$ 。

牛顿—莱布尼兹公式 若 $f(X)$ 在 $[a, b]$ 上连续, $F(X)$ 是它的任一原函数, 则 $\int_a^b f(X) dX = F(b) - F(a)$ 。

不定积分简表

$$\int k dX = kX + c \quad (k \text{ 为常数})$$

$$\int X^n dX = \frac{1}{n+1} X^{n+1}$$

$$\int \frac{1}{X} dX = \ln |X| + c$$

$$\int e^X dX = e^X + c$$

$$\int a^X dX = \frac{1}{\ln a} a^X \quad (a > 0)$$

$$\int \sin X dX = -\cos X + c$$

$$\int \cos X dX = \sin X + c$$

$$\int \tan X dX = -\ln |\cos X| + c$$

$$\int \cot X dX = \ln |\sin X| + c$$

$$\int shXdX = chX + c$$

$$\int chXdX = shX + c$$

$$\int thXdX = lnchX + c$$

$$\int cthXdX = lnshX + c$$

$$\int \frac{1}{a^2 + X^2} dX = \frac{1}{a} \arctg \frac{X}{a} + c (a > 0)$$

$$\int \frac{1}{a^2 - X^2} dX = \frac{1}{2a} \ln \frac{a+X}{a-X} + c$$

$$(|X| < |a|)$$

$$\int \frac{1}{X^2 - a^2} dX = \frac{1}{2a} \ln \frac{X-a}{X+a} + c$$

$$(|X| > |a|)$$

$$\int \frac{1}{\sqrt{a^2 - X^2}} = \arcsin \frac{X}{a} + c$$

$$(|X| > |a|)$$

$$\int \frac{dX}{\sqrt{X^2 - a^2}} = \ln(X + \sqrt{X^2 - a^2}) + c$$

$$= \operatorname{arch} \frac{X}{a} + c (|X| > |a|)$$

运算法则: α, β 为常数, 则 $\int [\alpha f(X) \pm \beta g(X)] dX = \alpha \int f(X) dX + \beta \int g(X) dX$ 。

积分方法: (1) 换元积分法: 令 $X = \varphi(t)$, 若 $f(X), \varphi'(t), t = \varphi^{-1}(X)$ 都连续, 则 $\int f(X) dX = \int f(\varphi(t)) \varphi'(t) \cdot dt$ 。对定积分: 设当 $\alpha \leq t \leq \beta$ 时, $a \leq \varphi(t) \leq b, \varphi(\alpha) = a, \varphi(\beta) = b$, 则

$\int_a^b f(X) dX = \int_a^{\beta} f(\varphi(t)) \varphi'(t) dt$; (2) 分部积分法: 若 $u(X), v(X)$ 可

微, $u'v$ 和 uv' 中至少有一个有原函数, 则

$$\int uv' dX = uv - \int vu' dX \text{ 或 } \int u dv = uv - \int v du, \text{ 定积分: } \int_a^b uv' dX = [uv] \Big|_a^b - \int_a^b u' v dX.$$

【定积分的近似计算】许多函数的积分是不容易或不可能求出来的, 而近似算法可以求出积分的近似值, 常用方法有:

矩形公式

$$\int_a^b f(X) dX \approx \frac{b-a}{n} \sum_{k=0}^{n-1} f(X_k + \frac{1}{2}h)$$

其中, $h = \frac{b-a}{n}, X_k = a + kh, k = 0, 1, \dots, n$ 。

梯形公式

$$\int_a^b f(X) dX \approx \frac{b-a}{2n} \{f(X) + 2 \cdot [f(X_1) + f(X_2) + \dots + f(X_{n-1})] + f(X_n)\}$$

其中, $X_k = a + \frac{b-a}{n}k, k = 0, 1, \dots, n$ 。

抛物线公式 (辛普森公式)

$$\int_a^b f(X) dX \approx \frac{b-a}{6n} \{f(X_0) + 4[f(X_1) + f(X_3) + \dots + f(X_{2n-1})] + 2[f(X_2) + f(X_4) + \dots + f(X_n)]\}$$

其中 $X_k = a + \frac{b-a}{2n}k, k = 0, 1, \dots, n$ 。

广义积分 (1) 无穷限积分: 设函数 $f(X)$ 在 $[a, +\infty]$ 有定义, 且对任意 $A (A > a)$ 在区间 $[a, A]$ 上可积, 若极限 $\lim_{A \rightarrow +\infty} \int_a^A f(X) dX$

存在,则定义 $\int_a^{+\infty} f(X)dX = \lim_{A \rightarrow +\infty} \int_a^A f(X)dX$ 。此时称广义积分 $\int_a^{+\infty} f(X)dX$ 收敛,若极限不存在,则称广义积分 $\int_a^{+\infty} f(X)dX$ 发散。若 $\int_a^{+\infty} |f(X)|dX$ 收敛,称

$\int_a^{+\infty} f(X)dX$ 为绝对收敛。 $\int_{-\infty}^b f(X)dX, \int_{-\infty}^{+\infty} f(X)dX$ 与此类似。广义积分具有线性性质: $\int_a^{+\infty} [\alpha f(X)dX \pm \beta g(X)]dX = \alpha \int_a^{+\infty} f(X)dX \pm \beta \int_a^{+\infty} g(X)dX$, 分部积分法的公式适用: $\int_a^{+\infty} u dv = uv|_a^{+\infty} - \int_a^{+\infty} v du$ 。

柯西收敛原理: $\int_a^{+\infty} f(X)dX$ 收敛 $\Leftrightarrow \forall \epsilon > 0, \exists A > 0$, 当 $A', A'' > A$ 时有 $|\int_{A'}^{A''} f(X)dX| < \epsilon$ 。还有与级数类似的判敛法,如:若 $|f(X)| \leq \varphi(X)$, 若 $\int_a^{+\infty} \varphi(X)dX$ 收敛,则

$\int_a^{+\infty} f(X)dX$ 绝对收敛;若 $|f| \geq \varphi \geq 0$, $\int_a^{+\infty} \varphi(X)dX$ 发散,则 $\int_a^{+\infty} |f(X)|dX$ 发散,基于这个比较原则,找一些简单的易于判定其敛散的积分作为比较函数,例如柯西判别法就是利用 $\varphi(X) = \frac{c}{X^p} (c > 0)$ 来作比较函数的。

(2) 无界函数的广义积分:若函数 $f(X)$ 在 $[a, b]$ 的端点 b 为无界的(称 b 为 $f(X)$ 的瑕点),但 $\forall \eta > 0, f(X)$ 在 $[a, b - \eta]$ 上可积,且极限 $\lim_{\eta \rightarrow 0} \int_a^{b-\eta} f(X)dX$ 存在,则定义 $\int_a^b f(X)dX = \lim_{\eta \rightarrow 0} \int_a^{b-\eta} f(X)dX$, 此时称广义积分(无界积分)收敛,若极限不存在,则称广义积分 $\int_a^b f(X)dX$ 发散。与此类似 a 为瑕点时 $\int_a^b f(X)dX = \lim_{\eta' \rightarrow 0} \int_{a+\eta'}^b f(X)dX$, a

$c < b$, c 为瑕点时, $\int_a^b f(X) dX = \lim_{\eta \rightarrow +0} \int_a^{c-\eta} f(X) dX + \lim_{\eta' \rightarrow 0} \int_{a+\eta'}^b f(X) dX$ 。同无穷限广义积分一样, 引进绝对收敛的概念, 主要的判敛法也是基于比较原则。

柯西主值(无穷限积分和无界函数的积分)

$$P. V. \int_a^b f(X) dX = \lim_{\eta \rightarrow 0} \left[\int_a^{c-\eta} f(X) dX + \int_{c+\eta}^b f(X) dX \right]$$

$$P. V. \int_{-\infty}^{+\infty} f(X) dX = \lim_{A \rightarrow +\infty} \int_{-A}^A f(X) dX$$

含参变量的积分 在矩形域: $a \leq X \leq b, c \leq y \leq d$ 上考虑二元函数

$f(X, y)$, 当固定 y 时, 它就成了 X 的函数, 若 $\int_a^b f(X, y) dX$ 有意义且在积分过程中仍固定 y , 它就是有意义的, 当 y 在 $[c, d]$ 内变动时, 积分的结果 $I(y) = \int_a^b f(X, y) dX$ 就是 y 的函数, 称这个积分为含参变量的积分, 于是又可讨论连续、可微、可积、广义积分等问题。基本结果是: 在上述矩形域上 (1) $f(X, y)$ 连续, 则 $I(y)$ 是连续函数; (2) 若 $f(X, y), f_y(X, y)$ 连续, 则 $I'(y)$ 存在且 $\frac{d}{dy} \int_a^b f(X, y) dX = \int_a^b \frac{\partial}{\partial y} f(X, y) dX$ 。可变上、下限的含参变量积分, 当 $c \leq y \leq d, a \leq a(y) \leq d, a \leq b(y) \leq b$ 时, $F(y) = \int_{a(y)}^{b(y)} f(X, y) dX$ 在 $[c, d]$ 上; (3) $f(X, y)$ 连续, 则 $F(y)$ 连续; (4) 若还有 $f_y(X, y)$ 连续及 $a'(y), b'(y)$ 存在, 则 $F'(y)$ 存在且 $F'(y) = \int_{a(y)}^{b(y)} f_y(X, y) dX + [f(b(y), y)] b'(y) - f[a(y), y] a'(y)$ 。

含参变量的广义积分 在 $G: a \leq X < +\infty, c \leq y \leq d$ 上考虑积分 $I(y) = \int_a^{+\infty} f(X, y) dX$ 。一致收敛的定义: $\forall \epsilon > 0, \exists A_0(\epsilon) > a$, 当

$A, A' \geq A_0$ 时, 对一切 $y \in [c, d]$ 有 $|\int_A^{A'} f(X, y) dX| < \epsilon$, 就称积分 $\int_a^{+\infty} f(X, y) dX$ 关于 $y \in [c, d]$ 一致收敛, 类似地讨论无界函数积分的一致收敛进而可讨论连续、可微等问题: 在 G 内, $f(X, y)$ 连续, $\int_a^{+\infty} f(X, y) dX$ 一致收敛。(1) $I(y)$ 连续; (2) $\int_c^d dy \int_a^{+\infty} f(X, y) dX = \int_a^{+\infty} dX \int_c^d f(X, y) dy$; (3) 若还有 $f_y(X, y)$ 连续, 则 $I'(y)$ 存在且 $I'(y) = \int_a^{+\infty} f_y(X, y) dX$ 。两个重要的函数:

$$B(p, q) = \int_0^1 X^{p-1} (1-X)^{q-1} dX$$

$$(p > 0, q > 0)$$

$$\Gamma(s) = \int_0^{+\infty} X^{s-1} e^{-X} dX (s > 0)$$

$$\Gamma(s+1) = s\Gamma(s), \Gamma(n+1) = n!$$

斯特林公式

$$\Gamma(X) = \sqrt{2\pi} X^{X-\frac{1}{2}} e^{-X} [1 + r(X)] \quad (|r(X)| \leq \frac{1}{12X} - 1)$$

$$n! \approx \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} \quad (n \rightarrow \infty)$$

【二重积分】考虑一个以函数 $z=f(X, y)$ 为上顶、下底为 Xy 平面区域 (D) , 母线与 Z 轴平行的柱体体积, 用与 X, Y 轴平行的直线划分 D 为许多小矩形 D_i (由 (D) 的边界与这些平行直线所构成的是“曲边矩形”), 求出其面积, 在每个 D_i 内任取一点 (X_i, y_i) , 求出平

顶小柱体的体积 $f(X_i, y_i) \triangle X_i \triangle y_i$, 作出和式 $\sum_{i=1}^n f(X_i, y_i) \triangle X_i \triangle y_i$, 这是体积的近似值, 设上述 D_i 中的直径的最大值为 λ , 若极限

$$\lim_{\lambda \rightarrow 0} \sum_{i=1}^n f(X_i, y_i) \Delta X_i \Delta y_i$$

存在,就称它为该物体的体积。一般,设在域 D 中定义一个函数 $f(X, y)$,用若干与 X, Y 轴平行的直线将 D 分成有限个小区域 $D_1, D_2, \dots, D_n$ 。在每个 D_i 中任取点 (X_i, y_i) 得函数值 $f(X_i, y_i)$,再乘以 D_i 的面积 $\Delta X_i \Delta y_i$,作和式 $\sigma = \sum_{i=1}^n f(X_i, y_i) \Delta X_i \Delta y_i$,设 λ 为 D_i 中的直径中的最大者,若 $\lim_{\lambda \rightarrow 0} \sigma$ 存在,则称它为函数 $f(X, y)$ 展布于 D 上的二重积分,记作 $V = \iint_D f(X, y) dX dy$ 此时称函数 $f(X, y)$ 为可积的,称 D 为积分区域。

性质与计算法则:(1)设 $f(X, y)$ 与 $g(X, y)$ 都可积, α, β 为任意常数,则

$$\iint_D [\alpha f(X, y) + \beta g(X, y)] dX dy =$$

$$\alpha \iint_D f(X, y) dX dy + \beta \iint_D g(X, y) dX dy; (2) \text{ 若 } D = D_1 \cup D_2, D_1 \cap D_2 =$$

$\emptyset$, 则 $\iint_D f(X, y) dX dy = \iint_{D_1} f(X, y) dX dy + \iint_{D_2} f(X, y) dX dy$ 。(3) 通常

把二重积分化为二次积分, 最简单的情况是 D 为矩形域: $a \leq X \leq b, c \leq y \leq d$, 这时有 $\iint_{[a, b; c, d]} f(X, y) dX dy = \int_c^d dy \int_a^b f(X, y) dX = \int_a^b dX$

$\int_c^d f(X, y) dy$ (4) 当 D 为 $a \leq X \leq b, y(X) \leq y \leq Y(X)$ 时, 则有

$$\iint_D f(X, y) dX dy = \int_a^b dX \cdot$$

$\int_a^b \int_{y(X)}^{Y(X)} f(X, y) dy$ 对其他的区域, 可适当分割, 化为若干个 (3) 与

(4) 型区域, 分别计算, 用 (2) 求和即可。(5) 变量替换, 设 $f(X, y)$ 在平面区域 D 中连续, $X = X(u, v), y = y(u, v), (X = X(u, v), y = y(u, v))$ 连续可微, 当 $Xy \in D$ 时, $(u, v) \in D'$ 则

$$\iint_D f(X, y) dX dy = \iint_{D'} f(X(u, v), y(u, v)) \left| \frac{D(X, y)}{D(u, v)} \right| du dv \quad \text{其中 } |J| =$$

$\left| \frac{D(X, y)}{D(u, v)} \right|$ 为雅可比行列式。特例: 化为极坐标计算时: $X = r \cos \theta, y$

$= r \sin \theta$, 此时 $|J| = r$, 于是 $\iint_D f(X, y) dX dy = \iint_D f(r \cos \theta, r \sin \theta) r dr d\theta$ 。

【三重积分】设在三维空间域 V 中定义了函数 $f(X, y, z)$, 用一组曲面将 V 分成几个小区域 $V_1, V_2, \dots, V_n$ 在每个 V_i 中任取一点 $(X_i,$

$y_i, z_i)$, 作和式 $\sigma = \sum_{i=1}^n f(X_i, y_i, z_i) V_i$, 这里 V_i 也表示小区域 V_i 的体积, 设 λ 为 V_i 中的直径中的最大者, 若极限 $\lim_{\lambda \rightarrow 0} \sigma$ 存在, 则称它为函数 $f(X, y, z)$ 展布于 V 上的三重积分, 记作 $\iiint_V f(X, y, z) dX dy dz$ 。

(1) 当积分区域是长方体 $[a, b; c, d; e, f]$ (即 $a \leq X \leq b, c \leq y \leq d, e \leq z \leq f$) 时, 化为

$$\iiint_{[a, b; c, d; e, f]} f(X, y, z) dX dy dz = \iint_{[a, b; c, d]} \left[\int_e^f f(X, y, z) dz \right] dX dy。$$

(2) 当积分区域 V 的曲面方程为

$$z = z_1(X, y), z = z_2(X, y), z_1 \leq z_2$$

V 在 XY 平面上的投影 σ_{XY} 由曲线

$y = y_1(X), y = y_2(X), y_1 \leq y_2, a \leq X \leq b$ 所围成时, 则化为三重积分:

$$\iiint_V f(X, y, z) dv = \int_a^b dX \int_{y_1(X)}^{y_2(X)} dy \int_{z_1(X, y)}^{z_2(X, y)} f(X, y, z) dz。$$

(3) 变量替换: $X = X(u, v, w), y = y(u, v, w), z = z(u, v, w)$ ($(u, v, w) \in V'$) $J = \frac{D(X, y, z)}{D(u, v, w)}$, 则 ($|J| \neq 0$) 有

$$\iiint_V f(X, y, z) dX dy dz = \iiint_{V'} f(X(u, v, w), y(u, v, w), z(u, v, w)) |J| du dv dw。$$

(4)特例:1. 柱面坐标系: $X=r\cos\theta, y=r\sin\theta, z=z$, 此时 $|J|=r$,

$$\iiint_V f(X, y, z) dv = \iiint_V f(r\cos\theta, r\sin\theta, z) r dr d\theta dz.$$

2. 球面坐标: $X=\rho\sin\varphi\cos\theta, y=\rho\sin\varphi\sin\theta, z=\rho\cos\theta$. 此时 $|J|=\rho^2\sin\varphi$, $\iiint_V f dv = \iiint_V f(\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\theta) \rho^2 \sin\varphi d\rho d\varphi d\theta$.

【第一型曲线积分】设给定函数 $f(X, y)$ 与可求长的平面曲线 L , 任意将 L 分为小弧段 $\overline{A_0A_1}, \overline{A_1A_2}, \dots, \overline{A_{n-1}A_n}$ ($A_n=B$), 在每段 $\overline{A_iA_{i+1}}$ 上取点 (X_i, y_i) , 作和式 $\sum_{i=0}^{n-1} f(X_i, y_i) \Delta s_i$, 其中 Δs_i 为 $\overline{A_iA_{i+1}}$ 之长度, 令 $\lambda = \max\{\Delta s_i \text{ 之长}\}$, 当 $\lambda \rightarrow 0$ 时。若极限 $\lim_{\lambda \rightarrow 0} \sum_{i=0}^{n-1} f(X_i, y_i) \Delta s_i$ 存在, 就称它为函数 $f(X, y)$ 沿 L 的(第一型)曲线积分, 记为 $I = \lim_{\lambda \rightarrow 0} \sum_{i=0}^{n-1} f(X_i, y_i) \Delta s_i = \int_L f(X, y) ds$. 计算方法:(1)当曲线 L 表为 $y=y(X)$ ($a \leq X \leq b$) 时,

$$I = \int_a^b f(X, y(X)) \sqrt{1 + [y'(X)]^2} dX;$$

(2)曲线 L 为: $X=\varphi(t), y=\psi(t)$ ($\alpha \leq t \leq \beta$),

$$I = \int_{\alpha}^{\beta} f(\varphi(t), \psi(t)) \cdot \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2} dt.$$

【第二型曲线积分】 f, L 同第一型曲线积分所述, 取 $\triangle s_i$ 在 X 轴上的射影 $\triangle X_i$, 作和式 $\sum_{i=0}^{n-1} f(X_i, y_i) \triangle X_i$, 若 $\mu = \max \overline{A_i A_{i+1}}$ 趋于零

时, 极限 $\lim_{\mu \rightarrow 0} \sum f(X_i, y_i) \triangle X_i$ 存在, 就称它为 $f(X, y)$ 沿曲线 $\widehat{AB}$ 的

(第二型) 曲线积分, 记为 $\int_L f(X, y) dX$ 。同理有 $\int_L f(X, y) dy$ 。计算方法:

设 $L: X = \varphi(t), y = \psi(t), \alpha \leq t \leq \beta$, 则

$$\int_L p(X, y) dX + Q(X, y) dy = \int_{\alpha}^{\beta} [p(\varphi(t), \psi(t)) \varphi'(t) + Q(\varphi(t), \psi(t)) \psi'(t)] dt$$

当以曲线长 s 为参数时, $L: X = X(s), y = y(s) \quad (0 \leq s \leq S)$

$$\text{则 } \int_L f(X, y) dX = \int_0^s f(X(s), y(s)) \cdot$$

$$X'(s) ds = \int_0^s f(X(s), y(s)) \cdot \cos \alpha ds。$$

【曲面积分】设 S 是由分段光滑的边界所包围的双侧曲面, 函数 $f(X, y, z)$ 定义于其上, 用任意的光滑曲线将 S 分为小块 $(s_1), (s_2),$

$\dots, (s_n)$ 。(1) 第一型曲面积分, 记 s_i 为 (s_i) 的面积, 作和式 $\sum_{i=0}^n f(X_i, y_i, z_i) s_i$,

记 $\lambda = \max \{s_i \text{ 的面积} \}$, 若极限 $\lim_{\lambda \rightarrow 0} \sum_{i=0}^n f(X_i, y_i, z_i) s_i$ 存在, 称它为函

数 $f(X, y, z)$ 沿曲面 S 的第一型曲面积分, 记作 $I = \iint_{(S)} f(X, y, z) ds$, 其

中 ds 表示面积元素。设曲面 S 由参数式给出: $X = X(u, v), y = y(u, v), z = z(u, v)$, 当 $(u, v) \in (\triangle)$ 时, $(X, y, z) \in S$, 有

$$\iint_{(S)} f(X, y, z) ds$$

$= \iint_{(\Delta)} f(X(u,v), y(u,v), z(u,v)) \cdot \sqrt{EG-F^2} du dv$ 。其中 $E = X_u^2 + y_u^2 + z_u^2$, $F = X_u X_v^1 + y_u y_v + z_u z_v$, $G = X_v^2 + y_v^2 + z_v^2$ 。(2) 第二型曲面积分

取 (s_i) 在 XY 平面上的射影 (D_i) , 作和式 $\sigma = \sum_{i=1}^n f(X_i, y_i, z_i) D_i$, 当各 (s_i) 中的最大直径 d 趋于零时, 若 $\lim_{d \rightarrow 0} \sigma$ 存在, 就称它为函数 $f(X, y, z)$ 沿 (S) 的第二型曲面积分, 记为

$I = \iint_{(S)} f(X, y, z) dX dy$ 。同理讨论 S 到 YZ 平面和 ZX 平面上的射影构成的第二型曲面积分, 当 S 由参数方程给出: $X = X(u, v)$, $y = y(u, v)$, $z = z(u, v)$, 有

$$\begin{aligned} & \iint_{(S)} P dy dz + Q dz dX + R dX dy \\ &= \pm \iint_{(\Delta)} (PA + QB + RC) du dv, \end{aligned}$$

其中

$$A = \begin{vmatrix} y'_u & z'_u \\ y'_v & z'_v \end{vmatrix}, \quad B = \begin{vmatrix} z'_u & X'_u \\ z'_v & X'_v \end{vmatrix},$$

$$C = \begin{vmatrix} X'_u & y'_u \\ X'_v & y'_v \end{vmatrix},$$

联系:

$$\begin{aligned} & \iint_{(S)} p dy dz + Q dz dX + R dX dy = \\ & \iint_{(S)} [P \cos \lambda + Q \cos \mu + R \cos \gamma] ds \end{aligned}$$

其中 λ, μ, γ 分别是曲面的法线与 X, Y, Z 轴正向所夹锐角。

【格林公式】设 D 是由光滑曲线 L 为边界的平面单连通区域, 函数 $P(X, y)$ 和 $Q(X, y)$ 在 D 及 L 上连续并有对 X 和 y 的连续偏导数, 则

$$\iint_D \left(\frac{\partial Q}{\partial X} - \frac{\partial P}{\partial y} \right) dX dy = \oint_{L^+} P dX + Q dy$$

其中右端积分路径方向是和区域的正向相联系的。

【高斯公式】设连通域 V 的边界曲面 s 是光滑的, 函数 $P(X, y, z)$, $Q(X, y, z)$, $R(X, y, z)$ 在 V 与 S 上连续可微, 则

$$\begin{aligned} \iiint_V \left(\frac{\partial P}{\partial X} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dX dy dz \\ = \iint_S P dy dz + Q dz dX + R dX dy \end{aligned}$$

其中右端的曲面积分是沿闭曲面 S 外侧的。

【斯托克斯公式】设光滑曲面 S 的边界为光滑曲线 L , $P(X, y, z)$, $Q(X, y, z)$, $R(X, y, z)$ 在 S 及 L 上连续可微, 则

$$\begin{aligned} \int_L P dX + Q dy + R dz \\ = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy dz \\ + \iint_S \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial X} \right) dz dX \\ + \iint_S \left(\frac{\partial Q}{\partial X} - \frac{\partial P}{\partial y} \right) dX dy. \end{aligned}$$

【微积分的几何应用】设函数 $y=f(X)$, 定义于 $[a, b]$: (1) 凸性, 若

对任意 $X_1, X_2 \in [a, b]$ 都有 $f\left(\frac{X_1+X_2}{2}\right) \geq \frac{f(X_1)+f(X_2)}{2}$ 则称 $f(X)$ 在 $[a, b]$ 上是向上凸的, 如图所示。当不等式成立时, 称为是严格上凸的, 判据: 在 (a, b) 内 $f''(X) < 0$, 则 $f(X)$ 在 (a, b) 上向上凸。下凸以此类推。(2) 设 $X_0 \in [a, b]$, $f(X) \in C^2(O(X_0, \delta))$, $f''(X_0) = 0$, $f'''(X_0) \neq 0$, 则称 X_0 为 $f(X)$ 的拐点。拐点改变上、下凸的方向。(3) 平面曲线的曲率: $k = \lim_{\Delta s \rightarrow 0} \Delta \varphi / \Delta s$, 其中 $\Delta \varphi$ 为在弧段 Δs 上切线方向的变化角度, 曲率公式 $k = |y'' / (1 + y'^2)^{\frac{3}{2}}|$ 。(4) 曲率半径 $\rho = 1/k$ 。(5) 弧长微分 $ds^2 = dX^2 + dy^2$ 。(6) 弧长的计算公式: $S = \int_a^b \sqrt{1 + f'^2(X)} dX$ 。当曲线由参数方程给出: $X = X(t)$, $y = y(t)$, $\alpha \leq t \leq \beta$, 则 $S = \int_\alpha^\beta [X'^2(t) + y'^2(t)]^{\frac{1}{2}} dt$; 若用极坐标给出: $X = r \cos \theta$, $y = r \sin \theta$, $\alpha \leq \theta \leq \beta$, 则 $S = \int_\alpha^\beta [r'^2(\theta) + r^2(\theta)]^{\frac{1}{2}} d\theta$ 。

当作 $y = f(X)$ 的图形时, 找出 $y = f(X)$ 的极值点、升降的区间、拐点、凸凹性等, 便可顺利作图。

空间曲线的切线与法平面: 设参数方程为 $X=X(t), y=y(t), z=z(t)$,

$X(t), y(t), z(t)$ 可微, 则过

$M_0(X_0, y_0, z_0)$ 的切线方程为 $(X-X_0)/X'(t_0) = (Y-y_0)/y'(t_0) = (Z-z_0)/z'(t_0)$

$$\frac{X-X_0}{X'(t_0)} = \frac{Y-y_0}{y'(t_0)} = \frac{Z-z_0}{z'(t_0)}$$

法平面方程为

$$X'(t_0)(X-X_0) + y'(t_0)(Y-y_0) + z'(t_0)(Z-z_0) = 0$$

曲面的切平面和法线: 设曲面方程为 $F(X, y, z) = 0$, 且 $F(X, y, z)$ 在 M_0 有一阶连续偏导数, 且不同时为零, 则过 $M_0(X_0, y_0, z_0)$ 的切平面方程为

$$F'_X(X_0, y_0, z_0)(X-X_0) + F'_Y(X_0, y_0, z_0)(Y-y_0) + F'_Z(X_0, y_0, z_0)(Z-z_0) = 0$$

法线方程为

$$\frac{X-X_0}{(F_X)_{M_0}} = \frac{Y-y_0}{(F_Y)_{M_0}} = \frac{Z-z_0}{(F_Z)_{M_0}} \quad ((F_X)_{M_0} = F_X(X_0, y_0, z_0))$$

当曲面由参数式给出时: $X=X(u, v), y=y(u, v), z=z(u, v)$, 则切平面:

$$\left[\frac{D(y, z)}{D(u, v)} \right]_{M_0} (X-X_0) + \left[\frac{D(z, X)}{D(u, v)} \right]_{M_0} (Y-y_0) + \left[\frac{D(X, y)}{D(u, v)} \right]_{M_0} (Z-z_0) = 0$$

法线:

$$\frac{X-X_0}{\left[\frac{D(y, z)}{D(u, v)} \right]_{M_0}} = \frac{Y-y_0}{\left[\frac{D(z, X)}{D(u, v)} \right]_{M_0}}$$

$$= \frac{Z - z_0}{\left[\frac{D(X, y)}{D(u, v)} \right]_{M_0}}$$

旋转体的体积与侧面积: $y = f(X)$ 在 $a \leq X \leq b$ 间的一段与 X 轴所围的面积, 绕 X 轴旋转, 则旋转体的

体积: $V = \pi \int_a^b f^2(X) dX$

侧面积: $S = 2\pi \int_a^b y \sqrt{1 + f'^2(X)} dX$

【场论】若在空间或空间的部分区域内任一点都有一个物理量与之对应, 称这种物理量的分布为场。当物理量为数量时称为数量场, 当为向量时称为向量场。因此, 相当于定义了一个点函数或向量函数: $u = u(X, y, z)$ 或 $\vec{A}(P) = P(X, y, z)\vec{i} + Q(X, y, z)\vec{j} + R(X, y, z)\vec{k}$ 。

(数量场的)梯度: 由数量场 $u = u(X, y, z)$ 在 $p(X, y, z)$ 点定义了一个向量函数, 记作 $\text{grad}u =$

$\left[\frac{\partial u}{\partial X}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right]$
 $= \frac{\partial u}{\partial X} \vec{i} + \frac{\partial u}{\partial y} \vec{j} + \frac{\partial u}{\partial z} \vec{k}$ 称为数量场在 P 点的梯度。梯度的方向是使 u (X, y, z) 的方向导数达到最大的方向。

(向量场的)散度:由向量场 $\vec{A} = P\vec{i} + Q\vec{j} + R\vec{k}$, 定义了一个数量函数, 记作: $\text{div} A = \frac{\partial P}{\partial X} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$, 称为这个向量场的散度。若 $\vec{A}$ 是流体流动速度, 则散度是流体在场中各点发散的速度。

(向量场的)旋度:向量场 $\vec{A} = A_x\vec{i} + A_y\vec{j} + A_z\vec{k}$ 定义了一个向量函数, 记作: $\text{rot} \vec{A} = \left[\frac{\partial A_y}{\partial z} - \frac{\partial A_z}{\partial y} \right] \vec{i} + \left[\frac{\partial A_z}{\partial X} - \frac{\partial A_x}{\partial z} \right] \vec{j} + \left[\frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial X} \right] \vec{k}$, 称为这个向量场的旋度, 它反映流体“旋转”的程度。

级数与傅里叶分析

【级数】级数理论是分析学的一个基本分支, 是重要的工具。数项级

$$\text{数: } \sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_n + \cdots \quad (1)$$

记其前 n 项部分和为 $S_n = \sum_{k=1}^n a_k$, 若 $\lim_{n \rightarrow \infty} S_n = s$ 存在, (s 为有限数)。

就称级数(1)为收敛的, 并称 s 为其和, 否则就称它是发散的。由于它是无限多项求和, 故一般的代数运算不能随意地向级数移置, 只有在收敛和(函数项级数的)一致收敛的基础上, 才有可能移置过

来,并在此基础上研究“和函数”的连续性、可微性等。从定义知,级数的收敛与数列收敛是相互转化的;又由于是无限项求和,又与“无限分割而求和”的积分有点类似,二者确也有密切联系。

柯西准则: $\sum_{h=1}^{\infty} a_n$ 收敛 $\Leftrightarrow \forall \epsilon > 0, \exists N > 0$, 当 $n > N$ 时, 对任意的正整数 k 有 $|a_{n+1} + \cdots + a_{n+k}| < \epsilon$ 。

收敛的级数的运算性质: (1) 若 $\sum_{n=1}^{\infty} a_n$ 收敛于 s , 则 $\sum_{n=1}^{\infty} cs$ (c 为常数)。 (2) $\sum_{n=1}^{\infty} a_n$ 和 $\sum_{n=1}^{\infty} b_n$ 都收敛, 则 $\sum_{n=1}^{\infty} (a_n \pm b_n)$ 收敛, 且 $\sum_{n=1}^{\infty} (a_n \pm b_n) = \sum_{n=1}^{\infty} a_n \pm \sum_{n=1}^{\infty} b_n$ 。 (3) 收敛的级数可在其项间任意加括号, 其收敛性不变, 和亦不变。 (4) 收敛级数的通项 a_n 当 $n \rightarrow \infty$ 时趋于 0。

正项级数的收敛准则: 设 $a_n \geq 0 (n=1, 2, \dots)$, 称 $\sum_{n=1}^{\infty} a_n$ 为正项级数。 (1) 若部分和有上界, 则必收敛。 (2) 比较判敛法: 1. 设 $a_n \leq b_n (n=1, 2, \dots)$, 若 $\sum_{n=1}^{\infty} b_n$ 收敛, 则 $\sum_{n=1}^{\infty} a_n$ 收敛; 若 $\sum_{n=1}^{\infty} a_n$ 发散, 则 $\sum_{n=1}^{\infty} b_n$ 发散; 2. 设 $a_{n+1}/a_n \leq b_{n+1}/b_n$: 若 $\sum_{n=1}^{\infty} b_n$ 收敛, 则 $\sum_{n=1}^{\infty} a_n$ 收敛, 或若 $\sum_{n=1}^{\infty} a_n$ 发散, 则 $\sum_{n=1}^{\infty} b_n$ 发散。以下各判别法中的式子均设从某项开始。 (3) 柯西根式判敛法: $\sqrt[n]{a_n} \leq q < 1$, $\sum_{n=1}^{\infty} a_n$ 收敛, $\sqrt[n]{a_n} \geq 1$ 则发散。 (4) (达朗贝尔) $a_{n+1}/a_n \leq r < 1$ 则 $\sum_{i=1}^X a_n$ 收敛, $a_{n+1}/a_n \geq 1$, 则

发散。(5)(拉阿伯判别法): $R_n = n(\frac{a_n}{a_{n+1}} - 1) \geq r > 1$, 则 $\sum_{n=1}^{\infty} a_n$ 收敛;
 $R_n \leq 1$, 则发散。(6)高斯判别法: $a_n/a_{n+1} = \lambda + \frac{\mu}{n} + \frac{\theta_n}{n^2}$, 其中 λ, μ 为
 常数, θ_n 为有界量, 当 $\lambda > 1$ 或 $\lambda = 1$ 且 $\mu > 1$, 级数 $\sum_{n=1}^{\infty} a_n$ 收敛; 当 λ
 < 1 或 $\lambda = 1$ 且 $\mu \leq 1$ 则发散。(7)柯西积分判别法, 设 $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} f$
 (n) , $f(n)$ 是 $f(X)$ 对 $X \geq N$ (自然数) 所定义的某函数, $f(X)F(X) =$
 $\int f(X)dX$, 若 $\lim_{X \rightarrow \infty} F(X) = A < +\infty$, 则 $\sum_{n=1}^{\infty} a_n$ 收敛, 若 $F(X) \rightarrow \infty$ 则
 发散。

【任意项级数】在级数(1)中正负项任意出现, 如果级数(1)的各项
 绝对值所组成的级数 $\sum_{n=1}^{\infty} |a_n|$ 收敛, 称(1)为绝对收敛的, 当(1)收

敛而 $\sum_{n=1}^{\infty} |a_n|$ 不收敛, (1)就称非绝对收敛的或条件收敛的。当然:

$\sum_{i=1}^X |a_n|$ 收敛, (1)必定收敛。绝对收敛的级数可以任意改变项的次
 序, 既不影响收敛性, 也不改变其和。而改变条件收敛的级数的项
 的次序可以使其收敛于事先给定的数(黎曼定理)。条件收敛的级
 数的正项和负项各组成一个发散级数。正负相间的级数叫交错级
 数。 $\sum_{n=1}^{\infty} (-1)^n u_n (u_n > 0)$ 。对此有莱布尼兹定理: $u_n \leq u_{n-1}$, 且 $u_n \rightarrow$
 $0 (n \rightarrow \infty)$, 则交错级数必收敛。考虑级数

$$\sum_{n=1}^{\infty} a_n b_n \quad (C)$$

(1)阿贝尔判别法:若 $\sum_{n=1}^{\infty} b_n$ 收敛,且

$|a_n| \leq k, (n=1, 2, \dots)$, 则 (C) 收敛。(2)狄里克雷判别法: $\sum_{n=1}^{\infty} b_n$

的部分和总有界: $S_n = \sum_{k=1}^n b_k \leq M, (n=1, 2, \dots)$, 又 $a_n \rightarrow 0 (n \rightarrow \infty)$, 则 (C) 收敛。

【函数级数】对 $\forall \epsilon > 0, \exists N$, 对一切 $X \in X$ 都有 $|u_{n+1}(X) + u_{n+2}(X) + \dots| < \epsilon$ 成立, 就称 $\sum_{n=1}^{\infty} u_n(X)$ 在 X 上一致收敛。 X 为某实数集, 是 $U_n(X) (h=1, 2, \dots)$ 的定义域。为方便计, 本段设 $X = [a, b]$ 。

(1)柯西准则, 级数 $\sum_{n=1}^{\infty} u_n(X)$ 一致收敛的充要条件是, $\forall \epsilon > 0, \exists N = N(\epsilon) > 0$, 当 $n > N$ 时,

$|\sum_{k=n+1}^{n+m} U_k| < \epsilon, (m=1, 2, \dots)$ 。(2)外尔斯特拉斯判别法, 设数项级

数 $\sum_{n=1}^{\infty} c_n$ 收敛, $|u_n(X)| \leq c_n, (n=1, 2, \dots) (\forall X \in X)$, 则 $\sum u_n(X)$

一致收敛, 称 $\sum c_n$ 为优级数。(3)阿贝尔判别法, 设

$\sum_{n=1}^{\infty} b_n(X)$ 在 X 上一致收敛, $\forall X \in X \{a_n(X)\}$ 单调且 $|a_n(X)| \leq M$,

$\forall X \in M$, 则 $\sum_{n=1}^{\infty} a_n(X) b_n(X)$ 在 X 上一致收敛。(4)狄里克雷判别

法, $|\sum_{k=1}^{\infty} b_k(X)| \leq M (\forall X \in X), \{a_n(X)\}$ 单调且 $a_n(X)$ 一致趋于

零, 则 $\sum_{n=1}^{\infty} a_n(X) b_n(X)$ 在 X 上一致收敛。

和函数的性质: (1) 设 $u_n(X) \in C[a, b]$, $\sum_{n=1}^{\infty} u_n(X)$ 在 $[a, b]$ 上一致收敛, 则 1. $S(X) \in C[a, b]$; 2. $\int_a^b S(X) dX = \sum_{n=1}^{\infty} \int_a^b u_n(X) dX$; 3. 再加上“ $\sum_{n=1}^{\infty} u'_n(X)$ 一致收敛”, 则有 $(\sum_{n=1}^{\infty} u_n(X))' = \sum_{n=1}^{\infty} u'_n(X)$ 。(2) 狄尼定理: $u_n(X) \in C[a, b]$, $u_n(X) > 0$, $S(X)$ 连续, 则 $\sum_{n=1}^{\infty} u_n(X)$ 一致收敛。

【幂级数】称 $\sum_{n=1}^{\infty} a_n(X-X_0)^n$ 为幂级数。每个幂级数的收敛区间都是关于点 s_0 对称的, 即存在 $R > 0$ 级数在 (X_0-R, X_0+R) 收敛, 称 R 为其收敛半径, 设 $\lim_{n \rightarrow \infty} |\frac{a_{n+1}}{a_n}| = \rho$, 则当 $\rho \neq 0$ 时, 收敛半径 $R = \frac{1}{\rho}$; 当 $\rho = 0$ 时, $R = +\infty$; 当 $\rho = +\infty$ 时, $R = 0$ 幂级数在其收敛区间内任一闭区间上必一致收敛且绝对收敛, 对 $|X-X_0| > R$ 的 X , 级数发散。在 (X_0-R, X_0+R) 内和函数连续, 可逐项积分与逐项求导。

泰勒公式与泰勒级数: (1) 设 $f(X)$ 在 X_0 点有 n 阶导数, 则 $f(X) =$

$\sum_{k=1}^{\infty} \frac{f^{(k)}(X_0)}{k!} (X-X_0)^k + o((X-X_0)^n)$ 。(2) 设 $f(X) \in C^n(a, b)$, $f^{(n+1)}(X)$ 存在, $X_0, X \in (a, b)$, 则 $f(X) =$

$\sum_{k=1}^n \frac{f^{(k)}(X_0)}{k!} (X-X_0)^k + f^{(n+1)}(\zeta)/(n+1)! \cdot (X-X_0)^{n+1}$, 其中 ζ 介于 X_0 与 X 之间。(1) 与 (2) 中的 $f(X)$ 的表示式称为泰勒公式。

(3) 若 $f(X)$ 在 (X_0-R, X_0+R) 无穷次可微, 则当余项 $R_n(X) \rightarrow 0$ (n

$\rightarrow \infty$) 时有展开式 $f(X) = \sum_{n=0}^{\infty} \frac{f^{(n)}(X_0)}{n!} (X - X_0)^n$ ($-R < X - X_0 < R$) 称此展开式为泰勒级数。(4) 当 $X_0 = 0$ 时, (3) 中的级数称为马克劳林级数。余项的形式:

拉格朗日形式:

$$R_n(X) = f^{(n+1)}(\theta X) / (n+1)! \cdot X^{n+1}$$

柯西形式:

$$R_n(X) = f^{(n+1)}(\theta X) / n! \cdot (1-\theta)^n X^{n+1} \quad (0 < \theta < 1)$$

【傅里叶分析】又称调和分析。是在三角级数基础上发展起来的、18 世纪后形成的重要的分支, 主要研究傅里叶变换及其性质。

正交系与傅里叶级数是向量垂直概念的推广。设在平面上两个向量 $\vec{a} = (a_1, a_2), \vec{b} = (b_1, b_2)$ 。 $\vec{a} \perp \vec{b}$, 即内积为零: $\langle a, b \rangle = a_1 b_1 + a_2 b_2 = 0$, 平方可积函数 $f(X)$ 与 $g(X)$ 的内积定义为 $\langle f, g \rangle = \int_a^b f(X) g(X) dX$ 。用 $\langle f, g \rangle = 0$ 来定义正交, 一个函数系 $\{\varphi_k(X)\} \subset L^2[a, b]$ 。 $\varphi_k(X) \neq 0 (k=0, 1, 2, \dots)$, 当 $\int_a^b \varphi_i(X) \varphi_j(X) dX = 0 (i \neq j)$ ($i, j=0, 1, 2, \dots$) 时, 称 $\{\varphi_k(X)\}$ 为正交系, 若还有 $\int_a^b \varphi_k^2(X) \equiv 1 (k=1, 2, \dots)$, 就称它为就范正交系。例如, 三角函数系

$$1, \cos X, \sin X, \cos 2X, \sin 2X, \dots, \cos nX, \sin nX, \dots \quad (1)$$

是正交系, 而 $1, \frac{1}{\sqrt{\pi}} \cos X,$

$$\frac{1}{\sqrt{\pi}} \sin X, \dots, \frac{1}{\sqrt{\pi}} \cos nX,$$

$\frac{1}{\sqrt{\pi}} \sin nX, \dots$ 是就范正交系, 在更抽象的空间, 也用内积来定义正

交性,当然,定义内积的方式有所不同。利用正交系,可将一般周期函数展成三角级数

$$f(X) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} [a_k \cos kX + b_k \sin kX] \quad (2)$$

称(2)为 $f(X)$ 的傅里叶级数。其系数分别为

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(X) dX, \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(X) \cos kX, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(X) \sin kX dX \end{aligned} \quad (3)$$

为 $f(X)$ 的傅里叶系数。傅里叶级数的复数形式为 $f(X) = \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n e^{inX}$

$C_n e^{inX}$ 部分和是 $\sum_{-n}^n C_k e^{ikX}$, 系数则为 $C_n =$

$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(X) e^{-inX} dX$ 。为讨论(2)的收敛性, 把它的部分和表为含参变

量的广义积分——狄里克雷积分:

$$\begin{aligned} S_n[f(X)] &= \frac{1}{2}a_0 + \sum_{k=1}^n [a_k \cos kX + b_k \sin kX] \\ &= \frac{1}{\pi} \int_0^{\pi} [f(X+u) + f(X-u)] \\ &\quad \times \frac{\sin(n+\frac{1}{2})u}{2\sin \frac{u}{2}} du \end{aligned}$$

$S_n[f(X)]$ 的收敛就转化为

$$S_n[f(X)] - S = \frac{1}{\pi} \int_0^{\pi} [f(X+u) - S] \frac{\sin(n+\frac{1}{2})u}{2\sin \frac{u}{2}} du$$

$$+f(X-u)-2s]\frac{\sin \frac{2n+1}{2}u}{2\sin \frac{1}{2}u}du$$

是否收敛于零的问题。

重要性质与判别法:(1)若 $f(X)$ 可积和绝对可积,其傅里叶系数 $a_n \rightarrow 0, b_n \rightarrow 0$ (当 $n \rightarrow \infty$),而其傅里叶级数收敛于 $f(X)$ 本身。

(2)狄尼判别法:对固定的 X ,若有适当的 s ,使 $\varphi(u) = f(X+u) + f(X-u) - 2s$ 满足:对某正数 h ,在 $[0, h]$ 上, $\varphi(u)/u$ 可积和绝对可积,则 $f(X)$ 的傅里叶级数在 X 点收敛于 s 。(3)李普希茨判别法 1. 若 $f(X)$ 在点 X 连续,并且对充分小的 u ,在点 X 处成立李普希茨条件:

$$|f(X \pm u) - f(X)| < Lu^\alpha (0 < u \leq h) \\ (L > 0, 0 < \alpha \leq 1)$$

则 $f(X)$ 在点 X 的傅里叶级数收敛于

$f(X)$ 。2. 若对充分小的 u 有 $|f(X \pm u) - f(X \pm 0)| < Lu^\alpha$,则 $f(X)$ 的傅里叶级数在点 X 上收敛于 $\frac{1}{2}[f(X+0) + f(X-0)]$ 。(4)狄里克雷判别法:若在 $[X-h, X+h]$ 上, $f(X)$ 是分段单调函数,则 $f(X)$ 的傅里叶级数在 X 点收敛于 $\frac{1}{2}[f(X+0) + f(X-0)]$ 。(5)1. 设 $f(X) \in C[-\pi, \pi], f(-\pi) = f(\pi)$,除有限个点外 $f'(X)$ 存在;2.

$\int_{-\pi}^{\pi} |f'(X)|^2 dX$ 收敛,则 $f(X)$ 的傅里叶级数一致收敛于 $f(X)$ 。(6)

帕舍伐尔等式:设 $f(X) \in L^2(-\pi, \pi)$,则

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(X)|^2 dX = 2 \sum_{k=-\infty}^{\infty} |C_k|^2 = \frac{1}{2} a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2)。(7)$$

(里斯-费希尔定理)给定双向数列 $\{c_k\} (k=0, \pm 1, \pm 2 \dots)$,若

$\sum_{k=-\infty}^{\infty} c_k^2 < \infty$, 则存在唯一的函数 $f(X) \in L^2(-\pi, \pi)$, 它的傅里叶系数等于 $\{c_k\}$ 。(8) 设 c 和 X 是 $[-\pi, \pi]$ 上任意两点, $[-\pi, \pi]$ 上的分段连续函数 $f(X)$ 的傅里叶级数可逐项积分得

$$\int_c^X f(X) dX = \frac{1}{2} a_0 (X - c) + \sum_{k=1}^{\infty} \int_c^X (a_k \cos kX + b_k \sin kX) dX$$

(9) 在 (5) 1. 之下, $f'(X)$ 可从其傅里叶级数逐项求导而得到。

【傅里叶积分】周期函数在一定条件下可展成傅里叶级数(2), 用类似于作积分和的办法转化为

$$f(X) \sim \frac{1}{\pi} \cdot \int_0^{\pi} du \int_{-\infty}^{\infty} f(t) \cos u(X-t) dt \quad (4)$$

称(4)为函数 $f(X)$ 的傅里叶积分, 当满足下述条件之一:

(1) (迪尼) 对某数 $h > 0$

$$\int_b^h |f(X+t) + f(X-t) - 2f(X)| t dt < \infty$$

(2) (狄里克雷-约当) $f(X)$ 在 $(X-h, X+h)$ 上分段连续时, $f(X)$ 的傅里叶积分在点 X 收敛于 $f(X)$ 。

(4) 的复数形式为

$$f(X) \sim \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iuX} du \int_{-\infty}^{\infty} f(t) e^{-iut} dt \quad (5)$$

称(5)的内层积分

$$\hat{f}(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-iut} dt$$

为傅里叶变换, 这是一种积分变换, 它具有下述性质: (1) $\hat{f}(u)$ 是 $u \in (-\infty, \infty)$ 的连续函数; (2) $\lim_{u \rightarrow \infty} \hat{f}(u) = 0$; (3) 线性性质; (4) 平移性质: $\hat{f}(X-s) = e^{-isu} \hat{f}(u)$; (5) 当 $f(X) \in L'(-\infty, \infty)$ 时, 逆变换 f

$$(X) = \int_{-\infty}^{\infty} f(u) e^{iXu} du a. e. \text{ 成立。}$$

函 数 论

实变函数论

【实变函数论】研究自变量取实数数值的函数的一般理论。在 18 世纪末、19 世纪初,微积分已基本成熟,在许多方面得到发展和应用。此时人们对它的概念和基础进行探究,发现有许多问题解释不清,或不便应用。其中主要涉及函数的连续性、可微性、极限和积分等问题。经过许多数学家的努力,这些问题都得到解决和发展。其中尤其是积分理论由 (R) (黎曼) 积分发展为 (L) (勒贝格) 积分。 (L) 积分和 (L) 测度的概念已在概率论及其它许多学科中找到了应用。在连续性方面,区间 $[a, b]$ 上的任何连续函数必为一列多项式一致收敛的极限;但连续函数列的极限却不一定是连续函数,从而引进了贝尔函数类的概念。由于直观,人们很可能认为,连续曲线就像我们用手所画的那样,并且只有有限个没有切线的点。其实,皮亚诺作出了这样的曲线:它充满了一个正方形;而外尔斯特拉斯作出了一个函数,它在 $[a, b]$ 上连续,但处处不可导。对于导数的研究,指出了有界变差函数类几乎处处有导数,从而具体找出了一类可求长的曲线。许多数学分支都与实变函数论有关。

【集合的势】设有集合 A, B , 若存在一个由 A 到 B 上的一一映射, 则称集合 A 与集合 B 对等, 记作 $A \sim B$ 。规定空集与其自身对等。对等具有下述性质: $A \sim A$ (自反性); $A \sim B, B \sim C$, 则 $A \sim C$ (传递性); $A \sim B$, 则 $B \sim A$ (对称性)。若两个集合对等, 则称它们的“势”相等。“势”是有限集的个数的推广, 集合 A 的势记为 $\bar{A}$ 。

伯恩斯坦定理指出: 设 A, B 为两个集合, $A_1 \subset A, B_1 \subset B$, 若 $A \sim B_1$, 且 $B_1 \sim A$, 则 $A \sim B$ 。

与自然数集 N 对等的集称为可列集(或可数集), 可列集的势记为 $\bar{N} = \aleph_0$ 读作阿列夫零。全体有理数组成的集合 Q 是可列集。可列集的子集至多可列。若 A 是可列集, B 为有限集, 则 $A \cup B$ 是可列集, 若 $A_i (i=1, 2, \dots)$ 是可列集, 则 $\bigcup_{i=1}^n A_i$ 和 $\bigcup_{i=1}^{\infty} A_i$ 也是可列集。若 $A_i (i=1, 2, \dots, n)$ 是可列集, 则 $A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_i \in A_i, i=1, 2, \dots, n\}$ 是可列集。整系数多项式全体是可列集。

区间 $[0, 1]$ 是不可列的。若集合 A 与 $[0, 1]$ 对等, 则称集合 A 的势为连续统, 记为 $\bar{A} = \aleph$, 读作阿列夫。直线上任一区间(包括 $(a, +\infty)$, 或 $(-\infty, b)$ 或 $(-\infty, +\infty)$)的势是 $\aleph$ 。实数列全体的势为 $\aleph$ 。二进制数全体的势为 $\aleph$ 。设集合 A 的势为 μ , 以 2^μ 表示 A 的一切子集所组成的集的势, 则 $\mu < 2^\mu$ 。

【康托尔三分集】将闭区间 $[0, 1]$ 三等分, 去掉中间的一个开区间

$$0 \quad \frac{1}{9} \quad \frac{2}{9} \quad \frac{1}{3} \quad \frac{2}{3} \quad \frac{7}{9} \quad \frac{8}{9} \quad 1$$
 $I_1^{(1)} = (\frac{1}{3}, \frac{2}{3})$, 把剩下的两个闭区间 $[0, \frac{1}{3}]$ 和 $[\frac{2}{3}, 1]$ 分别再三等分, 并各去掉中间的开区间: $I_1^{(2)} = (\frac{1}{9}, \frac{2}{9})$, $I_2^{(2)} = (\frac{7}{9}, \frac{8}{9})$; 余下四个闭区间 $[0, \frac{1}{9}]$, $[\frac{2}{9}, \frac{1}{3}]$, $[\frac{6}{9}, \frac{7}{9}]$, $[\frac{8}{9}, 1]$ 再分别三等分, 并各去掉中间的开区间: $I_1^{(3)} = (\frac{1}{27}, \frac{2}{27})$, $I_2^{(3)} = (\frac{7}{27}, \frac{8}{27})$, $I_3^{(3)} = (\frac{19}{27}, \frac{20}{27})$, $I_4^{(3)} = (\frac{25}{27}, \frac{26}{27})$; ……第 n 次去掉的开区间(称为第 n 级开区间)是: $I_1^{(n)} = (\frac{1}{3^n},$

$\frac{2}{3^n}$), $I_2^{(n)} = (\frac{7}{3^n}, \frac{8}{3^n})$, $\dots\dots I_{2^{n-1}}^{(n)} = (\frac{3^n-2}{3^n}, \frac{3^n-1}{3^n})$, 令 $G_0 = \bigcup_{n,k} I_k^{(n)}$ (求和号中 $k=1, 2, \dots, 2^{n-1}; n=1, 2, \dots$), 称 $P_0 = [0, 1] - G_0$ 为康托尔集。 P_0 是完全集、疏朗集; $\overline{P_0} = \mathbb{S}$ 。

【 R 中开集的构造】(1) 设 $G \subset R$ 为非空开集。 $I = (\alpha, \beta)$ 是一个开区间 $(-\infty \leq \alpha, \beta \leq +\infty)$, 称 I 是 G 的构成区间, 若: $I \subset G$; 若 α, β 有限, 则 α, β 不属于 G 。设 F 为直线上的闭集, 称 F^c (为开集) 的构成区间为 F 的余区间。(2) 直线上任一非空开集 G 可以表示为至多可列个互不相交的构成区间之并: $G = \bigcup_k (\alpha_k, \beta_k)$, 其中每一个 (α_k, β_k) 都是 G 的构成区间, $\bigcup_k$ 可以是 $\bigcup_{k=1}^n$ 或 $\bigcup_{k=1}^\infty$ 。

R^n 中非空开集 G 可表为可列个互不相交的半闭正方形之并。

【稠密与疏朗】 设 $A \subset R^n, B \subset R^n$, 若 $\forall X \in B, \forall N(X, r)$, 其内均含有 A 的点, 则称 A 在 B 中稠(密), 当 $B = R^n$ 时, 称 A 为稠密集, 设 $A \subset R^n$ 在任何邻域 $N(a, r) (a \in A)$ 中, 总有邻域 $N(y, \delta) \subset N(a, r)$, 且 $N(y, \delta) \cap A = \emptyset$, 则称 A 为疏朗集。

G_δ 型集 R^n 中一系列开集之交 $\bigcap_{k=1}^\infty G_k (G_k \text{ 为开集})$

F_δ 型集 R^n 中一系列闭集之并 $\bigcup_{k=1}^\infty F_k (F_k \text{ 为闭集})$ 。

【 R^n 的基本拓扑性质】(1) 完备性: 点列 $\{X_k\} \subset R^n$ 称为基本列(柯西列), 如果 $\forall \varepsilon > 0, \exists N \in \mathbb{N}$, 当 $l, k > N$ 时有 $d(X_k, X_l) < \varepsilon$ 。完备性是说: R^n 内任一基本列必为收敛列。(2) 列紧性: 任一有界点列必定含有收敛子列; 设 F 为 R^n 的有界闭集, 则 F 的任一开覆盖必有有

限于覆盖。(3)分离性:设 F_1, F_2 为 R^n 中的非空不交闭集,则存在开集 G_1, G_2 , 使: $G_1 \supset F_1, G_2 \supset F_2$, 且 $G_1 \cap G_2 = \emptyset$ 。

【勒贝格测度】区间长度、平面图形面积等的推广。即把区间的长度扩充到一般的点集上去,对 n 维空间来说,把只适用于“立方体”的体积概念扩充到一般点集上去。设 $E \subset R^n$, 称

$\inf \left\{ \sum_{k=1}^{\infty} |I_k| : \bigcup_{k=1}^{\infty} I_k \supset E, I_k \text{ 为 } R^n \text{ 中的开区间} \right\}$ 为集 E 的勒贝格外测度,记为 $m^* E$, 并规定 $m^* \emptyset = 0$ 。它有如下性质:(1) $m^* E \geq 0$; (2) 单调性, $E_1 \subset E_2$, 则 $m^* E_1 \leq m^* E_2$; (3) 半可加性:

$m^* \left(\bigcup_{k=1}^{\infty} E_k \right) \leq \sum_{k=1}^{\infty} m^* E_k$; (4) 距离可加性若 $d(A, B) > 0$, 则 $m^* (A \cup B) = m^* A + m^* B$ 。

$E \subset R^n$, 若对任意 $A \subset R^n$, 都有 $m^* A = m^* (A \cap E) + m^* (A \cap E^c)$, 则称 E 为勒贝格可测集, 记测度为 mE , 此时 $mE = m^* E$ 。上式称为卡拉特渥多利条件。测度的性质为:(1) E 可测时, E^c 必可测; (2) $\{E_k\}$ 为可测列, 则 $\bigcap_{k=1}^{\infty} E_k, \bigcup_{k=1}^{\infty} E_k$ 均可测; (3) 若 E_1 与 E_2 均可测且 $E_1 \subset E_2$, 则 $mE_1 \leq mE_2$; (4) 若 $\{E_k\}$ 为单增(减)

函数列 $E = \bigcup_{k=1}^{\infty} E_k (E = \bigcap_{k=1}^{\infty} E_k)$, 则 $mE = \lim_{k \rightarrow \infty} mE_k$; (5) 若 E 可测, 则存在 G_δ 型集 A 与 F_σ 型集 B , 使 $A \supset E \supset B$ 且 $mA = mB = mE$; (6) E 可测 $\Leftrightarrow \forall \varepsilon > 0, \exists$ 开集 G 与闭集 F , 使 $G \supset E \supset F$, 且 $m(G - F) < \varepsilon$ 。

设 $P(X)$ 是一个与自变量 $X (X \in E)$ 有关的命题。如果 E 中使这个命题不成立的 X 所组成的点集是 E 的一个测度为零的子集, 则说这个命题在 E 上几乎处处(a. e.)成立。设 f 为定义于 E 上的广义实值函数(即函数值可取为 $\pm \infty$)。若对任意实数 $c, E(f > c)$ 是可测集, 则称 $f(X)$ 为 E 上的可测函数。 f 在 E 上可测 $\Leftrightarrow \forall c \in R, E(f > c)$ 是可测集, 可测 $\Leftrightarrow \forall c \in R, E(f < c)$ 可测 $\Leftrightarrow \forall c \in R, E(f \leq c)$ 可测。连续

函数在其定义域内必可测。

可测函数的性质:(1)对四则运算封闭:设 f, g 为 E 上的可测函数 $a \in R$, 则 $af, f \pm g, f \cdot g$ 可测, 当 f/g a. e. 有意义时, f/g 可测。(2)对极限运算封闭: 设每个函数 f_k 都可测, 则 $\sup f_k(X), \inf f_k(X), \overline{\lim} f_k(X), \lim f_k(X) = \inf_k \sup_{m \geq k} f_m(X), \liminf f_k = \sup_k \inf_{m \geq k} f_m$ 及 $f(X) = \lim_{k \rightarrow \infty} f_k(X)$ 均为可测函数。(3) f 在 E 上可测 $\Leftrightarrow f$ 可表为一列简单函数的 a. e. 收敛的极限; (4) E 上的非负可测函数 $f(X)$ 必为非负单增函数列 $\{\psi_k(X)\}$ 的极限: $0 \leq \psi_1(X) \leq \psi_2(X) \leq \dots, \lim_{k \rightarrow \infty} \psi_k(X) = f(X)$ 。(5) $f: R^n \rightarrow R^* = R \cup \{\pm \infty\}$ 。 f 在 R^n 可测 $\Leftrightarrow \forall$ 开集 $G \subset R, f^{-1}(G)$ 在 R^n 可测。(6) f 在 R 连续, g 在 R^n 可测 $\Leftrightarrow fog$ 在 R^n 可测。

设 $f, g, f_k (k=1, 2, \dots)$ 定义于可测集 E 上及 f 可测, (1) a. e. 收敛与一致收敛的关系: 设 $mE < +\infty, f_k$ 和 f 均 a. e. 有限, 且 $f_k \rightarrow f$ (a. e.), 则 $\forall \delta > 0, \exists E_\delta \subset E$, 使 f_k 在 E_δ 上一致收敛于 f , 且 $m(E - E_\delta) < \delta$ 。(2) 依测度收敛: 设每个 f_k 是 a. e. 有限可测的, 若 $\forall \epsilon > 0, \lim_{k \rightarrow \infty} mE(|f_k - f| \geq \epsilon) = 0$ 。就称 f_k 依测度收敛于 f , 记为 $f_k \xrightarrow{m} f (k \rightarrow \infty)$, 它有以下性质: 极限函数是唯一的, 即若 $f_k \xrightarrow{m} f (k \rightarrow \infty)$, 又 $f_k \xrightarrow{m} g (k \rightarrow \infty)$, 则 $f \sim g$ (即 $E_1 = \{X | X \in E, f \neq g\}, mE_1 = 0$); 线性性质: $\forall \alpha, \beta \in R, f_k \xrightarrow{m} f, g_k \xrightarrow{m} g (k \rightarrow \infty)$, 则 $\alpha f_k + \beta g_k \xrightarrow{m} \alpha f + \beta g$ 且 $f_k^2 \xrightarrow{m} f^2$ 及 $f_k g_k \xrightarrow{m} f g (k \rightarrow \infty)$; 当 $f_k \xrightarrow{m} f$ 时必有 $|f_k| \xrightarrow{m} |f| (k \rightarrow \infty)$; (3) 关系: $f_k \rightarrow f$ (a. e.) $(k \rightarrow \infty)$, 必有 $f_k \xrightarrow{m} f (k \rightarrow \infty)$; (Riesz) 定理: 若 $f_n \xrightarrow{m} f (f \rightarrow \infty)$, 则存在子列 $\{f_{n_k}\}$ 使 $f_{n_k} \xrightarrow{a. e.} f(X) (k \rightarrow \infty)$ 。称 $f(X)$ 具有

性质 A, 若 $\forall \varepsilon > 0, \exists$ 闭集 $F \subset E$, 使 $m(E - F) < \varepsilon$ 且将 f 限制于 F 上时连续。鲁金定理, 设 f 定义在可测集 E 上, 且 a. e. 有限, 则 f 可测 $\Leftrightarrow f$ 具性质 A; $E \subset R, f$ 在 E 上是 a. e. 有限可测的, 则 $\forall \delta > 0$, 存在连续函数 $g(X)$, 使 $mE(f \neq g) < \delta$ 。

【勒贝格积分】(R) 积分基本上是为连续函数建立的, 有许多重要的函数不存在 (R) 积分, 例如 Dirichelet 函数便是, 在极限符号与积分符号可交换顺序方面, (R) 积分所要求的条件苛刻; 还有, 即使 $F'(X)$ 存在且有界时,

$F'(X)$ 也不一定可积, 从而牛顿-莱布尼茨公式 $F(X) - F(a) =$

$\int_a^X F'(X) dX$ 便不能应用。为克服上述局限性, 引进了 (L) 积分: 设

$y = f(X)$ 为定义于 $[a, b]$ 上的有界函数, $A \leq f(X) \leq B$, 分割其值域: $A = y_0 < y_1 < \dots < y_n = B$, 令 $e_k = \{X \mid y_{k-1} \leq y \leq y_k\}$, 论 $\lambda =$

$\max_{1 \leq k \leq n} (y_k - y_{k-1})$, 在 e_k 中任取点 g_i , 作和数 $\sigma = \sum_{k=1}^n f(g_k) m e_k$, 若

$\lim_{\lambda \rightarrow 0} \sigma$ 存在, 就称这个极限为 $f(X)$ 在 E 上的 (L) 积分, 记为 $\int_E f(X)$

dX 或 $\int_E f(X) dm$ 。(L) 积分的性质: (1) 若 $mE < +\infty$, 则 $f(X)$ 在 E

上有界可积 $\Leftrightarrow f(X)$ 在 E 有界可测; (2) 线性性质: f, g 均有界可

积, 则 $\forall \alpha, \beta \in R, \alpha f(X) + \beta g(X)$ 有界可积, 且 $\int_E (\alpha f + \beta g) dm = \alpha$

$\int_E f dm +$

$\beta \int_E g dm$; (3) 可加性: 1. (对集合的有限可加性): f 在 E 上可积, E

$= \bigcup_{k=1}^n E_k, E_k$ 可测, $E_i \cap E_j \neq \emptyset (i \neq j)$, 则 f 在每个 E_i 上可积, 且

$\int_E f dX = \int_{E_1} f dX + \int_{E_2} f dX + \cdots + \int_{E_n} f dX$ 。2. (对集合的 σ 可加性): f

在 E 上可积, $E = \bigcup_{k=1}^{\infty} E_k$, E_i 可测 $E_i \cap E_j = \emptyset (i \neq j)$, 则

$\int_E f dX = \sum_{k=1}^{\infty} \int_{E_k} f dX$ 。(4) 积分是绝对连续的, f 在 E 上可积, 则 $\forall \varepsilon$

$> 0, \exists \delta > 0$, 当 $m(e) < \delta (e \in E)$ 时, $|\int_e f dX| < \varepsilon$; (5) $f \geq 0, \int_E f dX$

$= 0 \Leftrightarrow f = 0, a. e.$; (6) $f = g, a. e.$, 当 f 可积时, g 必可积, 且积分值

相等; (7) $f \leq g, a. e.$, 则 $\int_E f dX \leq \int_E g dX$; (8) 若 g 可测, $0 \leq g \leq f$, 当

f 可积时, g 必可积且 $\int_E g dX \leq \int_E f dX$; (9) f 可积 $\Leftrightarrow |f|$ 可积; (10)

$mE = 0$, 则定义于 E 上的任一函数 f 必有 $\int_E f dX = 0$; (11) 可积函

数必 $a. e.$ 有限; (12) (L) 积分与 (R) 积分的关系: 在 $[a, b]$ 上 (R) 可

积的函数必 (L) 可积, 且积分值相等; $f(X)$ 广义 (R) 可积: f 是 (L)

可积 $\Leftrightarrow |f|$ 广义 (R) 可积, 且积分值相等, f 在 $[a, b]$ 上有界, 则:

(R) 可积 $\Leftrightarrow f$ 在 $[a, b]$ 上 $a. e.$ 连续。

无界可测集上的 (L) 积分。(1) 设 $A \subset R^n$ 为无界可测集, f 为 A 上的非负可测函数, 取

$$\triangle_k = [-k, k) \times \cdots \times [-k, k), A_k = A \cap \triangle_k,$$

则 $A_k \nearrow A$ (以上升方式趋于 A), 称 $\lim_k \int_{A_k} f dX$ 为 f 在 A 上的 (L) 积

分, 记为 $\int_A^* f dX$, 当 $\lim_k \int_A^* k f dm < +\infty$ 时, 就说 f 在 A 上 (L) 可积,

否则, 只说 f 在 A 上的 (L) 积分存在。(2) 设 f 为定义在集 A 上的

可测函数, 如果 f^+, f^- 均在集 A 上 (L) 可积, 则说 f 在集 A 上 (L)

可积, 其积分为

$\int_A^* f dm = \int_A^* f^+ dm - \int_A^* f^- dm$, 当 f^+, f^- 中只有一个 (L) 可积时, 仍记 $\int_A^* f dm = \int_A^* f^+ dm - \int_A^* f^- dm$, 此时, 只说 f 在 A 上的 (L) 积分存在。

【积分号下的极限运算】设可测集 $E \subset R^n, f, f_k (k=1, 2, \dots), g$ 均为 E 上的可测函数: (1) (单调收敛性) $0 \leq f_n \nearrow f$, 则 $\lim_{n \rightarrow \infty} \int_E f dX = \int_E f dX$; (2) f_n 非负可测, 则 $\int_E (\sum_1^\infty f_n) dX = \sum_{n=1}^\infty \int_E f_n dX$; (3) (法都 Fatou 引理): g, h 在 E 上可积, $\{f_n\}$ 是可测函数列. 1. 若对一切 $n \geq 1, g \leq f_n$, 则 $\int_E \lim_{n \rightarrow \infty} f_n dX \leq \lim_{n \rightarrow \infty} \int_E f_n dX$; 2. 若对一切 $n \geq 1, f_n \leq h$, 则 $\lim_{n \rightarrow \infty} \int_E f_n dX \leq \int_E \lim_{n \rightarrow \infty} f_n dX$; (4) (勒贝格控制收敛定理): g 可积, $|f_n| \leq g$,

$\lim_{n \rightarrow \infty} f_n = f$ a. e., 则 f 可积, 且

$\lim_{n \rightarrow \infty} \int_E f_n dX = \int_E f dX$; (5) 设 g 可积, $|f_n| \leq g$ a. e., $f_n \xrightarrow{m} f$, 则 f 在 E 上可积, 且 $\lim_{n \rightarrow \infty} \int_E f_n dX = \int_E f dX$; (6) 含参变量积分 $I(t) = \int_E f(t, X) dm$ 的连续性与可微性: (i) 设 $\forall f \in [a, b], f(t, X)$ 作为 X 的函数在 E 上可测, (ii) $|f(t, X)| \leq F(X), \forall t \in [a, b], F(X)$ (L) 可积, (iii) $\forall \{f_n\} \subset [a, b]$, 当 $t_n \rightarrow t_0 \in [a, b]$ 时有 $f(t_n, X) \rightarrow f(t_0, X)$

(a. e. 于 E), 则 (a) $I(t)$ 在 t_0 点连续; (b) 若再设 $\frac{\partial f(t, X)}{\partial t}$ a. e. 存在

且 $|\frac{\partial f}{\partial t}| \leq g(X), g(X)$ 是 (L) 可积的, 则 $I'(t_0) = \int_E \frac{\alpha}{\alpha t} f(t_0, X) dm$ 。

重积分化为累次积分的 Fubini 定理: 设 $f(X, y)$ 是 I 上的 (L) 可积函数, $I = I_p \times I_q$, I_p, I_q 分别为 R^p, R^q 的区间, 则 (1) a. e. $X \in I^n$, $f(X, y)$ 作为 y 的函数在 I_q 上可积; (2) $F_f(X) = \int_{I_q} f(X, y) dm_y$ 在 I_p 上可积; (3) $\int_{I_p} F_f(X) dm_X = \int_{I^n} f(X, y) dm$ 。

【维他利覆盖】 设 E 为直线上的非空集合, $M = \{d\}$ 是由非退化的闭区间组成的集合 (即 $\forall d \in M, |d| > 0$), 如果 $\forall X \in E, \exists \{d_n\}_1^\infty \subset M$, 使 $|d_n| \rightarrow 0 (n \rightarrow \infty)$, 且 $X \in d_n (n = 1, 2, \dots)$, 则称 M 在维他利意义下覆盖 E 。维他利覆盖定理。设 E 为有界集, $E \subset R, M$ 是 E 的维他利覆盖, 则存在至多可数个互不相交的闭区间列 $\{d_n\} \subset M$, 使 $m(E - \bigcup_n d_n) = 0$ 。

设 f 是定义在 $[a, b]$ 上的实值函数, 称 $\lambda \in [-\infty, +\infty]$ 为 f 在点 $X_0 \in [a, b]$ 的一个导出数, 如果存在数列 $\{h_n\} (h_n \neq 0, X_0 + h_n \in [a, b], h_n \rightarrow 0 (n \rightarrow \infty))$ 使 $\lim_n [f(X_0 + h_n) - f(X_0)]/h_n = \lambda$, 记为 $Df(X_0) = \lambda$ 。当 f 在 X_0 的一切导出数都相等时, 称 f 为在 X_0 可微 (可导), 记为 $f'(X_0)$ 。定理: $[a, b]$ 上的单调函数 $f(X)$ a. e. 存在有限导数。

【有界变差函数】 设 f 在 $[a, b]$ 上定义, 任取一组分点 $a = X_0 < X_1$

$< \dots < X_n = b$, 记 $V_i(X_0, X_1, \dots, X_n) = \sum_{i=1}^n |f(X_i) - f(X_{i-1})|$, 若对一切可能的分点组 $\{V_i(X_0, X_1, \dots, X_n)\}$ 有界, 即 $\sup_{X_0, X_1, \dots, X_n} V_i(X_0, X_1, \dots, X_n) < \infty$, 就称 f 是 $[a, b]$ 上的有界变差函数, 记作 $\bigvee_a^b(f) =$

$\sup_{X_0, X_1, \dots, X_n} v_i(X_0, \dots, X_n)$, 并称 $\bigvee_a^b(f)(X)$ 为 f 在 $[a, b]$ 上的全变差。 $[a,$

$b]$ 上全体有界变差函数记作 $V_{[a,b]}$ 。性质(1)单调函数必为有界变差的;(2)有界变差函数必有界;(3)有界变差函数具有线性性质;(4)一个函数,若其全变差为零,则它必为常数;(5)在 $[a,b]$ 上的有界变差函数在其闭子区间上也是有界变差的;(6)
$$\overset{b}{V}_a = \overset{c}{V}_a + \overset{b}{V}_c \quad (a < c < b);$$
(7)函数列 $g_n \rightarrow g, \{\overset{b}{V}_a(g)\}$ 有界,且 $g \in V[a,b]$ 则 $\overset{b}{V}_a(g) \leq \sup \{\overset{b}{V}_a(g_n)\}$;(8) $f(X)$ 是有界变差的,当且仅当 $f = \pi(X) - \nu(X)$,其中 $\pi(X)$ 与 $\nu(X)$ 均为单增函数;(9) $f \in V_{[a,b]}$,则 $f'(X)$ a. e. 存在。有限并可积,又 f 的不连续点集至多可列。

称定义于 $[a,b]$ 上的函数 $f(X)$ 为绝对连续的,若: $\forall \varepsilon > 0, \exists \delta > 0$,对 $[a,b]$ 中任一互不相交且 $\sum_i (b_i - a_i) < \delta$ 的至多可列个开区间列 $\{(a_i, b_i)\}$ 有 $\sum |f(b_i) - f(a_i)| < \varepsilon$ 。性质:(1)绝对连续函数必为有界变差的且为一致连续的,还是 a. e. 可微的函数;(2)对四则运算封闭(对 f/g ,要 $g(X) \neq 0$)。 $F(X) = \int_{[a,X]} f dm = \int_a^X f(X) dX$ 为 f 的勒贝格(L)不定积分。性质与计算法则:(1) $F(X)$ 绝对连续,从而 a. e. 可微;(2) $F'(X) = f(X)$ a. e. ;(3)一个函数 $F(X)$ 绝对连续,当且仅当存在(L)可积函数 $f(X)$ 使 $F(X) = \int_a^X f(X) dX$;(4)(分部积分公式) $F(X)$ 绝对连续, g 可积且 $G(X) = \int_{[a,X]} g dm$ ($X \in [a,b]$),则有 $\int_{[a,b]} Fg dm + \int_{[a,b]} F'G dm = F(b)G(b) - F(a)G(a)$;(5)(换元公式)设 g 在 $[a,b]$ a. e. 可微, f 可积且 $g([a,b]) \subset [c,d]$, $F(X) = \int_{[a,X]} f dm$,则 $fg(t)g'(t)$ 可积,且 $\int_{[g(a),g(b)]} f(X) dm = \int_{[\alpha,\beta]} f(gt) \cdot g'(t) dm$ 其中 $\alpha, \beta \in [a,b]$;(6)(第一中值定理) f

在 $[a, b]$ 连续, g 是 $[a, b]$ 上的非负可积函数, 则 $\exists \xi \in [a, b]$ 使 $\int_a^b f \cdot g dm = f(\xi) \cdot \int_{[a, b]} g dm$ 。(7)(第二中值定理) 若 f 在 $[a, b]$ 上 (L) 可积, g 是 $[a, b]$ 上的单调函数, 则 $\exists \xi \in [a, b]$, 使 $\int_{[a, b]} f \cdot g dm = g(a) \int_{[a, \xi]} f dm + g(b) \int_{[\xi, b]} f dm$ 。

【斯蒂尔吉斯积分】(1) 黎曼-斯蒂尔吉斯(R-S)积分, 设 f, g 定义于 $[a, b]$, 作分划 $\Delta: a = X_0 < X_1 < \cdots < X_n = b$, 任取 $\zeta_i \in [X_{i-1}, X_i]$ 作和式 $S(\Delta) =$

$\sum_{i=1}^n f(\zeta_i) [g(X_i) - g(X_{i-1})]$, 记 $\lambda = \max_{1 \leq i \leq n} (X_i - X_{i-1})$, 若极限 $\lim_{\lambda \rightarrow 0} S(\Delta) = I$ 存在, 则称 I 为 f 关于 g (R-S) 积分, 记为 $I = \int_a^b f(X) dg(X)$ 。性质 1. 线性性质 (a) $\int_a^b (\alpha f_1 + \beta f_2) dg = \alpha \int_a^b f_1 dg + \beta \int_a^b f_2 dg$; $\int_a^b f d(\alpha g_1 + \beta g_2) = \alpha \int_a^b f dg_1 + \beta \int_a^b f dg_2$; 2. $\int_a^b f dg = \int_a^c f dg + \int_c^b f dg$ ($\forall X \in [a, b]$); 3. f 关于 g [R-S] 可积, 当且仅当 g 关于 f (R-S) 可积, 且 $\int_a^b f dg + \int_a^b g df = f(b)g(b) - f(a)g(a)$ 。R-S 积分存在条件:

在 g 为单增函数的条件下, 下面各条之一便保证 $\int_a^b f dg$ 存在:

1. f 是 (R) 可积的, g 满足李普希茨条件; 2. f 连续, g' (R) 可积; 3. f 连续, g 绝对连续; 4. $f \in V[a, b]$, g 绝对连续。将 (L) 积分通过 R-S 积分来计算: 1. 设 f 是 E 上 a. e. 有限的 (L) 可测函数, $\alpha \in R$, $mE < +\infty$ 。称 $w_f(\alpha) = m(\{X \in E | f(X) > \alpha\})$ 为在 E 上的分布函数; 2. 若 f 在 E 上有界可测, $\forall X \in E$ 有 $a < f(X) \leq b$, 则 $\int_E f dm =$

$-\int_a^b \alpha dw(\alpha)$ 。可推广到当 f 在 E 上无界时的情形。(2) 勒贝格-斯蒂尔吉斯(L-S)积分。先定义(L-S)测度: 设 $f(X)$ 是定义在 R 上单增的实值函数, 对每个左开右闭的有限区间 (a, b) , 定义 $\lambda_f((a, b)) = f(b) - f(a)$ 。对 R 的任一非空子集 E , 称

$$\inf \left\{ \sum_{k=1}^{\infty} \lambda_f((a_k, b_k)) \mid \bigcup_{k=1}^{\infty} (a_k, b_k) \supset E \right\}$$

为相应于 f 的(L-S)外测度, 记作 $\wedge_f^*(E)$, 规定 $\wedge_f^*(\emptyset) = 0$ 。和勒贝格外测度一样, 具有非负性、单调性、半可加性和距离可加性。由此, 当 $E \subset R$, 若 $\forall A \subset R$ 总有 $\wedge_f^*(A) = \wedge_f^*(A \cap E) + \wedge_f^*(A \cap E^c)$, 则称 E 为相应于 f 的(L-S)可测集, 此时其(L-S)外测度称为相应于 f 的(L-S)测度, 记为 $\wedge_f$ 。由此, 与前面类似, 定义(L-S)可测函数, (L-S)积分等等。(R-S)积分和(L-S)积分的关系: 设 f 是 $[a, b]$ 上的递增的右连续函数, g 是 $[a, b]$ 上有界的 Borel 可测函数, 如果 g 关于 f 的(R-S)积分存在, 则 g 关于 $\wedge_f$ 的(L-S)积分也存在, 且 $\int_{[ab]} g d\wedge_f = \int_a^b g df$ 。(设 E 为 Borel 集, 称 $f: E \rightarrow R$ 为 Borel 可测函数, 如果对于任意实数 c , $E(f > c)$ 为 Borel 集)。

复变函数论

【复变函数论】定义域及值域均为复数集的函数称为复变函数。通常记作 $w = f(z)$ 。若设 $w = u + iv$, $z = X + iy$, 也可表示 $f(z) = u(X, y) + iv(X, y)$, 这表明一个单复变量函数关系等价于两个二元实变量函数关系, 关于复变函数的理论称为复变函数论。其中有关解

析函数的理论内容十分丰富,并得到广泛的应用,是复变函数论的核心部分。

【解析函数】定义于复平面区域 D 上的复变函数 $f(z) = u(X, y) + iv(X, y)$, 若对 D 内任一点 z 都有

$$\lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} = A \quad (h \text{ 是复变量}, A \text{ 是有限复数, 称之为 } f(z) \text{ 在 } z \text{ 点的复导数, 记作 } A = f'(z)),$$

则称 $f(z)$ 是 D 内的解析函数。解析函数一定是定义于开集上的, 有三种常用的等价定义揭示了解析函数三类不同的特征: (1) $u(X, y)$ 与 $v(X, y)$ 在 D 内处处可微, 且

$$\text{满足 } \frac{\partial u}{\partial X} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial X}. \quad (\text{称为柯西-黎曼条件, 简记为 } C-R \text{ 条件});$$

(2) 在 D 内任一点 z_0 的任意一个完全属于 D 的领域 $\Delta: |z - z_0| <$

R 内, $f(z)$ 可以展成泰勒 (Taylor) 级数, 即 $f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$; (3)

对 D 内任一条可求长约当闭曲线 L , 只要其内部全属于 D , 必有

$$\int_L f(z) dz = 0. \quad \text{命题(1)揭示了解析函数与微分方程的密切联系; 命题(2)揭示了解析函数的一种构造, 由此可导出解析函数是无限次}$$

可导函数, 且 $a_k = \frac{f^{(k)}(z_0)}{k!} (k=0, 1, \dots)$ 。由定义推证命题(3)称为

柯西积分定理, 它是解析函数理论的核心定理, 其重要应用之一是

导出了以下公式: $f(z)$ 在区域 D 解析, 在 $\bar{D} = D \cup L$ 连续, L 是 D 的

边界, 它由有限多条简单光滑闭曲线构成, 则有 $f^{(k)}(z) = \frac{k!}{2\pi i} \int_L$

$$\frac{f(\zeta)}{(\zeta - z)^{k+1}} d\zeta \quad (z \in D, k=0, 1, \dots)$$

这公式表明解析函数 $f(z) = f^{(0)}(z)$ 及其各阶导数在 D 内的值均可以用 $f(z)$ 在边界 L 上的值来表示, 而 $k=0$ 的情形称为柯西积分公式, 在复变函数论中占有重要地

位。

当 $f(z)$ 在去心的圆盘 $\triangle_0: 0 < |z - z_0| < R$ 上解析, 而在 z_0 点不解析(或无定义)时, $f(z)$ 在 $\triangle_0$ 可以表作 $f(z) = \sum_{k=-\infty}^{-1} a_k (z - z_0)^k + \sum_{k=0}^{\infty} a_k (z - z_0)^k$ 右边的级数称为罗朗级数, 它的非负幂部分称为解析部分, 它在 $|z - z_0| < R$ 上解析, 它的负幂部分称为主要部分, 其和函数在 $|z - z_0| > 0$ 上解析, z_0 称为 $f(z)$ 的孤立奇点, 罗朗级数的系数为 $a_k = \frac{1}{2\pi i} \int_L \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} d\zeta$ ($k = 0, \pm 1, \dots$), 其中 L 是 $|\zeta - z_0| = \rho$ ($0 < \rho < R$), 逆时针方向。 $f(z)$ 在孤立奇点 z_0 的不同性态是由主要部分的不同情形加以区别的: (1) 当主要部分的系数全为零时, 称 z_0 是 $f(z)$ 的可去奇点, 这时补充定义 $f(z_0) = a_0$, $f(z)$ 即是 $|z - z_0| < R$ 内的解析函数; (2) 当主要部分为 $\sum_{k=-n}^{-1} a_k (z - z_0)^k$ ($a_{-n} \neq 0$) 时, 称 z_0 是 $f(z)$ 的 n 阶极点, 这时有

$\lim_{z \rightarrow z_0} f(z) = \infty$, 它的另一个特征是 $\frac{1}{f(z)}$ 在 z_0 点有一 n 阶零点(即 $f^{(k)}(z_0) = 0, k = 0, 1, \dots, n-1$, 而 $f^{(n)}(z_0) \neq 0$); (3) 当主要部分有无限多项的系数非零时, 称 z_0 是 $f(z)$ 的本性奇点, 在本性奇点的任意小邻域内函数可以取到任意复数值, 最多只有一个例外值, 这就是著名的毕卡定理, 在罗朗级数的系数中 $a_{-1} = \frac{1}{2\pi i} \int_L f(\zeta) d\zeta$ 有重要的应用, 称它为 $f(z)$ 关于 z_0 点的留数, 记作 $Res[f(z), z_0]$ 。

初等解析函数是重要的解析函数类, 它包括多项式、有理函数、根式函数、指数函数、对数函数、三角函数、反三角函数以及由它们复合而成的解析函数。这些函数与同名实函数取相同记号, 并且包括同名实函数为其一部分。它们不但保持了同名实函数的几

乎全部性质,而且进一步揭示了初等函数的许多新的特性:例如欧拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$ 揭示了三角函数与指数函数的关系,并由此推知指数函数 $w = e^z$ 是以 $2k\pi i$ 为周期的周期函数,而它的反函数——对数函数则是多值函数,即对任意非零复数 z ,有 $\ln z = \ln|z| + i(\arg z + 2k\pi)$ (k 取任意整数)。所有单值初等函数都是其定义域上的解析函数,其求导法则也与同名实函数一样。多值初等函数总可以在特定的区域上被分解成单值解析分支,而加以应用或讨论其性质。

还有一些称之为特殊函数的解析函数类,它们不能用初等函数的代数运算或复合来生成,而通常是某些函数方程或微分方程的解,并以积分或某种展开式的形式出现。如 Γ 函数,开始作为函数方程 $X\Gamma(X) = \Gamma(X+1)$, $\Gamma(1) = 1$ 的解,它有性质: $\Gamma(n+1) = n!$,故它可看作是阶乘 $n!$ 的推广,一般以欧拉积分的形式表示 $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$,但这时仅当 $x > 0$ 时才有意义,后来发现当它的定义域开拓到复平面时才揭示了 Γ 函数的解析结构: $\Gamma(z) = z^{-1} e^{-\gamma z} \sum_{n=1}^{\infty} (1 + \frac{z}{n})^{-1} e^{\frac{z}{n}}$, 其中 $\gamma = \lim_{n \rightarrow \infty} (1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n)$ 称为欧拉常数,它的近似值为 0.57722, $\Gamma(z)$ 在全复平面除在 $0, -1, -2, \dots$ 有一阶极点外处处解析。再如椭圆函数,是复平面上双周期亚纯函数(除极点外没有其它奇点的函数称为亚纯函数),它的一个重要的子类称为维爾斯特拉斯函数,其基本周期为 ω_1 及 ω_2 ($\text{Im}\omega_2/\omega_1 \neq 0$),记 $\omega = n_1\omega_1 + n_2\omega_2$ (n_1, n_2 为整数),这时 $\zeta(z) = \frac{1}{z^2} + \sum \left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right]$, 其中 $\sum$ 表示 n_1, n_2 取遍全体整数之和,它在全平面除以全体 ω 为二阶极点外处处解析,已证明它满足微分方程 $(W')^2 = 4W^3 - g_2W - g_3$, 其中 g_2, g_3 是常数,因而 $W = \zeta(z)$ 的反函数可以表示

成一个椭圆积分: $z =$

$$\int_{w_0}^w \frac{dw}{\sqrt{4w^3 - g_2w - g_3}} + c, \text{ 此处 } c \text{ 是常数, } w_0 = \zeta(c).$$

【共形映照】一个区域 D 上的解析函数 $w = f(z)$ 当 $f'(z) \neq 0$ 时, 它实现的映照有以下性质: (1) 在 D 内任一点 z_0 的一个完全属于 D 的邻域内映照是双方单值的, 称之为局部单叶性; (2) 过 z_0 点的任意一条光滑弧 $C_z: z = z(t)$ 的象曲线为过 $w_0 = f(z_0)$ 点的光滑弧 $C_w: W = w(t) = f(z(t))$, 且 C_w 上任一点 $w = w(t)$ 到 w_0 点的距离 $|w - w_0|$ 与相应的原象 $z = z_1(t_1)$ 到 z_0 点的距离 $|z - z_0|$ 之比。当 $t \rightarrow t_0$ 时极限为 $|f'(z_0)|$, 它与 C_z 弧的取法无关, 称 $|f'(z_0)|$ 为 $f(z)$ 在 z_0 点的伸缩率, 称此性质为伸缩率的不变性; (3) 过 z_0 点的任意两条光滑弧 C_{z_1}, C_{z_2} 的夹角与象曲线 C_{w_1}, C_{w_2} 的夹角同向相等, 此处两弧之夹角是指在交点处两切线之夹角, 这一性质称为第一类保角性。按以上性质可推知在 z_0 点的足够小的邻域内一个三角形与其象 (一般来说是一个曲边三角形) 之间具有近似的同向相似性, 因而称具有上述三条性质的解析映照为共形映照。具有以上三条性质的函数也一定是导数处处不为零的解析函数。共形映照的基本定理是黎曼映照定理: D 是边界点集多于一个点的单连通区域, 则一定存在 $w = f(z)$ 将 D 双方单值共形映照为单位圆 $\Delta: |w| < 1$, 若再要求对取定的 $z_0 \in D$ 满足 $f(z_0) = 0, f'(z_0) > 0$, 则上述映照还是唯一确定的。这一定理在复变函数理论的发展上起着重要的作用。若单连通区域 D 的边界 C 是简单闭曲线, 则将 D 双方单值共形映照为单位圆的函数 $w = f(z)$ 一定可以连续开拓到闭区域 $\bar{D} = D \cup C$ 上, 且将 $\bar{D}$ 双方单值连续映照到闭单位圆 $|w| \leq 1$, 这一结论称为边界对应定理。由初等解析函数实现的映照称为初等共形映

照,其中特别重要的一类是分式线性映照 $w=w(z)=\frac{az+b}{cz+d}$ (a, b, c, d 为复数且 $ad-bc \neq 0$), 若补充定义 $w(-\frac{d}{c})=\infty$,

$w(\infty)=\frac{a}{c}$ 且合理地规定 $w=\frac{1}{z}$ 在 $z=0$ 及 $z=\infty$ 的邻域内也实现共形映照,则分式线性映照是扩充复平面(即加上 ∞ 点的复平面)上的共形映照,分式线性映照有以下性质:(1)若定义两个分式线性映照的“乘法”为 $w_1(z) \circ w_2(z) = w_1(w_2(z))$, 则全体分式线性映照关于此乘法构成变换群;(2)每个不为常数的分式线性映照在扩充复平面上最多有两个不动点,也可能只有一个二重的不动点;

(3)交比 $(z_1, z_2, z_3, z_4) = \frac{z_1 - z_3}{z_1 - z_4} : \frac{z_2 - z_3}{z_2 - z_4}$ 是分式线性映照 $w(z)$ 的不变量,即 $(w(z_1), w(z_2), w(z_3), w(z_4)) = (z_1, z_2, z_3, z_4)$; (4)扩充复平面上的“圆”(包括过 ∞ 点的“圆”——直线)经分式线性映照的象仍为扩充复平面上的“圆”; (5)关于“圆”对称的点经分式线性映照后的象点仍是关于象“圆”对称的点。除分式线性映照外其它初等函数也有重要的映照特性,如 $W=e^z=e^x \cdot e^{iy}$ 将带形区域 $\alpha < \text{Im} z < \beta$ ($\beta - \alpha < 2\pi$) 映为角型区域 $\alpha < \arg w < \beta$ ($\beta - \alpha < 2\pi$); $w=\ln z$ 为其逆映照,它的单值解析分支 $w_k=\ln|z|+i(\arg z+2k\pi)$ 映角形区域 $\alpha < \arg z < \beta$ ($\beta - \alpha < 2\pi$) 为带形区域 $\alpha+2k\pi < \text{Im} w < \beta+2k\pi$; 函数 $w=\frac{1}{2}(z+\frac{1}{z})$ 称为儒可夫斯基函数,映 $|z|<1$ 及 $|z|>1$ 为同样的象区域:沿 $[-1, 1]$ 割开的复平面,即 $C \setminus [-1, 1]$; 而映上,下半平面也为同样的象区域:沿 $(-\infty, -1]$ 及 $[1, +\infty)$ 割开的平面,即 $C \setminus ((-\infty, -1] \cup [1, +\infty))$; 三角函数 $w=\cos z = \frac{1}{2}(e^{iz}+e^{-iz})$ 可以

看成是 $w_1=\gamma z, w_2=e^{w_1}, w=\frac{1}{2}(w_2+\frac{1}{w_2})$ 的复合,它映带形域 $0 <$

$\operatorname{Re} z < \pi$ 为 $C \setminus [-\infty, -1] \cup [1, \infty]$ 。且所有的带形域 $k\pi \leq \operatorname{Re} z < (k+1)\pi$ 的象均相同。非初等解析函数的共形映照也有重要的结果, 如 $f(z) = c_0 \int_{z_0}^z \prod_{\mu=1}^m (a_\mu - z)^{\alpha_\mu - 1} dz + c_1$ (其中 c_0, c_1 为常数), 将单

位圆盘 $|w| \leq 1$ 映为多角形区域 $\bar{D}$; 即以 $b_\mu = f(a_\mu)$ ($\mu = 1, 2, \dots, m$) 为多角形顶点, 对应顶角为 $\alpha_\mu \pi$, 除 a_μ 外映照是共形的。

多连通区域同样可以被共形映照为多连通(连通数不能变)的典型区域, 这典型区域可以有多种形式, 例如可以是带有 $n-2$ 条同心圆弧割口的同心圆环。特别二连通区域, 当其两个连通的边界点集均多于一个点时, 它一定可以被共形映照为同心圆环; $1 < |w| < R$, 但 R 不能任意指定, 有如下结论: 圆环 $r_1 < |z| < r_2$ 与圆环 $R_1 < |w| < R_2$ 之间存在共形映照的充分必要条件是 $r_1/r_2 = R_1/R_2$ 。

【解析开拓】将一个定义于区域 D 上的解析函数 $w = f(z)$ 开拓为一个包含 D 的更大的区域上的解析函数, 称后者为前者的解析开拓。即 $f(z)$ 与 $g(z)$ 是分别定义于区域 D 和 G 上的解析函数若 $G \supset D$, 且在 D 上 $f(z) = g(z)$, 则称 $g(z)$ 是 $f(z)$ 由 D 到 G 上的解析开拓。例如 $g(z) = \frac{1}{1-z}$ 就是 $f(z) = \sum_{k=1}^{\infty} z^k$ 从 $D: |z| < 1$ 到 $G: C \setminus \{1\}$ 上的解析开拓。实现解析开拓的方法主要有两种: (1) 将 $f_0(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$ 在其收敛圆 $D_0: |z - z_0| < R_0$ ($R_0 > 0$) 内一点 z_1 ($\neq z_0$) 处展成幂级数 $\sum_{k=0}^{\infty} a_k^{(1)} (z - z_1)^k$, 设其收敛圆为 $D_1: |z - z_1| < R_1$ 且在 D_1 上收敛于解析函数 $f_1(z)$, 这时若 $D_1 \setminus D_0 \neq \emptyset$, 则

$g(z) = \begin{cases} f_0(z) & z \in D_0 \\ f_1(z) & z \in D_1 \end{cases}$ 就是 $f_0(z)$ 从 D_0 到 $D_0 \cup D_1$ 上的解析开拓。这种方法称为幂级数开拓法,用这一方法可以一直不断地开拓下去,直到后一个幂级数的收敛圆永远被前一个的收敛圆所包含,就认为开拓到“头”了,这时最后得到的函数称为完全解析函数,而它的边界被称为自然边界。(2)若 $f(z)$ 的解析性区域 D 在上半平面,且 D 的边界上有一段 γ 在实轴上, $f(z)$ 在 $D \cup \gamma$ 上连续,且在 γ 上

$f(z)$ 取实值,那么

$$F(z) = \begin{cases} f(z) & z \in D \cup \gamma \\ \overline{f(\overline{z})} & z \in D^* \end{cases}$$

就是 $f(z)$ 由 D 到 $D \cup \gamma \cup D^*$ 上的解析开拓,此处 D^* 是区域 D 关于实轴的对称区域,因而 $f(z)$ 的解析性已通过 γ 开拓到对称区域 D^* 上,这一方法称为对称原理,它的一般情形是要求区域 D 在一个“圆”的一侧,且边界的一段 γ 在“圆”上,于是以上结论也是适用的,不论用哪一种方法作解析开拓,多次开拓之后可能出现这样的情况:在相重叠的一些区域上各自经解析开拓得到的函数都不相同,这时解析开拓得到了多值解析函数,但若认为有不同解析开拓的相重叠的区域不是在同一平面上,称它们为不同的叶,按这一规定将开拓一直进行下去,直到得到完全解析函数,这时它的定义域是一个连通的多叶的抽象“曲面”,我们称这一曲面为黎曼曲面。因此,一个完全解析函数在其黎曼曲面上是一个单值解析函数。

【多复变函数】有多个自变量的复值函数。例如设 C^n 是 n 维复空间,表作 $C^n = \{z = (z_1, z_2, \dots, z_n) \mid z_k \in C, k=1, 2, \dots, n\}$, D 是 C^n 中的一个开集,则 D 到 C 中的一个单值对应关系就是一个 n 维多复变函数,记作 $w = f(z) = f(z_1, z_2, \dots, z_n)$,若对任意 k ,取定 $z_1, \dots,$

$z_{k-1}, z_{k+1}, \dots, z_n$ 的值时 $f(z)$ 是 z_k 变量的解析函数, 则称 $f(z)$ 是 D 上的多复变解析函数。它的等价的定义有两种形式: (1) $\forall z_0 \in D$, $z_0 = (z_{0_1}, z_{0_2}, \dots, z_{0_n})$, 都有 z_0 的一个邻域 $U(z_0) \subset D$, 和一个幂级数:

$$\sum_{p \in N^n} a_p (z - z_0)^p = \sum_{p_1, p_2, \dots, p_n \geq 0} a_{p_1, p_2, \dots, p_n} (z_1 - z_{0_1})^{p_1} (z_2 - z_{0_2})^{p_2} \cdots (z_n - z_{0_n})^{p_n},$$

式中 $P = (p_1, p_2, \dots, p_n)$, p_k 为自然数, 此级数在 $U(z_0)$ 上收敛于 f

(z) ; (2) $f(z)$ 在 D 上处处有 $\frac{\partial f}{\partial z_k} = \frac{1}{2} \left(\frac{\partial}{\partial X_k} + i \frac{\partial}{\partial Y_k} \right) f(z) = 0$ ($k=1, 2, \dots, n, z_k = X_k + iY_k$)。满足条件(2)的 $f(z)$ 有时也称为全纯函数。多

复变函数理论的研究几乎和单复变函数同时开始, 但它却远不像单复变函数理论那样系统和完整, 即使多复变解析函数, 除了一些最基本的性质与单复解析函数相同外, 在许多重要的性质上都与单复变解析函数有本质的区别, 例如 $f(z) = z_1 \cdot z_2$ 是 C^2 上的解析函数, 其零点集为 $\{(0, z_2) \cup \{(z_1, 0)\}$, 显然, 对于多复变解析函数来说, 零点不一定孤立, 同样, 极点也不一定是孤立的, 如 $f(z) =$

$\frac{1}{z_1 + z_2 - 1}$ 的极点集是一二维超曲面 $z_1 + z_2 - 1 = 0$ 。于是有关孤立奇

点的概念和性质都难以推广到多复变解析函数中来。再如多复变解析函数按其定义很难简单推出它一定是连续函数。海尔托克斯 (Hartogs) 用了很长时间才解决了这一难题, 从而证明了全纯函数一定是解析函数。在单复变函数中极其重要的柯西积分公式到多复变函数类中也发生了根本的变化; 定义于 $D = \{z \mid |z_k - a_k| \leq r_k, k=1, 2, \dots, n\}$ 上的解析函数 $f(z), \forall z \in D$, 都有积分表示式: f

$$(z) = (2\pi i)^{-n} \int_{|\zeta_1 - a_1| = r_1} \cdots \int_{|\zeta_n - a_n| = r_n} \frac{f(\zeta_1, \dots, \zeta_n)}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)} d\zeta_1 \cdots d\zeta_n, \text{ 这一公}$$

式我们称它为多圆柱 D 上的柯西积分公式, 但我们看到它的积分限并非区域 D 的边界, 而只是边界的一小部分, 称 $(|\zeta_1 - a_1| = r_1)$

$\times (|\zeta_2 - a_2| = r_2) \times \cdots \times (|\zeta_n - a_n| = r_n)$ 为区域 D 的一个框架, 可见多圆柱 D 上多复变解析函数在区域内的值仅仅由其在框架上的值就可以完全确定了。实际上, 多复变解析函数有多种形式的积分表示, 它们甚至与区域的形状有关。对于单复变解析函数来说有这样的性质: D 是 C 中的区域, 若有 $a \in C \setminus D$, 则一定有 D 上的解析函数, 不能解析开拓到 a 点, 例如 $f(z) = \frac{1}{z-a}$ 就是这样的函数。但在多复变解析函数中这一性质不再成立。却有这样的结果: 设 $\triangle = \{z = (z_1, \cdots, z_n) \mid |z_k| < 1, k = 1, 2, \cdots, n\}$, U 是 $\triangle$ 的邻域, 使 $\triangle \cap U$ 是连通的, 则 $\triangle$ 上的解析函数 $f(z)$ 一定可以解析开拓到 $\triangle \cup U$ 上, 即一定存在一个在 $\triangle \cup U$ 上解析的函数 $F(z)$, 使在 $\triangle$ 上 $F(z) = f(z)$ 。单复变解析函数的黎曼映照定理对多复变解析函数也不再成立, 已证明双圆柱区域 $D = \{(z_1, z_2) \mid |z_1| < 1, |z_2| < 1\}$ 与超球 $G = \{(z_1, z_2) \mid |z_1|^2 + |z_2|^2 < 1\}$ 这样两个简单的单连通区域之间不存在双方单值共形映照 (共形映照的定义可仿照解析函数的定义给出)。

微分方程论

【微分方程】常微分方程与偏微分方程的总称。大致与微积分同时产生。力学、天文学、几何学等领域的许多问题都可以归结为微分方程。20 世纪以来,一些社会科学领域的问题也导致微分方程,如人口发展的模型等。在微分方程发展的早期,数学家们集中注意于求通解,由于这对大多数方程不能实现,于是逐步放弃了这一奢望,而转向“定解问题”(初值问题、边值问题、混合问题等),随之产生存在性、唯一性和连续依赖性等,即使对定解问题,表示为初等函数的解一般也不可能,于是数值计算方法、定性理论等便应运而生。

常微分方程

【常微分方程】含有一个自变量、未知函数及其导数的关系式

$$F\left(X, y, \frac{dy}{dX}, \frac{d^2y}{dX^2}, \dots, \frac{d^ny}{dX^n}\right) = 0 \quad (1)$$

叫做常微分方程,其中出现的最高阶导数的阶数,叫做常微分方程的阶。如 $\frac{dy}{dX} + ay = f(X)$, $y \cdot \frac{dy}{dX} + \frac{d^3y}{dX^3} + \frac{d^5y}{dX^5} = 0$ 分别是 1 阶和 5 阶的常微分方程。设 $y = \varphi(X)$ 定义于区间 J 上,有直到 n 阶的导数,

将它代入(1),将(1)变成一个关于 X 的恒等式,即

$$F \left[X, \varphi(X), \frac{d\varphi(X)}{dX}, \frac{d^2\varphi(X)}{dX^2}, \dots, \frac{d^n\varphi(X)}{dX^n} \right] \equiv 0,$$

$X \in J$

就称 $y = \varphi(X)$ 为(1)的一个定义于 J 上的一个解。称 J 为它的定义区间。 n 阶方程的含有 n 个独立常数的解 $y = \varphi(X, C_1, C_2, \dots, C_n)$ 叫做它的通解,不含任意常数的解叫做特解。这里说 $C_1, C_2, \dots, C_n$ 是独立的,意即雅可比行列式

$$\frac{D[\varphi, \varphi^1, \dots, \varphi^{(n-1)}]}{D[C_1, C_2, \dots, C_n]} \neq 0,$$

其中 $\varphi^1, \dots, \varphi^{(n-1)}$ 分别是 φ 关于自变量 X 的相应阶导数。在 (X, y) 平面上, $y = \varphi(X)$ 所描绘的曲线叫做微分方程的积分曲线。方程的隐式解叫做积分。

【常微分方程的初等解法】①可分离变量的方程:其形式为 $f_1(X)f_2(y)dX = g_1(X)g_2(y)dy$, 当 $f_2(y)g_1(X) \neq 0$ 时, 以其除全式得 $f_1(X)/g_1(X)dX = g_2(y)/f_2(y)dy$ 两边分别积分就得到原方程的通积分

$$\int \frac{f_1(X)}{g_1(X)} dX = \int \frac{g_2(y)}{f_2(y)} dy + c$$

能使 $g_1(X) = 0$ 的 $X = X_i$ 和 $f_2(y) = 0$ 的 $y = y_i$ 均为原方程的解。注意 $\frac{dy}{dX} = f(X)g(y)$ 也是分离变量型的方程。

②齐次方程:其形式为

$$P(X, y)dX + Q(X, y)dy = 0 \quad (2)$$

其中 $P(X, y)$ 和 $Q(X, y)$ 是 m 次的齐次函数。令 $y = Xu$, 便可将

(2)化为分离变量型的方程 $[p(1,u)+uQ(1,u)]dX+XQ(1,u)du=0$ 。最简单的齐次方程形式为: $\frac{dy}{dX}=f(\frac{y}{X})$ 。

③一阶线性方程,其形式为

$$\frac{dy}{dX}+P(X)y=Q(X) \quad (3)$$

当 $Q(X) \neq 0$ 时,称为线性非齐次方程,当 $Q(X) \equiv 0$ 时,叫做线性齐次方程。线性齐次方程是变量分离方程,其通解为 $y=c e^{Xp\{-\int p(X)dX\}}$,利用它,令 $y=c(X)e^{Xp\{-\int p(X)dX\}}$ 为非齐次方程(3)的解,将它代入(3)中,可以求出 $C'(X)$,对 X 积分之后得到 $C(X)$,由此得到(3)的通解为

$$y=e^{-\int p(X)dX}\left\{C_1+\int Q(X)e^{\int p(X)dX}dX\right\}$$

这种变任意常数 C 为 X 的函数的方法,叫常数变易法。上式可作为公式来用,要想求当 $X=X_0$ 时 $y=y_0$ (即初值问题)的解,则有公式

$$y=e^{-\int_{X_0}^X p(X)dX}\left\{y_0+\int_{X_0}^X Q(\tau)e^{\int_{X_0}^{\tau} p(s)ds}d\tau\right\}$$

④伯努利方程,形状为 $\frac{dy}{dX}+P(X)y=Q(X)y^\alpha$ ($\alpha \neq 0,1$) 令 $u=y^{1-\alpha}$,就可以把伯努利方程化为一阶线性方程。

⑤黎卡提方程:其形状为 $\frac{dy}{dX}=p(X)y^2+Q(X)y+R(X)$,若已知它的一个特解 $y_1=\varphi(X)$,做变换 $y=u+\varphi(X)$,就变为伯努利方程,因而可解。一般,黎卡提方程没有初等解:1841年,刘维尔证明黎卡提方程 $y'+ay^2=bX^m$,除 $m=0,-2,-4k/(2k+1),-4k/(2k-1)$ 外,不存在初等形式的解。⑥恰当方程:若(2)的左端的式子恰为某二元函数 $u(X,y)$ 的全微分表达式,即 $du=PdX+Qdy=0$,就称(2)为全微分方程或恰当方程,此时 $u(X,y)=c$ 是(2)的通积分。

(2)为恰当方程的充要条件是: $\frac{\partial p(X,y)}{\partial y} \equiv \frac{\partial Q(X,y)}{\partial X}$, 当判定(2)是全微分方程之后, 由公式

$\int_{X_0}^X p(X,y)dX + \int_{y_0}^y Q(X_0,y)dy = c$ 或 $\int_{X_0}^X p(X,y_0)dX + \int_{y_0}^y Q(X,y)dy = c$ 便得到通积分。

⑦当(2)不是恰当方程, 但存在函数 $\mu(X,y) \neq 0$, 使 $\mu PdX + \mu Qdy = 0$ 是恰当方程时, 就称函数 $\mu(X,y)$ 为(2)的一个积分因子。为求积分因子, 利用 $\frac{\partial(\mu p)}{\partial y} \equiv \frac{\partial(\mu Q)}{\partial X}$, 即

$$\mu\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial X}\right) \equiv Q \frac{\partial \mu}{\partial X} - p \frac{\partial \mu}{\partial y} \quad (4)$$

也就是说, 需要解一个偏微分方程, 这反倒使问题复杂化。为此, 抛开“一般性”, 去求特殊的积分因子, 如 $\mu(X)$ 型, 或 $\mu(y)$ 型, $\mu(X \pm y)$ 型等, 有时是可能的, 例如, 为求 $\mu(X)$ 型的积分因子, 此时有 $\frac{\partial \mu}{\partial y} = 0$, 并且 $\frac{\partial \mu}{\partial X} = \frac{d\mu}{dX}$, (4) 成为

$$\frac{1}{\mu} \frac{d\mu}{dX} \equiv \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial X}\right) / Q \quad (5)$$

左端与 y 无关, 因而(5)的右端表达式仅为 X 的函数, 可以证明, 这个条件也是充分的。其他几种类型的积分因子可仿此推导。

⑧一阶隐方程形如

$$F(X, y, \frac{dy}{dX}) = 0 \quad (6)$$

的方程, 如有可能, 可以用参数法求解。在三维空间中, $F(X, y, z) = 0$ 一般是一个曲面, 设其参数表达式为

$$\begin{aligned} X &= f(u, v), & y &= g(u, v), \\ z &= h(u, v) \end{aligned} \quad (7)$$

由基本公式 $dy = y' dX$, 得到 u, v 的一阶显方程(注意以 y' 代 z)

$$\left(\frac{\partial g}{\partial u} - h \frac{\partial f}{\partial u}\right)du + \left(\frac{\partial g}{\partial v} - h \frac{\partial f}{\partial v}\right)dv = 0 \quad (8)$$

如能求得(8)的通积分 $y'(u, v, c) = 0$ 则 $X = f(u, v), y = g(u, v), \Phi(u, v, c) = 0$ 就是(6)的通积分。

⑨当(6)具有特殊形状如: $y = f(X, y')$, 令 $X = X, y' = p$, 由 $dy = pdX$, 得 $\frac{\partial f}{\partial X}dX + \frac{\partial f}{\partial p}dp = pdX$, 此时已化为显式方程了; 形如 $X = f(y, y')$ 的方程, 令 $y = y, y' = p$, 得到 $(1 - p \frac{\partial f}{\partial y})dy = \frac{\partial f}{\partial p}pdp$, 已呈显式方程。对 $F(X, y') = 0$ 及 $F(y, y') = 0$, 要先找 $F = 0$ 的参数表达式, 再行计算。

⑩奇解, 设有(6)的某解 $\Gamma: y = \varphi(X) \quad (X \in J)$, 如果对每一点 $P(X, y) \in \Gamma$, 在 P 点的任何邻域内, (6)有一个不同于 Γ 的解在 P 与 Γ 相切, 则称 Γ 是(6)的奇解。 p 判别曲线法(奇解的必要条件): 设函数 $F(X, y, p)$ 在 $(X, y, p) \in G$ 内连续, 且对 y 和 p 有连续的偏导数 F_y 和 F_p , 若 $y = \varphi(X)$ 是(6)的一个奇解, 当 $X \in J$ 时 $(X, \varphi(X), \varphi'(X)) \in G$, 则 $y = \varphi(X)$ 必满足 p 判别式

$$F(X, y, p) = 0, \quad F'_p(X, y, p) = 0$$

其中 $p = \frac{dy}{dX}$ 。

⑪包络, 设给出单参数 C 的曲线族

$$V(X, y, c) = 0 \quad (9)$$

及连续可微的曲线 Γ 。如果对于 Γ 中任意一点 P , 都有(9)中的一条曲线 $K(c^*)$ 在 P 点与 Γ 相切, 并在 P 的某一邻域内不同于 Γ , 就称 Γ 为(9)的一支包络。包络必满足 c 判别式:

$$V(X, y, c) = 0 \quad V'_c(X, y, c) = 0 \quad (10)$$

又若由(10)确定一支连续可微的曲线

$$\Lambda: \quad X = \varphi(c), \quad y = \Psi(c)$$

$(c \in J)$ 且满足非蜕化性条件: $(\varphi'(c), \Psi'(c)) \neq (0, 0)$ 以及 $V'_x + V'_y \neq 0$, 其中 $V'_x = V'_x(\varphi(c), \Psi(c), c)$ 和 $V'_y = V'_y(\varphi(c), \Psi(c), c)$, 则 Λ 是 (9) 的一支包络。

【高阶方程的几种降阶法】(1) $y^{(n)} = f(X)$ 对 X 积分 n 次, 运用累次积分的性质, 得通解

$$y = \frac{1}{(n-1)!} \cdot \int_{X_0}^X (X-\tau)^{n-1} f(\tau) d\tau + \sum_{i=1}^n c_i (X-X_0)^{n-i}; (2) y'' =$$

$f(y)$, 以 $2y'$ 乘全式再对 X 积分可得 $(y')^2 = 2 \int f(y) dy + c_1$ 已化为一阶方程了; (3) $F(X, y^{n-k}, \dots, y^{(n)}) = 0$, 令 $z = y^{(n-k)}$, 化为 $F(X, z, \dots, z^{(k)}) = 0$; (4) $F(y, y', \dots, y^{(n)}) = 0$, 令 $y' = u$, 取 y 为自变量, 则 $\frac{d^2 y}{dX^2} = u \frac{du}{dy}, \frac{d^3 y}{dX^3} = u^2 \frac{d^2 u}{dy^2} + u \left(\frac{du}{dy}\right)^2$, 利用数学归纳法可证 $\frac{d^k y}{dX^k} = \Phi(u, \frac{du}{dy}, \dots, \frac{d^{k-1} y}{dy^{k-1}})$, 因之可将原方程降低一阶。

【方程组与首次积分】设给出微分方程组

$$\frac{dy_i}{dX} = f_i(X, y_1, y_2, \dots, y_n) \quad (i=1, 2, \dots, n) \quad (11)$$

其中右端函数 $f_1, f_2, \dots, f_n$ 在某个区域 $D \subset R^{n+1}$ 内对 $(X, y_1, \dots, y_n)$ 连续, 对 $(y_1, y_2, \dots, y_n)$ 连续可微, 设函数 $\Phi(X, y)$ 在 $G \subset D$ 内连续可微, 不是常数, 但沿着 (11) 在 G 内的任一积分曲线

$$\Gamma: y_1 = y_1(X), y_2 = y_2(X), \dots, y_n = y_n(X) \quad (X \in J)$$

函数 $\Phi(X, y, \dots, y) (X, y_1, \dots, y_n)$ 都取常数值; 亦即 $\Phi(X, y_1(X), \dots, y_n(X)) \equiv \text{常数} (X \in J)$, 则称 $\Phi(X, y_1, \dots, y_n) = c$ 为 (11) 的在 G 内的一个首次积分。有时也称 $\Phi(X, y_1, \dots, y_n)$ 为首次积分。

判定准则: 在 G 内连续可微的函数 $\Phi(X, y_1, \dots, y_n)$ 为 (11) 的首次

积分的充要条件是: $\frac{\partial \Phi}{\partial X} + \frac{\partial \Phi}{\partial y_1} f_1 + \cdots + \frac{\partial \Phi}{\partial y_n} f_n \equiv 0, (X, y_1, \cdots, y_n) \in G$ 。

首次积分与解的关系: ①若知道(11)的一个首次积分, 便可将它降低一阶。②在 G 内若已知(11)的 n 个独立的首次积分

$$\begin{aligned} \Phi_i(X, y_1, y_2, \cdots, y_n) &= c_i \\ (i=1, 2, \cdots, n) \end{aligned} \quad (12)$$

则由它们可确定(11)在 G 内的通解:

$$\begin{aligned} y_i &= \varphi_i(X, c_1, c_2, \cdots, c_n) \\ (i=1, 2, \cdots, n) \end{aligned}$$

所谓 $\Phi_i (i=1, 2, \cdots, n)$ 相互独立, 指雅可比行列式

$$\frac{D(\Phi_1, \cdots, \Phi_n)}{D(y_1, \cdots, y_n)} \neq 0。$$

③(11)最多有 n 个相互独立的首次积分。首次积分的求法: 一般采用拼凑“可积组合”的办法, 例如, $\frac{dX}{dt} = y, \frac{dy}{dt} = X$, 将两个方程相加

得 $\frac{d(X+y)}{dt} = X+y$, 于是便得到一个首次积分: $(X+y)e^{-t} = c$ 又

例如, 对于对称形式的方程组

$$\frac{dX_1}{X_1(X_1, \cdots, X_n)} = \frac{dX_2}{X_2(X_1, \cdots, X_n)} = \cdots = \frac{dX_n}{X_n(X_1, \cdots, X_n)}$$

若能运用比例的法则凑成

$$\frac{dX_1}{X_1(X_1, \cdots, X_n)} = \cdots = \frac{d\Phi_1}{\Phi_1} = \frac{d\Phi_2}{\Phi_2}$$

则由此可得到一个首次积分: $\Phi_2 = c\Phi_1$, 凑成其他可以积分的形式也可。

【线性方程组】称

$$\frac{d\vec{y}}{dX} = A(X)\vec{y} + \vec{f}(X) \quad (13)$$

为线性微分方程组,其中 $A(X)=$

$$\begin{pmatrix} a_{11}(X) & a_{12}(X) & \cdots & a_{1n}(X) \\ a_{21}(X) & a_{22}(X) & \cdots & a_{2n}(X) \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}(X) & a_{n2}(X) & \cdots & a_{nn}(X) \end{pmatrix}$$

$$\vec{f}(X) = \begin{pmatrix} f_1(X) \\ f_2(X) \\ \cdots \\ f_n(X) \end{pmatrix}$$

也就是说,未知函数和它们的导数都是一

次的,

$a_{ij}(X), f_i(X) (i, j=1, 2, \cdots, n)$ 均在区间 J 上有定义。(13) 中的 $\vec{f}(X) \not\equiv \vec{0}$ 时, 称为非齐次线性微分方程组, 当 $\vec{f}(X) \equiv \vec{0}$ 时, 即

$$\frac{d\vec{y}}{dX} = A(X)\vec{y} \quad (14)$$

称为齐次线性方程组, 可以仿照线性代数的方法定义函数 $\vec{y}_1(X), \vec{y}_2(X), \cdots, \vec{y}_m(X)$ 的线性相关性。而当 $m=n$ 时, 定义 $\vec{y}_1(X), \vec{y}_2(X), \cdots, \vec{y}_n(X)$ 的朗斯基行列式为

$$W(X) = \begin{vmatrix} y_{n1}(X) & y_{12}(X) & \cdots & y_{n1}(X) \\ y_{21}(X) & y_{22}(X) & \cdots & y_{2n}(X) \\ \cdots & \cdots & \cdots & \cdots \\ y_{n1}(X) & y_{n2}(X) & \cdots & y_{nn}(X) \end{vmatrix}$$

当 $\vec{y}_1(X), \cdots, \vec{y}_n(X)$ 为 (14) 的解时, 有公式:

$$W(X) = W(X_0) e^{\int_{X_0}^X \sum_{i=1}^n a_{ii}(\tau) d\tau}$$

可见, $W(X_0) = 0 \Leftrightarrow W(X) \equiv 0 \quad \forall X \in J \Leftrightarrow \vec{y}_1(X), \cdots, \vec{y}_n(X)$ 线性相关。(14) 的解的研究: ① $\vec{y}(X) \equiv 0$ 是它的解, 称为平凡解; ② 存在 n 个线性无关解, 称做一个基本解组, (14) 的任意一个解都可用基本解组的线性组合来表示; ③ (14) 的解的全体构成一个 n 维线性空间; ④ 基本解组中的 n 个函数的 n^2 个元素排成的方阵称为它的一个基本方阵, 通常记为 $Y(X)$; ⑤ (13) 的通解等于它的一个特解与

(14)的通解之和,并有公式

$$\vec{y}(X) = y(X) \vec{y}_0 + \int_{X_0}^X y(X) y^{-1}(\tau) \cdot \vec{f}(\tau) d\tau, \text{ 其中 } Y(X) \text{ 为 (14)}$$

的基本方阵。

【高阶线性方程】是线性方程组的特殊情况,其形式为

$$y^{(n)} + a_1(X)y^{(n-1)} + a_2(X)y^{(n-2)} + \cdots + a_n(X)y = f(X) \quad (15)$$

其中 $a_i(X) (i=1, 2, \cdots, n)$ 与 $f(X)$ 均于区间 J 上定义, 令 $y_1 = y$, $y_2 = y'$, $y_3 = y''$, $\cdots$, $y_n = y^{(n-1)}$, 则可将(15)化为线性方程组(13), 其中 $A(X) =$

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_1(X) & -a_2(X) & -a_3(X) & \cdots & -a_{n-1}(X) & -a_n(X) \end{pmatrix}$$

$$\vec{f}(X) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(X) \end{pmatrix}$$

此时它的 n 个解的朗斯基行列式为

$$W(X) = W[y_1(X), \cdots, y_n(X)] = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \cdots & \cdots & \cdots & \cdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}$$

特例: (i) 一阶线性方程, 有公式解(参见“常微分方程的初等解法”) (ii) 二阶线性方程 $y'' + p(X)y' + q(X)y = 0$, 若已知它的一个非零解 $y_1(X)$ 利用朗斯基行列式可得到与 $y_1(X)$ 线性无关的特解,

进而可得通解:

$$y(X) = y_1(X) [C_2 + C_1 \int_X^X \frac{1}{y_1^2(s)} e^{-\int_{X_0}^s P(s) ds} d\tau]$$

【常系数线性方程与方程组】当(15)中的所有的系数 $a_i(X)$ 均为常数时, 即方程为

$$\frac{d^n y}{dX^n} + a_1 \frac{d^{n-1} y}{dX^{n-1}} + \cdots + a_n y = 0 \quad (16)$$

叫做常系数线性方程, 求(16)的形如 $y = e^{\lambda X}$ 的解则寻求 λ 所应满足的代表方程

$$\lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1} \lambda + a_n = 0 \quad (17)$$

称之为特征方程, 它的根叫做特征根。①若 n 个特征根 $\lambda_1, \lambda_2, \dots, \lambda_n$ 互不相同, 则 $y_1 = e^{\lambda_1 X}, y_2 = e^{\lambda_2 X}, \dots, y_n = e^{\lambda_n X}$ 是(16)的一个基本解组; ②若 λ_i 为 n_i 重根 ($i=1, 2, \dots, k, k < n$), 则

$$\begin{array}{cccc} e^{\lambda_1 X} & X e^{\lambda_1 X} & X^{n_1-1} e^{\lambda_1 X} & \\ \dots & \dots & & \\ e^{\lambda_k X} & X e^{\lambda_k X} & X^{n_k-1} e^{\lambda_k X} & \end{array}$$

是一个基本解组, ③当某特征根为复数 $\lambda = \alpha + i\beta$ 并且是 L 重根时, 则

$$\begin{array}{cccc} e^{\alpha X} \cos \beta X, & X e^{\alpha X} \cos \beta X, & \cdots, & X^{L-1} e^{\alpha X} \cos \beta X \\ e^{\alpha X} \sin \beta X, & X e^{\alpha X} \sin \beta X, & \cdots, & X^{L-1} e^{\alpha X} \sin \beta X \end{array}$$

是与特征根 $\lambda = \alpha \pm i\beta$ 相对应的 $2L$ 个实的解。

常系数线性方程组, 当(13)中所有的 $a_{ij}(X)$ 都是常数时, 便叫做常系数线性齐次方程组:

$$\frac{d\vec{y}}{dX} = \vec{A} \vec{y} \quad (18)$$

定义 A 的模为 $\|A\| = \sum_{i,j=1}^n |a_{ij}|$, 显然 $\|A\| \geq 0$, 当且仅当 $A=0$

(即 A 为零矩阵)时; $\|A\| = 0$ $\|A+B\| \leq \|A\| + \|B\|$; $\|AB\| \leq \|A\| \cdot \|B\|$ 矩阵 A 的幂级数

$$e^A = E + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{k!}A^k + \cdots$$

称为矩阵 A 的指数函数,其中 E 为单位矩阵,① $\Phi(X) = e^{XA}$ 是 (18) 的一个标准基解矩阵,即具有 $\Phi(0) = E$; ② 常系数非齐次方

程组 $\frac{d\vec{y}}{dX} = \vec{A}y + \vec{f}(X)$ 。它的通解为

$$\vec{y} = e^{XA} \vec{c} + \int_{X_0}^X e^{(X-\tau)} \vec{A} f(\tau) d\tau \quad (19)$$

其中 c 为任意的常数列向量,当把 (19) 中的 $\vec{c}$ 换为 $\vec{y}_0$ 时,即得到方程组具有初值 $\vec{y}(X_0) = \vec{y}_0$ 的解; ③ 存在非奇异矩阵 P , 使 $A = PJP^{-1}$, 其中 J 为约当矩阵

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}, \text{ 而}$$

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & \lambda_i \end{pmatrix}$$

J_i 为 n_i 阶的 ($i = 1, 2, \cdots, k, n_1 + n_2 + \cdots + n_k = n$), $e^{XA} = Pe^{XJ}P^{-1}$ 。其中

$$e^{XJ} = \begin{pmatrix} e^{XJ_1} & & & \\ & e^{XJ_2} & & \\ & & \ddots & \\ & & & e^{XJ_k} \end{pmatrix} \text{ 而 } e^{XJ_i} =$$

$$e^{\lambda_i X} \begin{pmatrix} 1 & X & \frac{X^2}{2!} \cdots & \frac{X^{n_i-1}}{(n_i-1)!} \\ & 1 & X \cdots & \frac{X^{n_i-2}}{(n_i-2)!} \\ & & \ddots & \ddots \\ & & & 1 \\ & & & & 1 \end{pmatrix}$$

$$(i=1, 2, \dots, k)$$

【用待定指数函数法求解常系数线性方程组】①当 A 的特征根 $\lambda_1, \lambda_2, \dots, \lambda_n$ 均为单根, 基解矩阵为: $\Phi(X) = (e^{\lambda_1 X} \vec{\Gamma}_1, e^{\lambda_2 X} \vec{\Gamma}_2, \dots, e^{\lambda_n X} \cdot \vec{\Gamma}_n)$

其中 $\vec{\Gamma}_i$ 是与 λ_i 相应的特征向量; ②设 λ_i 的重数为 $n_i (n=1, 2, \dots, k, k < n)$, 则与 λ_i 相应的解为

$$\vec{y} = e^{\lambda_i X} + \left(\vec{\Gamma}_0 + \left[\frac{X}{1!} \vec{\Gamma}_1 + \frac{X^2}{2!} \vec{\Gamma}_2 + \cdots + \frac{X^{n_i-1}}{(n_i-1)!} \vec{\Gamma}_{n_i-1} \right] \right)$$

$$(20)$$

其中 $\vec{\Gamma}_0$ 满足

$$(A - \lambda_i E)^{n_i} \vec{\Gamma}_0 = \vec{0}$$

$$(i=1, 2, \dots, k) \quad (21)$$

依次计算 $\vec{\Gamma}_1 + (A - \lambda_i E) \vec{\Gamma}_0, \vec{\Gamma}_2 = (A - \lambda_i E) \vec{\Gamma}_0, \dots, \vec{\Gamma}_{n_i-1} = (A - \lambda_i E)^{n_i-1}$

$$(i=1, 2, \dots, k) \quad (22)$$

(21) 必有 n_i 个 $\vec{\Gamma}_0$ 从每个 $\vec{\Gamma}_0$ 可由 (22) 得到 $\vec{\Gamma}_1, \dots, \vec{\Gamma}_{n_i-1}$ 由各组 $\vec{\Gamma}_j$ 构

成 n_i 个线性无关解。将这 $n_1, n_2, \dots, n_k$ 个解合并在一起即得到一个基解矩阵

$$(e^{\lambda_1 X} P_1^{(1)}(X), \dots, e^{\lambda_1 X} P_{n_1}^{(1)}(X); \dots, e^{\lambda_k X} \vec{P}_1^{(k)}(X), \dots, e^{\lambda_k X} \vec{P}_{n_k}^{(k)}(X))。$$

【常微分方程的幂级数解法】①设 $f(X, y)$ 在矩形区域 $R(R: |X - X_0| \leq a, |y - y_0| \leq b, a, b > 0)$ 内可以展成 $(X - X_0)$ 和 $(y - y_0)$ 的收敛的幂级数。则初值问题

$$\frac{dy}{dX} = f(X, y), y(X_0) = y_0 \quad (23)$$

在 X_0 点的邻域 $|X - X_0| < \rho$ 内有唯一的解析解, 其中数 $\rho > 0$; ②二阶方程

$$y'' + p(X)y' + q(X)y = 0 \quad (24)$$

的幂级数解法。设 $p(X)$ 和 $q(X)$ 在区间 $|X - X_0| < p$ 内可展成 $(X - X_0)$ 的收敛幂级数。则方程(24)在 $|X - X_0| < p$ 内有收敛的幂

级数解 $y = \sum_{n=0}^{\infty} c_n (X - X_0)^n$ 其中 c_0, c_1 是两个任意常数, $c_i (i \geq 3)$ 则

由递推公式确定。例: 勒让德方程

$(1 - X^2)y'' - 2Xy' + n(n+1)y = 0$ (n 为常数, $-1 < X < 1$) 在 $X = 0$ 的邻域内有级数解

$$y = \sum_{k=0}^{\infty} [c_{2k} X^{2k} + c_{2k+1} X^{2k+1}] = c_0 y_1(X) + c_1 y_2(X), \text{ 其中}$$

$$c_{2k} = (-1)^k [(n-2k+2) \cdots (n-2) \cdot n(n+1)(n+3) \cdots (n+2k-1)] \cdot \frac{c_0}{(2k)!}$$

$$c_{2k+1} = (-1)^k [(n-2k+1) \cdots (n-3)(n-1)(n+2) \cdots (n+2k)] \cdot \frac{c_1}{(2k+1)!}$$

当 n 为偶数时, y_1 是 n 次多项式; n 为奇数时, $y_2(X)$ 是 n 次多项

式, 顺次令 $n=1, 2, \dots$, 就得到著名的勒让得多项式: $p_0(X), p_1(X), p_2(X), \dots$, 这是一族正交函数系, 对一个定义于 $[-1, 1]$ 上的函数 $f(X)$, 可作关于 $p_n(X)$ 的广义傅里叶级数:

$f(X) \sim \sum_{n=0}^{\infty} a_n p_n(X)$, 其中广义傅里叶系数为 $a_n = \frac{2n+1}{2} \int_{-1}^1 f(X) p_n(X) dX$ 当 X_0 是 $f(X)$ 的连续点时, 它的广义傅里叶级数收敛到 $f(X_0)$, ③对首项系数不为 1 的方程

$$A(X)y'' + B(X)y' + R(X)y = 0 \quad (25)$$

设 X_0 为正则奇点, 即 $A(X) = (X - X_0)^2 P(X), B(X) = (X - X_0)Q(X)$ 这里

$\rho(X), Q(X)$ 与 $R(X)$ 在 X_0 点邻域内解析, $P(X_0) \neq 0$, 则 (25) 有收

敛的广义幂级数解 $y = \sum_{k=0}^{\infty} c_k (X - X_0)^{k+\rho}$, ($c_0 \neq 0$), 称 ρ 为指标, ρ 由指标方程确定, $c_k (k \geq 1)$ 可由递推公式来确定。例如, 贝塞尔方程:

$$X^1 y'' + X y' + (X^2 - n^2) y = 0 \quad (n \geq 0) \quad (26)$$

$X=0$ 为它的正则奇点, 它有指标方程 $(\rho+n)(\rho-n)=0$ $a. \rho_1 = n$, 相应的广义幂级数解为

$$J_n(X) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(n+k+1)\Gamma(k+1)} \cdot \left(\frac{X}{2}\right)^{2k+n}$$

其中 $\Gamma(k+1)$ 是 Γ 函数, 称 $J_n(X)$ 为第一类贝塞尔函数。 $b. \rho_2 = -n$, 而 $2n \neq$ 整数, 或 $2n$ 为奇数时, 有

$$J_{-n}(X) =$$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(-n+k+1)\Gamma(k+1)} \left(\frac{X}{2}\right)^{2k-n}$$

当 $2n$ 为偶数时, 则有解

$$y_n(X) = \lim_{a \rightarrow n} \frac{J_a(X) \cos a\pi - J_{-a}(X)}{\sin a\pi}$$

称之为诺伊曼函数或(相应于整数 n 的)第二类贝塞尔函数。 $J_n(X)$ 有无穷多个零点,依次排列为

$$0 < \beta_1 < \beta_2 < \cdots < \beta_k < \cdots, (\beta_k \rightarrow \infty) \quad (27)$$

在 $0 \leq t \leq 1$ 上,函数系

$J_n(\beta_1 t), J_n(\beta_2 t), \cdots, J_n(\beta_n t), \cdots$ 是正交的。若 $f(X)$ 在 $0 \leq X \leq 1$ 上定义,在适当条件下可展为级数:

$$f(X) \sim \sum_{k=1}^{\infty} a_k J_n(\beta_k X),$$

其中傅里叶系数

$$a_k = \frac{2}{[J_n(\beta_k)]^2} \int_0^1 t f(X) J_n(\beta_k t) dt$$

当 $f(X)$ 在 X_0 点连续,则广义幂级数就收敛于 $f(X_0)$ 。

【常微分方程初值问题】为了求特定的解,需要“定解条件”,初始条件是一类重要的定解条件,一阶微分方程为

$$\frac{dy}{dX} = f(X, y) \quad (E)$$

它的初始条件是

$$y(X_0) = y_0 \quad (I)$$

方程和微分方程初始条件一起,就构成了常微分方程的初值问题,也叫柯西问题,即求方程(E)的满足初始条件(I)的解,有时简记为(E)+(I)

已知以 (X_0, y_0) 为中心的矩形域 $R: |X - X_0| \leq a, |y - y_0| \leq b, a > 0, b > 0$

李普希茨条件:若存在 $L > 0$,使不等式

$$|f(X, y_1) - f(X, y_2)| \leq L |y_1 - y_2|$$

对任何 (X, y_1) 和 $(X, y_2) \in R$ 都成立。则称函数 $f(X, y)$ 在 R 内关于 y 满足李普希茨条件。

皮卡定理: 设微分方程 (E) 中的函数 $f(X, y)$ 在 R 内连续, 并且关于 y 满足李普希茨条件, 则 $(E) + (I)$ 的解于 $|X - X_0| \leq h$ 上存在, 而且是唯一的。

柯西-皮亚诺定理: 若 $f(X, y)$ 在 R 内连续, 则 $(E) + (I)$ 的解于 $|X - X_0| \leq h$ 上存在。

两个定理中的 $h = \min(a, \frac{b}{M})$, $M = \max_{(X, y) \in R} |f(X, y)|$ 。

解的延拓: 上述存在定理, h 往往“很短”, 延拓就是为了解决这个问题, 办法是反复运用存在定理, 把积分曲线一段一段光滑地延伸出去。从皮卡定理, $(E) + (I)$ 的右行解(从 X_0 出发, 在 $X > X_0$ 方向的解)到达 $X_1 = X_0 + h$, 当 $(X_1, \varphi(X_1)) \in G$ 时, 可作矩形 $R_1: |X - X_1| \leq a, |y - y_1| \leq b_1$, 使 $R_1 \subset G$, 记 $y(X_1) = y_1$, 则又组成了初值问题: $\frac{dy}{dX} = f(X, y), y(X_1) = y_1$ 应用皮卡定理, 他的解 $y = \varphi_1(X)$

于 $|X - X_1| \leq h_1$ 上存在, $h_1 = \min(a, \frac{b_1}{M_1})$, $M_1 = \max_{(X, y) \in R_1} |f(X, y)|$,

把两次所得的解“接起来”, 即

$$y = \begin{cases} \varphi(X) & X_0 - h \leq X \leq X_0 + h \\ \varphi_1(X) & X_0 + h \leq X \leq X_0 + h + h_1 \end{cases}$$

右行解已延伸到 $X_0 + h + h_1$, 这就叫做解的延拓。左行解可类似说明。现在的问题是: 这种延拓的“最终结局”怎样呢? 为此先给出饱和解的定义。若一个解不能再行延拓, 就称它为饱和解。饱和解相应的区间称为该解的最大存在区间。一个结论: 最大存在区间必为开区间。

定理: 设 $(E) + (I)$ 的解 $y = \varphi(X)$

的最大存在区间为 $\alpha < X < \beta$, 则当 G 为有限区域时, 有

$$\lim_{X \rightarrow \alpha+} \rho((X, \varphi(X)), \partial G) = 0;$$

$$\lim_{X \rightarrow \beta-} \rho((X, \varphi(X)), \partial G) = 0$$

式中 $\rho(\cdot, \cdot)$ 表示欧氏距离, ∂G 表示 G 的边界, 含义是: 当自变量 X 趋于最大存在区间的左、右端点时, 积分曲线趋近于区域 G 的边界。

当 G 是无限区域时, 解可以延拓到任意接近 ∂G , 或趋向无穷远的地方。

利用柯西-皮亚诺定理同样可以作解的延拓。

【解对初值和参数的连续相依性】这里讨论的问题是: 初值(或参数)有微小的变化时, 解是否也只有微小的变化? ①首先设函数 $f(X, y, \lambda)$ 在区域 $G: |X - X_0| \leq a, |y - y_0| \leq b, |\lambda - \lambda_0| \leq c$ 上连续, 关于 y 满足李普希茨条件: $|f(X, y_1, \lambda) - f(X, y_2, \lambda)| \leq L(y_1 - y_2)$, 其中 $L > 0$, 则初值问题: $\frac{dy}{dX} = f(X, y, \lambda), y(X_0) = y_0$ 的解 $y = \varphi(X, \lambda)$ 在区域 $D: |X - X_0| \leq h, |\lambda - \lambda_0| \leq c$ 上是连续的, 其中 $h = \min(a, \frac{b}{M})$, $M = \max_{(X, y, \lambda) \in G} |f(X, y, \lambda)|$ ②设函数 $f(X, y)$ 在 G 上连续, 对 y 满足李普希茨条件, $y = \xi(X)$ 是 (E) 的一个解, 其存在区间为 J , 任取 $[a, b] \subset J$, 则 $\exists \delta > 0$, 使对任何初值 $(X_0, y_0): a \leq X_0 \leq b, |y_0 - \xi(X_0)| \leq \delta$, 柯西问题 $(E) + (I)$ 的解 $y = \varphi(X, X_0, y_0)$ 也在 $a \leq X \leq b$ 上存在, 并且对 (X, X_0, y_0) , 在区域 $D_\delta: a \leq X \leq b, a \leq X_0 \leq b, |y_0 - \xi(X_0)| \leq \delta$ 上是连续的。

【解对初值和参数的可微性】这里讨论的是比连续依赖性更进一步

的问题,即解作为初值和参数的函数是不是可微的? 设函数 $f(X, y, \lambda)$ 以及 $f(X, y, \lambda)$ 对 X, y 和 λ 的偏导数都连续。(在区域 $D_\lambda: (X, y) \in G, \lambda \in I_\lambda$), 则初值问题

$\frac{dy}{dX} = f(X, y, \lambda), y(X_0) = y_0$ 的解 $y = \varphi(X, X_0, y_0, \lambda)$ 作为 X, X_0, y_0, λ

的函数,在其定义域内连续可微。并且 $\frac{\partial \varphi}{\partial X_0}$ 和 $\frac{\partial \varphi}{\partial y_0}$ 均满足“变分方程”:

$$\frac{dz}{dX} = f'_y(X, \varphi(X, X_0, y_0, \lambda), \lambda)z$$

的解,初值分别是 $z(X_0) = -f(X_0, y_0)$ 及 $z(X_0) = 1, \frac{\partial \varphi}{\partial \lambda}$ 则是初值问题

$$\begin{cases} \frac{dz}{dX} = f'_y(X, \varphi(X, X_0, y_0, \lambda), \lambda)z + f'_\lambda \\ (X, \varphi(X, X_0, y_0, \lambda), \lambda) \\ z(X_0) = 0 \end{cases}$$

的解。

【第一比较定理】设 $f(X, y)$ 与 $F(X, y)$ 都在平面区域 G 内连续,并且满足不等式

$$f(X, y) < F(X, y),$$

$$(X, y) \in G$$

又设 $y = \varphi(X)$ 与 $y = \Phi(X)$ 在区间 $a < X < b$ 上分别是初值问题

$$\frac{dy}{dX} = f(X, y), y(X_0) = y_0(E_1)$$

$$\text{与 } \frac{dy}{dX} = F(X, y), y(X_0) = y_0(E_2)$$

的解,其中 $(X_0, y_0) \in G$,则有

$$\varphi(X) < \Phi(X), \text{当 } X_0 < X < b,$$

$$\varphi(X) > \Phi(X), \text{当 } a < X < X_0.$$

【初值问题的推广】设 $y = \varphi(X)$ 是区间 I 上的绝对连续函数,对 $X \in I$ 时除了在一个测度为零的集合上以外,满足 (E) ,就称 $y = \varphi(X)$ 是 (E) 的卡拉西奥多里解。定理:设 $f(X, y)$ 在 G 上定义,对每个固定的 y ,关于 X 可测,对每个固定的 X 关于 y 连续,又对任何有界闭域 $D \subset G$,存在勒贝格可积函数 $m(X)$,使当 $(X, y) \in D$ 时, $|f(X, y)| \leq m(X)$,则 $(E) + (I)$ 有一个卡拉西奥多里解。特别,当 $f(X, y)$ 是 G 上的连续函数时,卡拉西奥多里解就是通常意义下的解。

【常微分方程边值问题】这是另一类定解问题,是求给定的微分方程满足边值条件的解,它在流体力学,波动力学,现代控制论中有重要的应用。

斯图姆比较定理:讨论下面的方程

$$y'' + P(X)y' + q(X)y = 0 \quad (1)$$

与

$$y'' + P(X)y' + Q(X)y = 0 \quad (2)$$

设 $P(X), q(X)$ 与 $Q(X)$ 在区间 J 上连续,则有:①(1)(及(2))的任意非零解在 J 内的零点都是孤立的;②(1)(及(2))的任意两个线性无关解的零点都是相互交错的;③设 $Q(X) \geq q(X)$,则(1)的任意非零解的两个相邻零点之间必有(2)的任意非零解的至少一个零点。

含参数的二阶线性微分方程

$$\begin{aligned} & [p(X)y'(X)]' + [q(X) + \lambda r(X)] \\ & y(X) = f(X). \end{aligned} \quad (3)$$

第一边界条件: $U_1(X) \equiv y(a) = r_1$,

$$U_2(X) \equiv y(b) = r_2 \quad (I)$$

第二边界条件: $U_1(X) \equiv y'(a) = r_1$,

$$U_2(X) \equiv y'(b) = r_2 \quad (II)$$

第三边界条件:

$$\begin{cases} U_1(X) \equiv \alpha_1 y(a) + \alpha_2 y'(a) = r_1 \\ U_2(X) \equiv \beta_1 y(b) + \beta_2 y'(b) = r_2 \end{cases} \quad (III)$$

(3)与(I)—(III)分别组成第一、第二与第三边值问题。当 $r_1 = r_2 = 0$ 时,与 $f(t) = 0$ 时的(3)分别称为第一、第二、第三齐次边值问题。

【特征值问题】含参数的齐次边值问题,称为特征值问题。(3)与(III)通常称为斯图姆-刘维尔边值问题,简称为 $S-L$ 问题。若当 $\lambda = \lambda_0$ 时,此边值问题有非零解 $y = \varphi_0(X)$,则称 λ_0 为这个边值问题的特征值,相应的函数 $y = \varphi_0(X)$ 称为特征函数。

例:求边值问题 $y'' + \lambda y = 0, y(0) = 0, y(1) = 0$ 的特征值和特征函数。① $\lambda \leq 0$ 时无解,即一切负的或等于零的常数 λ 都不是这个边值问题的特征值;② $\lambda > 0$ 时,方程的通解

$$y = c_1 \cos \sqrt{\lambda} X + c_2 \sin \sqrt{\lambda} X,$$

把边值条件代入其中得: $y(0) = c_1 = 0; y(1) = c_1 \cos \sqrt{\lambda} 1 + c_2 \sin \sqrt{\lambda} 1 = 0$,故 $\sin \sqrt{\lambda} 1 = 0$,于是 $\lambda = \lambda_n = \left[\frac{n\pi}{1} \right]^2, (n=1, 2, \dots)$ 。相应的特征函数为 $y = \varphi_n(X) = \sin \frac{n\pi}{1} X, (n=1, 2, \dots)$ 。由此可得到特征值

与特征函数的一些重要特性:① $S-L$ 边值问题有无限多个(简单的)特征值,而且可以排列为:

$\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < \cdots$ 并且 $\lim_{k \rightarrow \infty} \lambda_k = \infty$; ②与 $\{\lambda_k\}$ 相应的特征函数构成的特征函数系 $\{\varphi_k(X)\}$ 是正交系。这个例子虽然特殊,但已经证明这两条重要性质对一般的 $S-L$ 问题也是成立的。一般,可把(3)与(Ⅲ)化为如下典型的 $S-L$ 边值问题:

$$\begin{cases} y'' + [\lambda + q(X)]y = 0 \\ y(0)\cos\alpha + y'(0)\sin\alpha = 0, \quad y(1)\cos\beta - y'(1)\sin\beta = 0 \end{cases}$$

其中 $q(X)$ 于 $0 \leq X \leq 1$ 上有定义而且连续。则特征值与特征函数有上面例子中的两条性质,于是可考虑定义于 $0 \leq X \leq 1$ 上的函数 $f(X)$ 展成 $\{\varphi_k(X)\}$ 的(广义的)傅里叶级数的问题。

【常微分方程定性理论】1841 年刘维尔证明黎卡提方程一般无初等解,从理论上结束了对每个微分方程都去求初等解的探索。但是许多实际问题(天文的,力学的……)归结为微分方程,需要掌握它的解的性质,因此,定性理论便成了重要的方法和工具。 H . 庞加莱的《微分方程所定义的积分曲线》开拓了常微分方程定性理论的方向,它一度成为研究常微分方程的主流方向之一。1926 年荷兰工程师范德坡建立了三极管振荡电路的数学模型,即著名的范德坡方程: $\ddot{X} + \mu(X^2 - 1)\dot{X} + X = 0 (\mu > 0)$ 。前苏联的数学家们则把这项技术工作和庞加莱的定性理论联系起来,大大促进它的理论发展和广泛应用。以二维情况为例,称

$$\frac{dX}{dt} = P(X, y), \quad \frac{dy}{dt} = Q(X, y) \quad (1)$$

为自治(定常)的系统,(1)的右端在 (X, y) 平面上定义了一个向量场。称 (X, y) 平面为相平面,解 $X = X(t), y = y(t)$ 在相平面上描出

的光滑曲线叫做轨线。

奇点,使 $P(X_0, y_0) = Q(X_0, y_0) = 0$ 的点 (X_0, y_0) 叫做系统(1)的奇点,否则就称为常点。关于奇点的研究是定性理论的基础工作之一。对于常点有一个重要定理:(1)的轨线在常点附近是一族(在拓扑意义下的)平行直线。而奇点附近的轨线则是多姿多彩的。例如① $\dot{X} = X, \dot{y} = -y, (0, 0)$ 是奇点,叫做鞍点(图 1);② $\dot{X} = -X, \dot{y} = -2y, (0, 0)$ 是奇点,叫做稳定结点(图 2);③ $\dot{X} = -y, \dot{y} = X, (0, 0)$ 是奇点,叫做中心(图 3)。还有一些其他类型的奇点。

极限环:系统(1)的孤立的闭轨线称做极限环,例如系统 $\dot{X} = y + X(1 - X^2 - y^2), \dot{y} = -X + y(1 - X^2 - y^2)$ 化为极坐标: $\frac{dr}{dt} = r(1 - r^2),$

$\frac{d\theta}{dt} = -1$ 奇点为 $X = y = 0, r = 1$ 是极限环,若初值点 P_1 在 $r = 1$ 内

部,当 t 增加时,由 $\frac{dr}{dt} > 0$ 知, r 是增加的,并且 $\lim_{t \rightarrow +\infty} r = 1$ 。极限环:

系统(1)的孤立的闭轨线称做极限环,例如系统 $\dot{X} = y + X(1 - X^2 - y^2), \dot{y} = -X + y(1 - X^2 - y^2)$ 化为极坐标: $\frac{dr}{dt} = r(1 - r^2), \frac{d\theta}{dt} = -$

1 奇点为

图 5

图 6

图 7

$X=y=0, r=1$ 是极限环, 若初值点 P_1 在 $r=1$ 内部, 当 t 增加时, 由 $\frac{dr}{dt} > 0$ 知, r 是增加的, 并且 $\lim_{t \rightarrow +\infty} r = 1$ 。

若初始点 P_2 在 $r=1$ 的外部, $\frac{dr}{dt} < 0$, 当 t 增加时, r 减少, 并有 $\lim_{t \rightarrow +\infty} r = 1$, $r=1$ 这个极限环, “吸引”了从它附近出发的所有轨线, 叫做稳定的极限环。此外有不稳定极限环, 半稳定极限环等。关于极限环所研究的问题是: 极限环的存在性、稳定性、唯一性或个数和极限环的相对位置等。

全局结构的研究: 为了研究全局性态, 需研究“无穷远”处的情况, 庞加莱建立了一个单位圆模型, 并且给出了庞加莱变换, 把整个二维平面, 包括所有无穷远点, 变换到这个单位圆上, 其圆周便是无穷远点的集合。

其他如周期系统、环面上的微分方程、高维系统等。都有相当程度的进展。

【常微分方程的稳定性理论】这是由俄国数学家李亚普诺夫在 19 世纪 90 年代开创的数学分支。它研究干扰性因素对运动系统所产生的影响。设 $X \in R^n, t \in R^1$, 把给定系统

$$\frac{dX}{dt} = X(t, X) \quad (1)$$

的解叫做“运动”, 设已知某运动 $X=f(t)$, 称它是 (在李亚普诺夫意义下) 稳定的, 指当 X_0 与 $f(t_0)$ 距离很近时, $X=X(t, X_0)$ 总离 $X=f(t)$ 很近, 其中 $X=X(t, X_0)$ 是在 t_0 时过 X_0 的解。即: $\forall \epsilon > 0$, $\exists \eta > 0$, 使当 $\|X_0 - f(t_0)\| < \eta$ 时, 总有

$$\|X(t, X_0) - f(t)\| < \epsilon$$

($t \geq t_0$) 就称 (未受干扰运动) $X=f(t)$ 是 (在李亚普诺夫意义下) 为

稳定的,否则就称为不稳定的。若稳定的运动还具有

$$\lim_{t \rightarrow +\infty} \|X(t, X_0) - f(t)\| = 0 \quad \text{就称运动 } X=f(t) \text{ 是(在李亚}$$

普诺夫意义下)渐近稳定的。

在提出概念的同时,李亚普诺夫创造了两种解决问题的方法,即第一方法和第二方法。第二方法即李亚普诺夫直接法,是把系统(1)的解的稳定与否和某些具有特殊性质的函数 $V(t, X_1, \dots, X_n)$ 的存在性联系起来,给出一些判据。李亚普诺夫本人给出几个定理,已成为运动稳定性理论中的经典定理。对自治系统

$$\frac{dX}{dt} = X(X) \quad (2)$$

设 $X=0$ 为其解。如能找到定正函数 $V(X)$,①使 $\frac{dv}{dt}|_{(2)} = \frac{\partial v}{\partial X} X$

(X) 为常负的或恒为零,则(2)的零解是稳定的;②若 $\frac{dv}{dt}|_{(2)} = \frac{\partial v}{\partial X} X$

$X(X)$ 是定负的,则(2)的零解是渐近稳定的;③若

$\frac{dv}{dt}|_{(2)} = \frac{\partial v}{\partial X} X(X) > 0$,则(2)的零解是不稳定的。

例,考虑系统

$$\begin{cases} \dot{X} = y - aX(X^2 + y^2) \\ \dot{y} = -X - ay(X^2 + y^2) \end{cases} \quad (3)$$

取定正函数 $V(X, y) = X^2 + y^2$,

$$\frac{dv}{dt}|_{(3)} = \frac{\partial v}{\partial X} \frac{dX}{dt} + \frac{\partial v}{\partial y} \frac{dy}{dt} = 2X\{y - aX(X^2 + y^2)\} + 2y\{-X - ay(X^2 + y^2)\} = -2a(X^2 + y^2)^2。$$

故得: $a > 0$ 时,(3)的零解是渐近稳定的; $a = 0$ 时,(3)的零解是稳定而非渐近稳定的; $a < 0$ 时零解是不稳定的。

李亚普诺夫之后,第二方法得到很大的发展。除了推广李亚普

诺夫的结果之外,还发展出许多新的概念,如全局稳定性、关于部分变元的稳定性、结构稳定性等,并在自动控制、星际航行等领域有重要应用。

【动力系统】考虑系统

$$\frac{dX}{dt} = X(X) \quad (1)$$

其中 $X(X) \in c(G \subseteq R^n)$, $\frac{dX_i}{dt} (i=1, 2, \dots, n)$ 表示速度分量,称 R^n 为相空间, $R \times R^n = \{(t, X)\}$ 叫做增广相空间。运动(即(1)的解)在相空间中描出的图形叫做(质点运动的)轨线。以 $f(p, t)$ 记当 $t=0$ 时过 P 点的解,其定义区间为 $(-\infty, +\infty)$, 则对每个固定的 t , $f(p, t)$ 定义了开区域 G 到自身的一个变换,当 $t \in R$ 时, $\forall P \in G$, 有 $f(p, t) \in G$, 即 $f(\cdot, t): G \rightarrow G, t \in R$, 或 $f: G \times R \rightarrow G$, 变换 f 具有下述性质: I. $f(p, 0) = p$; II. $f(p, t)$ 关于 p, t 一并连续; III. $f(f(p, t_1), t_2)$

$= f(p, t_1 + t_2)$ 。这些变换的全体叫做一个动力系统,或直接称(1)是动力系统。抽象动力系统是抛开微分方程(1),只要 f 是 $G \times R$ 到 G 的变换,并且满足条件 I、II 和 III, f 就叫做 G 上的一个动力系统,或拓扑动力系统,这是因为对固定的 t , $f(\cdot, t)$ 是从 G 到 G 自身的拓扑变换。

拓扑动力系统,主要研究各种轨线的类型及它们之间的关系。为研究轨线的分类,必须研究轨线当 $t \rightarrow \pm\infty$ 时的状况,于是有极限集,不变集,极小集等的研究。 $t \rightarrow +\infty$ 时的轨线则分为正向远离的、正向渐近的及 P^+ (泊松)稳定的,负向类似。对双侧远离,渐近和 P 式稳定也需要研究。

微分动力系统,是具有可微性质的动力系统,它研究常微系统

【流形上的微积分】

$$\lim_{h \rightarrow 0} \frac{f(X+h) - f(X)}{h} = f'(X)$$
$$\left\{ \begin{array}{cccc} \frac{\partial f_1}{\partial X_1}(X) & \frac{\partial f_1}{\partial X_2}(X) & \dots & \frac{\partial f_1}{\partial X_n}(X) \\ \frac{\partial f_2}{\partial X_1}(X) & \frac{\partial f_2}{\partial X_2}(X) & \dots & \frac{\partial f_2}{\partial X_n}(X) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m}{\partial X_1}(X) & \frac{\partial f_m}{\partial X_2}(X) & \dots & \frac{\partial f_m}{\partial X_n}(X) \end{array} \right\}$$

上式右端的矩阵叫做 f 的雅可比矩阵。从几何上来说,用过 P 点的切线来逼近的函数 $y=f(X)$, 是最佳的线性逼近;对 $z=f(X, y)$ 来说,在 P 点的切平面是对 $z=f(X, y)$ 的最佳线性逼近。再从运算关系来看,微分是满足对加法、乘法和以数去乘时的规则的,进一步

还满足微分的莱布尼茨法则,当把微分概念向流形上推广时,就以这些要求作为基础。流形上的概念,要通过与其同胚的欧氏空间相应的概念来表现,即通过同胚转移到流形上。

设 M 是某 n 维流形,即它是一个拓扑空间并且是豪斯多夫空间,设 $\{U_\alpha\}$ 是 M 的开覆盖,对任意 U_α ,联系着一个映射 $\varphi_\alpha: U_\alpha \rightarrow V_\alpha$, V_α 是 n 维欧氏空间内的非空开集, φ_α 是同胚映射。就是说, M 虽然不是欧氏空间,但可以把它局部地欧几里得空间化,即可以认为(在拓扑意义上) U_α 和 V_α 相同; M 虽然没有坐标,但可以给它“装上”局部坐标: $\forall p \in U_\alpha \xrightarrow{\varphi_\alpha} X \in V_\alpha$, 可以把 X 的坐标看成 p 的局部坐标。称 U_α 是坐标邻域, φ_α 为坐标映射, $(U_\alpha, \varphi_\alpha)$ 是坐标图, $\{(U_\alpha, \varphi_\alpha)\}$ 叫坐标图册。若 U_α, U_β 都在 $\{U_\alpha\}$ 内,并且 $U_\alpha \cap U_\beta \neq \emptyset$ ($\emptyset$ 为空集),对 $p \in U_\alpha \cap U_\beta$,在 U_α 之下, $p \rightarrow X$,在 U_β 之下, $p \rightarrow y$,即同一个点,既可用坐标 X 表示,又可用坐标 y 表示,两个坐标间有变换式 $\varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$,使 $X \mapsto y$,映射 $\varphi_\beta \circ \varphi_\alpha^{-1}$ 是 $\varphi_\alpha(U_\alpha \cap U_\beta)$ 到 $\varphi_\beta(U_\alpha \cap U_\beta)$ 的双射,若它可微并且其逆可微,就称它为微分流形。

流形上的可微函数: 设 M 是 n 维 C^∞ 流形, 坐标图册是 $\{(U_\alpha, \varphi_\alpha)\}$, G 是 M 中的开集, $p \in G$. 若 $f \circ \varphi_\alpha^{-1}$ 可微, 就称 f 在点 p 可微, 当 f 在每个点 $p \in G$ 可微时, 就称 f 在 G 可微。

切空间: 设 $p \in M$, F_p 是在点 p 附近有定义且在 p 点可微的函数组成的空间, 具有: $(f+g)(p') = f(p') + g(p')$; $(\alpha f)(p') = \alpha f(p')$; $(fg)(p') = f(p')g(p')$ 。 $\forall \varphi \in F_p, \varphi: F_p \rightarrow \mathbb{R}$ 具有: (I) 线性: $\varphi(f+g) = \varphi(f) + \varphi(g)$, $\varphi(\alpha f) = \alpha \varphi(f)$; (II) 莱布尼茨法则: $\varphi(fg) = f(p)\varphi(g) + g(p)\varphi(f)$, 就称 φ 是 M 在点 p 的一个切向量, 所有这种 φ 组成的空间称为 M 在点 P 的切空间, 记为 T_p , 这是切线, 切平面的推广。切空间 T_p 是 n 维线性空间, 在给定的坐标系下, 它的基是 $\frac{\partial}{\partial X_1} \Big|_x, \dots, \frac{\partial}{\partial X_n} \Big|_x$, 其中 $X = (X_1, \dots, X_n)$ 是点 p 的局部坐标, 任一切向量均可表为

$$\varphi = \alpha_1 \frac{\partial}{\partial X_1} \Big|_x + \dots + \alpha_n \frac{\partial}{\partial X_n} \Big|_x.$$

积分: 在 $\mathbb{R}^3$, 三个向量 $\vec{a}_i = a_{i1}\vec{i} + a_{i2}\vec{j} + a_{i3}\vec{k}$ 的外积 $a_1 \wedge a_2 \wedge a_3 = \det(a_{ij})$, 其中 $(a_{ij}) = A(i, j=1, 2, 3)$ 为三阶方阵, 它具有: 重线性、反交换性及 $a \wedge a = 0$, 设 V 为 n 维实线性空间, V 中的元素以 u, v 来表示, V 的基以 $e_1, e_2, \dots, e_n$ 表示。构造新的线性空间 Λ^k 如下: $\Lambda^0 = \mathbb{R}, \Lambda^1 = V, \Lambda^2 = \left\{ \sum a_i(u_i \wedge v_i) \right\}$, 其中的和式均取有限项。 $a_i \in \mathbb{R}, u_i, v_i \in V$, 运算 $\wedge$ 满足: (I) 重线性: $(\alpha_1 u_1 + \alpha_2 u_2) \wedge v = \alpha_1(u_1 \wedge v) + \alpha_2(u_2 \wedge v)$; $u \wedge (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1(u \wedge v_1) + \alpha_2(u \wedge v_2)$; (II) 反交换性: $u \wedge v = -v \wedge u$; (III) $u \wedge u = 0$ 。由 $u = \sum_{i=1}^n \alpha_i e_i$ 和 $v =$

$\sum_{i=1}^n \beta_i e_i$, 有 $u \wedge v = \sum_{i,j} \alpha_i \beta_j (e_i \wedge e_j) = \sum_{i < j} (\alpha_i \beta_j - \alpha_j \beta_i) e_i \wedge e_j$ (其中用到了

$e_i \wedge e_i = 0$ 及 $e_j \wedge e_i = -e_i \wedge e_j$)。一般, $\Lambda^k = \left\{ \sum \alpha_{i_1 \dots i_k} (v_{i_1} \wedge v_{i_2} \wedge \dots \wedge v_{i_k}) \right\}$, 其中的和式为有限的, $\alpha_{i_1 \dots i_k} \in R, v_{i_1}, \dots, v_{i_k} \in V, \Lambda$ 满足: (I) 重线性: $v_{i_1} \wedge \dots \wedge (\alpha v_{i_j} + \alpha' v'_{ij}) \wedge \dots \wedge v_{i_k} = \alpha v_{i_1} \wedge \dots \wedge v_{i_j} \wedge \dots \wedge v_{i_k} + \alpha' v_{i_1} \wedge \dots \wedge v'_{ij} \wedge \dots \wedge v_{i_k}$; (II) 反交换性: $v_{i_1} \wedge \dots \wedge v_{i_j} \wedge \dots \wedge v_{i_e} \wedge \dots \wedge v_{i_k} = -v_{i_1} \wedge \dots \wedge v_{i_e} \wedge \dots \wedge v_{i_j} \wedge \dots \wedge v_{i_k}$; (III) 若 $v_{i_1}, \dots, v_{i_k}$ 中有两个元素相同, 则 $v_{i_1} \wedge \dots \wedge v_{i_k} = 0$, 在基 $e_1, e_2, \dots, e_n$ 之下, Λ^k 的元素的一般形式是

$\sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}$ 。这种形式叫做 k 形式, 其中 $a_{i_1 \dots i_k} \in R$,

记 $\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \dots \oplus \Lambda^n$, 称 Λ 是由 V 生成的格拉斯曼代数。

设 X 是 n 维线性空间, X^* 为 X 上所有实线性泛函组成的空间, 称它为 X 的对偶空间。 T_p 是流形 M 在点 $P \in M$ 处的切空间, T_p 上所有实线性泛函组成的空间 T_p^* 是 T_p 的对偶空间。当 P 的局部坐标是 $(X_1, X_2, \dots, X_n)$ 时, T_p^* 的元素一般形式是 $a_1 dX_1 + \dots + a_n dX_n$, 其中 $a_1, \dots, a_n$ 是在 P 的某邻域内的可微函数。当某函数可微时, $df = \frac{\partial f}{\partial X_1}(P) dX_1 + \dots + \frac{\partial f}{\partial X_n}(P) dX_n$ 。 T_p^* 生成一个格拉斯曼代数 Λ , 其中的 k 形式是 $\sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} dX_{i_1} \wedge \dots \wedge dX_{i_k} (1 \leq k \leq n)$ 称它为 k 微分形式。

外微分, 在 Λ 上引进映射 d , 它把每个 k 形式映为 $(k+1)$ 形式: $\omega = a(X_1, \dots, X_n) dX_{i_1} \wedge \dots \wedge dX_{i_k}, d\omega = da \wedge dX_{i_1} \wedge \dots \wedge dX_{i_k} = \left(\frac{\partial a}{\partial X_1} dX_1 + \dots + \frac{\partial a}{\partial X_n} dX_n \right) \wedge dX_{i_1} \wedge \dots \wedge dX_{i_k}$ 。称这种 d 为微分形式上的外微分。

设 M 为 n 维流形, 其坐标图册由坐标图 (M, φ) 构成, 又设 ω 是 M 上的一个 n 形式, 在坐标 $(X_1, X_2, \dots, X_n)$ 之下, ω 表为 $\omega = \omega$

$(X_1, \dots, X_n)dX_1 \wedge \dots \wedge dX_n$, ω 在 M 上的积分定义为 $\int_M \omega = \int_{\varphi(M)} \omega$
 $(X_1, \dots, X_n)dX_1 \cdots dX_n$ 。等式右端是欧氏空间上的重积分, 例如(第
 二型)曲面积分 $\iint_S f(X, y, z)dXdy$, $dXdy$ 理解成 $dX \wedge dy$, 当 $X = X$

(u, v) , $y = y(u, v)$, $z = z(u, v)$, $(u, v) \in D$ 时, 因为 $dX = X_u du + X_v dv$, $dy = y_u du + y_v dv$, 故 $dX \wedge dy = (X_u du + X_v dv) \wedge (y_u du + y_v dv) = \begin{vmatrix} X_u & X_v \\ y_u & y_v \end{vmatrix} du \wedge dv$, 于是

$$\iint_S f(X, y, z)dX \wedge dy = \iint_D f(X(u, v),$$

$$y(u, v), z(u, v)) \left(\pm \begin{vmatrix} X_u & X_v \\ y_u & y_v \end{vmatrix} du \wedge dv \right)$$

这就相当于化曲面积为二重积分。

问题在于, s 的坐标图册不是“整体”的, 或“完整的一块”的, 而是由局部坐标“粘连”起来的, 为将其整体化, 需要借助于“1 的分解”, 结果是: 若 M 是 n 维定向的紧微分流形, 其坐标图册是: $(U_1, \varphi_1), (U_2, \varphi_2), \dots, (U_n, \varphi_m)$, 又 Π_i 是属于坐标邻域 $\{U_i\} (i=1, \dots, n)$ 的 1 的 C^∞ 分解, 若 ω 为 n 形式, 在 U_i 内表为

$$\omega = \omega(X_1, \dots, X_n)dX_1 \wedge dX_2 \wedge \dots \wedge dX_n$$

定义 ω 在 M 上的积分 $\int_M \omega$ 为 $\int_M \omega \Pi_i$ 的和, 即有

$$\begin{aligned} \int_M \omega &= \sum_{i=1}^m \int_{U_i} \omega \Pi_i \\ &= \sum_{i=1}^m \int_{\varphi_i(U_i)} a_i(X_1, \dots, X_n) \cdot \end{aligned}$$

$$\Pi_i(X_1, \dots, X_n)dX_1 \cdots dX_n。$$

右端是 n 维欧氏空间上的 n 重积分。

例,分别令 $\omega = PdX + Qdy$, $\omega = Pdy \wedge dz + Qdz \wedge dX + RdX \wedge dy$ 及 $\omega = PdX + Qdy + Rdz$, 则格林公式、高斯公式和斯托克斯公式均呈下面统一的形式: $\int_D d\omega = \int_{\partial D} \omega$

当然其优点绝不仅仅是将以前的结果“统一”,而是有许多更深刻的论题。它着眼于拓扑结构、微分结构及复结构等。

【非标准分析】这是美国逻辑学家鲁宾逊于 1960 年创立的新学科。他首先把实数结构 R 扩张为包含无穷小数和无穷大数的结构 R^* , 在一定意义下 R^* 与 R 有相同的性质。称 R^* 中的元素为超实数。形象地说,是在普通实数中又加进了无穷小数(其绝对值小于任何实数)及无穷大数(其绝对值大于任何实数)。若 $\alpha = a + \varepsilon$, 其中 ε 为无穷小, a 为“标准”实数(即 $a \in R$)。当两个超实数 α 与 β 相差为无穷小时,就称 α 为无限接近于 β 的,记为 $\alpha \simeq \beta$, 可以证明这是一个等价关系。每个等价类中包含唯一的标准实数 a 。称等价类 $\mu(a)$ 为一个单子。单子不是 R^* 中的数,而相当于 R 中的数。可进行四则运算,但不满足阿基米德公理(即 $\forall a, b > 0, \exists n \in N$, 使得 $na > b$)。在这些规定之下,标准分析里的许多概念、定理等可以自然地扩张。如区间 $[a, b]$ 扩张为 $[a, b]^*$; R 中的函数可以自然地扩张为 $f^*(X)$ 等;连续性:函数 $f(X)$ 在标准点 X_0 连续 $\Leftrightarrow X \simeq X_0$ 时, $f^*(X) \simeq f^*(X_0)$ 。有一个特别重要的定理——转换定理:每个关于 R 可形式化的命题,对 R 成立时,经过适当解释(即将 R 中的对象,解释为 R^* 中的对象)对 R^* 也成立,反之亦然。

非标准分析使“无穷小”获得新生,重新用它刻画微积分问题。这不仅能表明状态,还能表明过程,使具有直观与简洁性。非标准分析一经问世便得到迅速发展,用它解决了许多问题。如希尔伯特

空间上的多项式紧算子的不变子空间的存在问题(在此之前,已有几十年未能解决)等。卢森堡构造出了非标准分析模型,也有其他人构造出非标准分析模型。此外,发展出非标准群论、非标准泛函分析、……学科,而且有不少应用。

偏微分方程

【偏微分方程】含有未知函数的偏导数的方程。如方程

$$u_t - a^2 u_{xx} = 0 \quad (1)$$

其中 a 为已知常数, $u = u(X, t)$ 是未知函数。方程中出现的未知函数的最高阶偏导数的阶数,称为方程的阶数。如方程(1)是二阶方程。

18 世纪初,微积分形成不久,人们就开始用偏微分方程研究物理学问题。如用方程(1)描述了热传导规律,其中 $u(X, t)$ 表示在 X 点 t 时刻的温度。随着科学技术的发展,偏微分方程已在多种学科中广泛应用。除物理学、技术科学外,在化学、生物学、生态学等多种领域中也有所应用。人们一方面研究偏微分方程的一般理论,一般方程的求解、解的性质等,但即使形式很简单的方程也常常很复杂。因此人们更多地研究其他学科中感兴趣的偏微分方程,针对各种问题,给出各种解决办法。

一般,一个偏微分方程有许多解,可含有任意函数。如二阶方程 $u_{tt} - a^2 u_{xx} = 0$, 有解 $u(X, t) = f(X - at) + g(X + at)$ 其中 $f(\zeta)$ 、 $g(\eta)$ 是二次连续可微的任意函数。

对于偏微分方程,通常注重求满足某些附加条件的特解。未知

函数在初始时刻所满足的条件称为初始条件,在所给区域的边界上满足的条件称为边界条件。这些条件统称为定解条件,它们通常依据实际问题的需要而提出。方程和定解条件组成定解问题。只含有初始条件的定解问题称为初始值问题或柯西问题;只含有边界条件的定解问题称为边值问题;既有边界条件又有初始条件的称为初边值问题或混合问题。

若定解问题的解是存在的、唯一的、且对定解条件是连续依赖的(即定解条件中的量变化不大时,解的改变也不大),称定解问题是适定的定解问题。在 20 世纪以前,人们认为适定的定解问题才有意义,但随着科学技术的发展,特别是在探矿、大气物理等一些很有意义的实际问题中,遇到了不适定的问题,如连续依赖性不成立。不适定的定解问题也称为病态问题,20 世纪中叶以来,不适定问题的研究已逐渐成为偏微分方程的一个重要方面。

若函数 u 在某区域 D 上连续,方程中出现的各偏导数也连续,代入方程使方程成为恒等式,就称 u 是方程的经典解。经典解对解函数光滑性的严格要求——偏导数也连续,常使一些很有意义的定解问题没有解。比如集中力作用的弦振动问题。因此需要扩充解的概念。扩充的方法一般有两类:其一是把经典解在一定意义下的极限算作解;另一方法是把广义导数意义下满足方程或定解问题的函数作为解。它们统称为广义解。相对于广义解,经典解也称为狭义解。

【数学物理方程】主要是指从物理学及其他各门自然科学、技术科学中出现的偏微分方程,有时也包括与此有关的积分方程、微分积分方程和常微分方程。

早在 18 世纪,达朗贝尔把弦的振动规律归结为波动方程 u_{tt}

$=a^2u_{xx}$, 付里叶用热传导方程 $u_t = a^2(u_{xx} + u_{yy} + u_{zz})$ 描述固体中的热传导规律, 而拉普拉斯方程 $u_{xx} + u_{yy} + u_{zz} = 0$ 描述了定常的温度场、速度势和引力势等。同一个偏微分方程常常可描述许多不同领域中的问题。

经典的数理方程大量的是二阶线性偏微分方程。所谓线性, 是指方程中出现的未知函数及其导数都是一次的。波动方程、热传导方程和拉普拉斯方程都是二阶线性方程。若线性方程中, 有不含未知函数及其导数的项, 如 $u_{tt} = a^2u_{xx} + f(X, t)$ 中的 $f(X, t)$, $f(X, t)$ 为已知函数, 这项称为非齐次项或自由项, 此时方程称为非齐次线性方程, 否则称齐次线性方程。齐次线性方程的解符合叠加原理, 即若 u_1, u_2 是方程的解, 则 $\alpha u_1 + \beta u_2$ 也是方程的解, 其中 α, β 为任意常数。分离变量法、齐次化原理等是解数学物理方程定解问题的重要方法, 都依据叠加原理。

【波动方程】 $u_{tt} - a^2u_{xx} = 0$ 是一维波动方程。它可描述弦的微小横振动, 其中 $M(X, t)$ 表示弦的横向位移, 故又称弦振动方程。它的解可表示成 $u = f(X - at) + G(X + at)$ ($F(\zeta), G(\eta)$ 为任意的二次连续可微函数), $F(X - at), G(X + at)$ 分别称为右传播波和左传播波。直线族 $X - at = c_1$ 和 $X + at = c_2$ 称方程的特征线族, 右传播波沿 $X - at = c$ 值不变左传播波沿 $X + at = c_2$ 不变。

给定弦的初始位移和速度 $u(X, 0) = \varphi(X), u_t(X, 0) = \psi(X)$, 它们和方程一起组成初始值问题, 其解表示为 $u(X, t) = \frac{1}{2}[\varphi(X - at) + \varphi(X + at)] + \frac{1}{2} \int_{X-at}^{X+at} \psi(\alpha) d\alpha$ 称为达朗贝尔公式。

对于有限长的弦的振动问题, 如长为 l 两端固定的弦的振动问题, 除初始条件外, 还要加端点条件, 即边界条件 $u(0, t) = 0$ 和

$u(l, t) = 0$, 成为初边值问题。利用分离变量法可得到它的级数形式的解。分离变量法的物理意义是把弦的振动分解成基音和泛音的叠加。分离变量法的基本步骤是: 设方程有形为 $u(X, t) = X(X)T(t)$ 的解, 将它代入方程, 可以分离变量而得到两个常微分方程 $X'' + \lambda X = 0, T'' + \lambda T = 0, \lambda$ 是参数。由边界条件得到 $X(X)$ 的边界值, 这就构成了 $X(X)$ 的特征值问题, 解之, 得到一系列特征函数 $X_n(X)$, 再解出相应的 $T_n(t)$, 从而得到方程满足边界条件的一系列的解 $u_n = X_n T_n$, 把它们叠加起来, 确定系数使满足初始条件, 便得到原问题的解。初边值问题的边界条件也可能是其他形式的。如在 l 端自由时, 边界条件形式为 $\frac{\partial u}{\partial X}(l, t) = 0$, 分离变量法还可用到其他方程和别的定解问题。

对非齐次方程, 如 $u_{tt} - a^2 u_{xx} = f(X, t), f(X, t)$ 的物理意义是持续外力作用, 把它看成各时刻所受冲击力的叠加, 利用相应齐次方程问题的解, 叠加起来——积分来得到非齐次方程的特解, 这种方法称为齐次化原理或 *Duhamel* 原理。齐次化原理也适用于热传导方程和一些更为复杂的线性方程。

二维、三维的波动方程 $u_{tt} = a^2(u_{xx} + u_{yy}), u_{tt} = a^2(u_{xx} + u_{yy} + u_{zz})$, 分别描述膜振动和声波、电磁波的传播。在实际中, 二维波与三维波的传播有很大的区别, 三维波有惠更斯原理, 有局部振源的声波, 波在传播过程中, 传到和传过都有明显的界面, 分别称为前阵面和后阵面, 无后效性。而由初始的局部扰动形成的二维波, 如平静的水面投一石块, 水面波有明显的前阵面, 而没有后阵面, 它具有持续后效的特性, 只是当 $t \rightarrow \infty$ 时, 振动减小到零。二维、三维波动方程的柯西问题解的泊松公式分别刻划了这一特性。

【拉普拉斯方程】又称调和方程。二维、三维及 n 维的拉普拉斯方程 $u_{xx} + u_{yy} = 0$, $u_{xx} + u_{yy} + u_{zz} = 0$ 及 $\sum_{i=1}^n \cdots u_{x_i x_i} = 0$, 均可用 $\Delta u = 0$ 或 $\nabla^2 u = 0$ 表示。它们相应的非齐次方程是泊松方程 $\Delta u = f$ 。 f 常表示成 $-4\pi\rho$, 其中 ρ 表示密度。

拉普拉斯方程通常描述稳定状态, 一般研究它的边值问题。基本的边界条件有三类, 在区域的边界上: $u = \varphi$ 称为第一类; $\frac{\partial u}{\partial n} = g$ 称为第二类; $\frac{\partial u}{\partial n} + \alpha u = g$ ($\alpha \geq 0$) 称为第三类。其中 $\frac{\partial u}{\partial n}$ 表示外法线方向导数, φ, g, α 是定义在区域边界上的函数。附加第一类边界条件问题称为第一类边值问题, 附加第二类边界条件的称为第二类边值问题。

拉普拉斯方程的二次连续可微解称为调和函数。调和函数是非常光滑的, 区域 Ω 上的调和函数不但有任意阶偏导数, 而且是实解析的函数, 即在 Ω 中任一点, 都存在它的一邻域, 在这邻域上调和函数可表示成一致收敛的幂级数。复变量 z 的解析函数 $f(z) = u + iv$ 它的实部 u 与虚部 v 都是二维的调和函数。

调和函数具有极值性质, 即在区域 Ω 上不恒为常数的调和函数, 在区域 Ω 的内部不能取得最大值和最小值。

$\frac{1}{2\pi} \ln \frac{1}{r}$ 与 $\frac{1}{4\pi r}$ 分别是二维和三维拉普拉斯方程的基本解, $r=0$

是它们的奇点。由它们可以构造出方程的其他解。以三维情形来说, 积分

$$\iint_s \frac{\mu(\zeta, \eta, \zeta)}{r} d\sigma, \quad (\text{式中 } s \text{ 为一曲面, } \mu \text{ 是定义在 } s \text{ 上的连续函数})$$
 确定的函数在 s 以外满足方程。在三维区域 Ω 上的积分 $u(X, y, z) =$

$\iiint_{\Omega} \frac{p(\zeta, \eta, \zeta)}{r} d\zeta d\eta d\zeta$ 是泊松方程的重要特解, 它是以 ρ 为密度的体位势。利用格林公式还可以得到调和函数的积分表示式 $u(X, y, z) =$

$$\frac{1}{4\pi} \iint_{\partial\Omega} \left[\frac{1}{r} \frac{\partial u}{\partial n} - u \frac{\partial}{\partial n} \left(\frac{1}{r} \right) \right] d\sigma, \text{ 其中 } \partial\Omega \text{ 是区域 } \Omega \text{ 的边界。}$$

拉普拉斯方程的狄里希勒问题的格林函数是 $(\zeta, \eta, \zeta, X, y, z) = v(\zeta, \eta, \zeta, X, y, z) + \frac{1}{4\pi r}$, 其中 v 是调和函数, 满足 $v|_{\partial\Omega} = \frac{1}{4\pi r}$ 。由格林函数可表示出狄里希勒问题的解 $u(X, y, z) = - \iint_{\partial\Omega} Q(\zeta, \eta, \zeta) \frac{\partial}{\partial n} G(\zeta, \eta, \zeta, X, y, z) d\sigma$ 。对于一些特殊区域, 如球、半空间、半球等用镜象法可求出格林函数, 从而得到狄里希勒问题的解。这种解狄里希勒问题的方法称格林函数方法。

对于某些特殊区域的边值问题, 也可用分离变量法求解。

【热传导方法】最早由描述热传导过程得到。它的一维形式是 $u_t - a^2 u_{xx} = f(X, t)$, u 是温度函数, $a^2 = \frac{k}{c\rho}$, k 是热传导系数, c 是比热, ρ 是密度。 F 是外热原密度 $f = \frac{F}{c\rho}$ 。三维介质中的热传导方程是 $u_t - a^2(u_{xx} + u_{yy} + u_{zz}) = f$ 。热传导方程也可描述其他的物理现象, 如扩散过程等。

方程附加初始条件 $u|_{t=0} = \varphi$ 成为初始值问题。一维齐次方程的初始值问题的解是 $u(X, t) = \frac{1}{2a\sqrt{\pi t}}$

$\int_{-\infty}^{\infty} \varphi(\xi) e^{-\frac{(X-\xi)^2}{4a^2 t}} d\xi$ 。其中函数 $\frac{1}{2a\sqrt{\pi t}} e^{-\frac{(X-\xi)^2}{4a^2 t}}$ 是初始值问题的基本

解,由它还可组成非齐次方程的特解。

对于有界区域上的热传导问题,为了确定解,除初始条件外;还要附加边界条件。边界条件通常有三类,设 $\partial\Omega$ 为区域的边界,第一类边界条件: $u|_{\partial\Omega} = \varphi$, 即已知界面上温度;第二类边界条件: $\frac{\partial u}{\partial n}|_{\partial\Omega} = g$, n 为外法线方向,其意义是通过边界的热量已知;第三类边界条件: $\left(\frac{\partial u}{\partial n} + \alpha u\right)|_{\partial\Omega} = g$ ($\alpha \geq 0$), 即物体表面给定热交换规律。对于一维情形,如杆的热传导,边界条件给在杆的两端,其中 $\frac{\partial u}{\partial n}$ 为 $\frac{\partial u}{\partial X}|_{X=0}$ 和 $\frac{\partial u}{\partial X}|_{X=1}$ 。

方程连同初始条件及边界条件中的任意一类,组成初边值问题。和波动方程的初边值问题类似,可以用分离变量法得到级数形式的解。

一个内部没有热源的热传导过程(即方程中 $f=0$),内部的温度不能超过和低于初始时刻及边界上的最高温度和最低温度,这种性质称为极值原理。利用极值原理直接可得初边值问题解的唯一性和稳定性。

【二阶线性偏微分方程】二个自变量的二阶线性方程的一般形式是 $a_{11}u_{XX} + 2a_{12}u_{XY} + a_{22}u_{YY} + b_1u_X + b_2u_Y + cu = f(1)$, 其中 $a_{ij}, b_i, c, f(ij=1,2)$ 是 X, y 的已知函数。许多经典的物理现象是由二阶线性方程描述的。如波动方程、热传导方程和拉普拉斯方程,它们是偏微分方程中研究最早、最有广泛应用的方程。方程中的二阶导数部分称为方程的主部, $a_{12}^2 - a_{11}a_{22}$ 为判别式。在 X, y 平面区域 Ω 上,以方程的判别式为正、为负或为零,分别称方程为椭圆型偏微分方程、双曲型偏微分方程和抛物型偏微分方程。简称为椭圆型方程、双曲

型方程和抛物型方程。各类方程都可以通过适当的可逆的变量变换,使其主部呈典型形式: $u_{xx} + u_{yy}$ (椭圆型); $u_{xx} - u_{yy}$ 或 u_{xy} (双曲型); u_{xx} (抛物型)。拉普拉斯方程、波动方程和热传导方程分别是椭圆型方程、双曲型方程和抛物型方程的典型代表。同一类方程常有许多相近的性质,不同类型的方程所描述的物理现象则有很大区别,从而解的性质和定解问题的提法及解法也常有很大区别。

对于多于二个自变量的方程,可以它的主部的系数组成的二次型的特性作类似的分类,但一般没有可逆的变量变换使方程成为典型形式。

在区域 Ω 内,有不同型的部分,如在空气动力学中迂到的特里柯米方程 $y u_{xx} + u_{yy} = 0$,在 $y = 0$ 附近。称这种方程为混合型方程。对混合型方程的研究 20 世纪才开始。

若在区域 Ω 内,仅有某些点集或边界上发生型的变化,其余部分均属一种类型,就称方程是退缩的,如退缩椭圆型、退缩双曲型和退缩抛物型。

【一阶偏微分方程】以两个自变量为例,它的一般形式是 $F(X, y, u, u_x, u_y) = 0$,其中一阶拟线性方程为 $a(X, y, u)u_x + b(X, y, u)u_y = c(X, y, u)$ (1). 把方程的解 $u = u(X, y)$ 看成 (X, y, u) ,空间的曲面,称为积分曲面。其法线方向数是 $(u_x, u_y, -1)$ 。当 u 是(1)的解时,它的几何意义是,积分曲面的法线与方程系数组成的向量 (a, b, c) 垂直,即积分曲面必与向量 (a, b, c) 相切。以 (a, b, c) 为方向数的方向标组成方向场,称为方程的特征方向场,相应的常微分方程组 $\frac{dX}{a} = \frac{dy}{b} = \frac{du}{c}$ 是特征方程组,特征方程组的积分曲线叫特征线,光滑曲面 $u = u(X, y)$ 成为(1)的解的充要条件是它由特征曲线组

成。

光滑曲线 $r: X=X(t), y=y(t)$ 上给定 u 的值 $u|_{\gamma}=\varphi(t)$, 称为初始值, 求方程 (1) 满足初始值的解称为解初始值问题或柯西问题。当 r 与特征线在 (X, y) 平面上投影不相切时, 过 $\Gamma: X=X(t), y=y(t), u=u(t)$ 每点作特征线, 这些特征线织成曲面 $u=u(X, y)$, 当 Γ 是光滑曲线时, $u(X, y)$ 就是光滑曲面且成为方程 (1) 的初始值问题的解。由此看出一阶方程的求解问题与常微分方程组有密切关系。

柯西-柯娃列夫斯卡娅定理, 证明了解析系数方程的解析初始值问题的解析解的存在、唯一性。这个定理条件要求苛刻, 得到的结果又是局部的, 但它在偏微方程的理论中仍是很重要的。

柯西-黎曼方程组 $\frac{\partial u}{\partial X} = \frac{\partial v}{\partial y}, \frac{\partial v}{\partial X} = -\frac{\partial u}{\partial y}$ 是最重要的一阶线性方程组之一, 它在流体力学中起着重要作用, 它的解 u 和 v 组成一对共轭调和函数, 即它们是某个解析函数 $f(z)$ 的实部和虚部。

积分方程

【积分方程】它是包含有积分运算且在积分号下含有未知函数的方程。描述地球引力场中质点下落轨迹的阿贝尔方程就是最早的积分方程。在积分方程的研究中线性积分方程(即未知函数在方程中以线性形式出现)的理论较为丰富, 其中最为著名的是以瑞典数学家弗雷德霍姆命名的第一、二、三型方程, 它们分别是 (1)

$$\int_a^b k(X, y)\varphi(y)dy = \psi(X); (2) \psi(X) - \lambda \int_a^b k(X, y)\varphi(y)dy = \psi(X);$$

(3) $A(X)\varphi(X) - \lambda \int_a^b k(X, y)\psi(y)dy = \psi(X)$, 其中 $k(X, y)$ 为在 $a \leq X, y \leq b$ 上连续的已知函数, 称为积分方程的核, $A(X)$ 与 $\psi(X)$ 都是在区间 $a \leq X \leq b$ 上连续的已知函数, $\varphi(X)$ 是未知函数, λ 是参数。弗雷德霍姆为其第二型的方程建立了求解的弗雷德霍姆方法及讨论可解性的一组重要定理: (1) 在 λ 平面的有限区域内, 最多有有限多个 λ 值, 使齐次方程 $\varphi(X) - \lambda \int_a^b k(X, y)\varphi(y)dy = 0$ 有非零解, 这样的 λ 称为特征值; 对应的非零解称为特征函数; (2) 每个特征值 λ 至多对应有限多个特征函数, 其个数称为 λ 的秩; (3) 当 λ 为特征值时, 非齐次方程 (即 $\psi(X) \neq 0$ 的情形) 可解的充分必要条件为其自由项 $\psi(X)$ 与共轭齐次方程 $\Phi(X) - \bar{\lambda} \int_a^b \overline{k(y, X)}\Phi(y)dy = 0$ 的特征函数 $\Phi_i(X)$ 正交, 即 $\int_a^b \psi(X)\Phi_i(X)dX = 0 (i=1, 2, \dots, r)$, r 是 λ 的秩。 (4) 齐次方程与其共轭齐次方程特征值互为共轭复数, 且对应的秩相等。这四个定理常被简述为弗雷德霍姆备择定理: 第二型弗雷德霍姆方程或者对任意自由项有唯一解, 或者相应的齐次方程有非零解。另一类重要的线性方程称为沃尔泰拉方程, 它是当 $k(X, y) = 0 (X < y)$ 的弗雷德霍姆方程, 这时方程的上限 b 可改为变量 X , 沃尔泰拉方程也相应地具有三种不同形式, 它的理论和弗雷德霍姆方程有了很大差异。如第二型沃尔泰拉方程对一切的 λ 总可解; 第一型沃尔泰拉方程在一定条件下总可以化为第二型沃尔泰拉方程, 而弗雷德霍姆第一型的方程则不具有这些性质, 而其可解性的研究则相当困难。以上两类方程都可以推广到关于未知函数向量 $\varphi(X) = (\varphi_1(X), \varphi_2(X), \dots, \varphi_n(X))$ 的方程组的情形, 这时核 $k(X, y)$ 看作是 n 阶方阵 $(k_{ij}(X, y)) (i, j=1, 2, \dots, n)$, 自由项也是 n 维向量 $\psi(X) = (\psi_1(X), \psi_2(X), \dots, \psi_n(X))$ 。当 $k(X,$

y)不是连续核,而具有某种奇异性时,称相应的方程为奇异积分方程,例如 $k(X,y)=\frac{1}{X-y}$,称其为柯西核, $\int_a^b \frac{dy}{X-y}$ 当 $y \in (a,b)$ 时,在柯西主值意义下存在。希尔伯特在研究解析函数边值问题时及庞加莱在研究潮汐现象时都发现了这类积分方程,其一般形式为 $A(t)\varphi(t) + \frac{1}{\pi i} \int_L \frac{\varphi(\tau)}{\tau-t} dZ + \int k(t,\tau)\varphi(\tau)d\tau = \psi(t)$,前两项称为方程的特征部分,第三项 $k(t,\tau)$ 是连续核。这类方程称其为具柯西核的奇异积分方程,对于这类方程的可解性有诺特定理。关于非线性积分方程直到目前还没有系统的理论,但有许多数学家在从事有关课题的研究。在近代数学中积分方程被看成抽象空间中的算子方程,如第二型弗雷德霍姆方程可写成 $\varphi - \lambda T\varphi = \psi$, φ, ψ 都是巴拿赫空间中的元素, T 是巴拿赫空间中的全连续算子, $T\varphi \equiv \int_a^b k(t,\tau)\varphi(\tau)d\tau$,特征值 λ 称为算子 T 的点谱,特征函数称为特征元素。

泛 函 分 析

【泛函分析】不是孤立地研究某个函数或若干个函数之间的关系和方程,而是把许多函数做为总体来研究。即研究函数空间和函数变换的数学分支。由于它把具体问题抽象成一种纯粹代数的与拓扑的结构形式以进行研究,因此形成了种种综合运用代数与几何的手段去处理分析问题的新方法。泛函分析研究各种具有拓扑的线性空间,性质及解决问题的方法。

【拓扑向量空间】(又称拓扑线性空间)设 X 是向量空间,又是拓扑空间,如果其代数运算按其拓扑连续,就说 X 是一个拓扑向量空间。即 X 要具有以下性质:(1)对于每一对元素 $x, y \in X$, 及其和 $x+y$ 的任何邻域 V_{x+y} , 总能找到 x 的邻域 V_x, y 的邻域 V_y , 使得 $V_x + V_y \subset V_{x+y}$ 。这里 $V_x + V_y = \{u+v \mid u \in V_x, v \in V_y\}$; (2)对于每个元素 $x \in X$ 及数值 λ , 元素 λx 的任何邻域 $V_{\lambda x}$, 总能找到 x 的邻域 V_x 及数 $\delta > 0$, 使得当 $|\mu - \lambda| < \delta$ 时, $\mu V_x \subset V_{\lambda x}$ 。

【拓扑向量量子空间】设 X 为拓扑向量空间。将 X 中的拓扑及代数运算诱导到 X 的线性子集 X_1 上, 则 X_1 也是一个拓扑向量空间, 称 X_1 为 X 的拓扑向量量子空间。

【赋范向量空间】又称赋范线性空间, X 为数域 K 上的向量空间。 $\|\cdot\|$ 为 X 到 R 的映射, 满足条件。(1) $\|x\| \geq 0 \forall x \in X; \|x\| = 0 \Leftrightarrow x = 0$; (2) $\|x+y\| \leq \|x\| + \|y\|$; (3) $\alpha \in K, x \in X \Rightarrow \|\alpha x\| = |\alpha| \|x\|$ 称 $\|\cdot\|$ 是 X 上的范数。 X 关于 $\|\cdot\|$ 为赋范向量空间简记作 $(X, \|\cdot\|)$ 。完备的赋范向量空间叫 *Banach* 空间, 记作 B 空间。

【赋准范向量空间】在赋范向量空间条件中, 如果条件(1), (2)成立, 而将(3)代以(3)₁: $\|-x\| = \|x\|$, (3)₂. $\{X_n\} \subset X, \alpha \in K$ 则 $\|X_n\| \rightarrow 0 \Rightarrow \|\alpha X_n\| \rightarrow 0$, (3)₃, $\alpha_n \in K, x \in K$ 则 $\alpha_n \rightarrow 0 \Rightarrow \|\alpha_n x\| \rightarrow 0$ 。则称 $\|\cdot\|$ 为 X 上的准范数。称 X 为赋准范向量空间。简称为赋准范空间。范数必为准范数; 赋范空间亦为赋准范空间。完备的赋准范空间叫做 *Frechet* 空间(F 型空间)。

【距离向量空间(度量线性空间)】若向量空间 X 上有距离函数 ρ (即(1) $\rho(X, X)=0, \rho(X, y)>0 \forall X \neq y$, (2)

$\rho(X, y)=\rho(y, X)$, (3) $\rho(X, z) \leq \rho(X, y)+\rho(y, X)$), 如果 X 上的运算关于 ρ 连续, 即(1) $\rho(X_n, X) \rightarrow 0, \rho(y_n, y) \rightarrow 0 \Rightarrow \rho(X_n + y_n, X + y) \rightarrow 0$; (2) $X_n \rightarrow X, \alpha \in K \Rightarrow \alpha X_n \rightarrow \alpha X$; (3) $\alpha_n \rightarrow \alpha, X \in X \Rightarrow \alpha_n X \rightarrow \alpha X$ 。

则称 X 为距离向量空间。记作 (X, ρ) 对 $(X, \| \cdot \|)$, 令 $\rho(X, y) = \| X - y \|$, 则 $(X, \| \cdot \|)$ 构成 (X, ρ) 。

【不变距离向量空间】具有平移不变性: $\rho(X + a, y + a) = \rho(X, y) \forall X, y, a \in X$ 的距离空间 (X, ρ) 赋范向量空间 $(X, \| \cdot \|)$, 以 $\rho(X, y) = \| X - y \|$ 为距离, 就是不变距离向量空间, 反之亦然。

【希尔伯特空间】与 n 维欧氏空间最相类似, 且极为有用的一种空间。在向量空间 X 上定义了二元运算内积: (X, y) , 它满足(1) $(X, y) = (y, X), \forall X, y \in X$; (2) $(X, X) \geq 0 \forall X \in X; (X, X) = 0 \Leftrightarrow X = 0$, 则称 X 为内积空间, 记作 $(X, (,))$ 如果令 $\| X \| = (X, X)^{1/2}$, 则 $(X, \| \cdot \|)$ 构成赋范空间。若 $(X, \| \cdot \|)$ 完备, 则 $(X, (,))$ 为希尔伯特空间。对内积空间 $(X, (,))$ 。 $X, y \in X$, 其夹角的余弦为 $\cos(X, y) = (X, y) / \| X \| \cdot \| y \|$ 由此, 二向量 X, y 正交(或垂直): $X \perp y$ 当且仅当 $(X, y) = 0$ 。若 $X_1, X_2, \dots, X_n (\in X)$ 两两正交, 则 $\sum_{i=1}^n X_i$ 的长度之平方为 $\| \sum_{i=1}^n X_i \|^2 = \sum_{i=1}^n \| X_i \|^2$, 即毕氏定理成立。设 $\{X_n\}_{i=1}^\infty$ 两两正交, 如果 $\lim_{n \rightarrow \infty}$

$\| X - \sum_{i=1}^n \alpha_i X_i \| = 0$, 即级数 $\sum_{i=1}^n \alpha_i X_i$ 前 n 项部分和之序列(平均)

收敛到 X , 记作 $X = \sum_{n=1}^{\infty} \alpha_n X_n$, 可证 $\alpha_n = (X, X_n)$ 。对于 $L^2(-\pi, \pi)$ 来说, 函数系 $\frac{1}{\sqrt{2\pi}}, \frac{\cos X}{\sqrt{\pi}}, \frac{\sin X}{\sqrt{\pi}}, \frac{\cos 2X}{\sqrt{\pi}}, \frac{\sin 2X}{\sqrt{\pi}}, \dots, \frac{\cos nX}{\sqrt{\pi}}, \frac{\sin nX}{\sqrt{\pi}}, \dots$ 是一个直交规范 (范数为 1) 函数系。任何函数 $X = X(t) \in L^2(-\pi, \pi)$ 可以展成关于它们的 Fourier 级数: $X(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt)$, 这里 $a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} X(t) dt$, $a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} X(t) \cos ntdt$, $b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} X(t) \sin ntdt$ 。

【线性算子】 设 X, Y 为向量空间。 T 为一映射, 其定义域记作 $D(T)$, 包含在 X 内: $D(T) \subset X$, 而其值域 (记作 $R(T)$) 包含在 Y 内: $R(T) \subset Y$ 。如果 T 满足条件: (1) $D(T)$ 构成向量空间; (2) 对任何 $\alpha \in K$ (标量域) 有 $T(X+y) = TX + Ty$ 及 $T(\alpha X) = \alpha TX$, 则称 T 为 $\mathcal{D}(T) \subset X$ 到 Y 内的线性算子。条件 (2) 也可以写成 $T(\alpha X + \beta y) = \alpha TX + \beta Ty$ 。线性算子把 X 中的零向量变为 Y 中的零向量, 集 $\{X \in$

$D(T) \mid TX = \theta\}$ 构成一个向量空间, 记作 $N(T)$, 称为 T 的零空间。 T 的所有值的全体 $\{y \in D(T) \subset Y, \exists X \in R(T), TX = y\}$ 也构成向量空间, 记作 $R(T)$, 称为 T 的值域。

线性算子的例子: (1) 恒等算子 $I_X: X \rightarrow X, I_X X = X \forall X \in X$; (2) 零算子 $O: X \rightarrow Y, OX = \theta \forall X \in X$; (3) 设 X 是 $[a, b]$ 上所有多项式组成的向量空间, 在 X 上定义算子 $T: (T_X)(t) = X'(t) = \frac{dX(t)}{dt} \forall X \in X$ 则 $R(T) = X, R(T) = X; N(T) = K$, 称此算子为

微分算子; (4) $X = C[a, b]$, $(TX)(t) = \int_a^t X(t) dt$, $t \in [a, b]$ 叫积分算子, $D(T) = X$; (5) $X = C[a, b]$, $(T_X)(t) = tX(t)$ 叫做乘法算子。

设 X, Y 同是实数域或复数域上的向量空间: $T: D(T) \rightarrow R(T)$ 为线性算子, $D(T) \subset X, R(T) \subset Y$, 若 T 为内射, 即 $X_1 \neq X_2 \Rightarrow TX_1 \neq TX_2$, 则存在算子 $S: D(S) = R(T), R(S) = D(T), Sy = X \Leftrightarrow TX = y$, 称 S 为 T 的逆算子, 记作 T^{-1} 。线性真子的逆存在的充要条件是 $TX = \theta \Rightarrow X = \theta$ 。线性算子的逆算子仍为线性算子。设 $(X, \| \cdot \|_1), (Y, \| \cdot \|_2)$ 都是 K 上的赋范空间。 $T: D(T) \rightarrow Y$ 是线性算子, $D(T) \subset X$ 。称 T 为有界算子, 如果存在实数 c , 使 $\| TX \|_2 \leq c \| X \|_1, \forall X \in X(T)$ 。称 $\sup \frac{\| TX \|_2}{\| X \|_1}$ 为算子 T 的范数, 记作 $\| T \|$ 。显然, $\| TX \|_2 \leq \| T \| \| X \|_1 \forall X \in D(T)$ 。

记 $B(X, Y)$ 为所有由 X 到 Y 内的有界线性算子组成的集。二算子 T, S 的和 $T+S: (T+S)X = TX + SX$, 数 α 与 T 的乘积 $\alpha T: (\alpha T)X = \alpha(TX)$, 则 $B(X, Y)$ 构成向量空间。又规定 $\| T \| = \sup_{X \in X, X \neq \theta} \frac{\| TX \|_2}{\| X \|_1}$, 则 $(B(X, Y), \| \cdot \|)$ 构成赋范空间。当 Y 完备时, $(B(X, Y), \| \cdot \|)$ 是完备赋范空间 (Banach 空间)。

赋范空间 $(X, \| \cdot \|_1)$ 到 $(Y, \| \cdot \|_2)$ 的线性算子 (2) 连续, 假如 $\forall X_n \in D(T), X_n \rightarrow X_0$ 有 $TX_n \rightarrow TX_0$ 。称 T 在 $D(T)$ 内连续假如 T 在每点 $X \in D(T)$ 连续。 $D(T)$ 连续当且仅当 T 在 θ 点连续。 T 在 $D(T)$ 连续, 当且仅当 T 有界。取值于实 (复) 数的线性算子叫做实 (复) 泛函。记作 $f(X)$ 或 $\langle f, X \rangle, X \in K$ 。由上可知, X 上的所有实 (复) 有界泛函全体, 以 $\| f \| = \sup_{X \in X, X \neq \theta} \frac{|f(X)|}{\| X \|} = X \in X, \sup_{\| X \| = 1} |f(X)|$ 为范数, 构成一个 B 空间。称为 X 的共轭空间, 记作 X^* 。若考察 X^* 上的所有界线性泛函的全体, 则得 X^* 的共轭空间, 记作

X^{**} , 称为 X 的二次共轭空间。当将 $X \in X$ 固定, f 变动时, $f(X)$ 就是 X^* 上的一个泛函, 记作 $F(f) = f(X)$ 它也有界, 且 $\|F\| = \|X\|$ 。作由 X 到 X^{**} 的映射 $\pi: X \rightarrow FX \in X^{**}$, 则 π 是 X 到 $\pi(X) \subset X^{**}$ 的线性等距同构。所以可以把 X 看做是 X^{**} 的一个子空间。若 X 是 B 空间, 则 $\pi(X)$ 是 X^{**} 的闭子空间。若 $\pi(X) = X^{**}$ 则说 X 自反。

对于一个空间来说, 首先要确定有没有 X 上的线性连续泛函。*Hahn-Banach* 定理指出, 若 $(X, \|\cdot\|)$ 是赋范空间, X_0 是它的线性子空间, 若 f 是 X_0 上的线性连续泛函, 则必能将 f 保范延拓到全空间 X 上。即: $\exists F \in X^*$, 使当 $X \in X_0$ 时, $F(X) = f(X)$, 且 $\|F\| = \|f\|$ 。由此即知, 当 $X \neq \{\theta\}$ 是赋范空间时, 必有足够多的泛函存在。

具体空间上线性连续泛函的一般形式为: (1) $C[a, b]$ 上的连续线性泛函的一般形式为 $f(X) = \int_a^b X(t)dv(t)$ 之形式。其中 v 是 $[a, b]$ 上的有界变差函数; (2) $L^p(a, b)$, $(1 < p < \infty)$ 上的线性连续泛函的一般形式为 $f(X) = \int_a^b X(t)y(t)dt$, 其中 $y \in L^{p'}, \frac{1}{p} + \frac{1}{p'} = 1$, $L^1(a, b)$ 的线性连续泛函的一般形式为 $f(X) = \int_a^b X(t)y(t)dt, y(t) \in L^\infty(a, b)$, 即本性有界可测函数。

设 $(X, \|\cdot\|_1)$ 是 B 空间, $(Y, \|\cdot\|_2)$ 是赋范空间。 $\{T_\alpha\}_{\alpha \in A}$ 是由 X 到 Y 内的线性连续算子的族, A 为指标集。对于每一个 $X \in X$, $\{\|T_\alpha X\|_2\}_{\alpha \in A}$ 是有界数集, 则 $\{\|T_\alpha\|_1\}_{\alpha \in A}$ 是有界数集, 这叫做一致有界定理或共鸣定理。由此一致有界定理可得许多有用的结果。

开映射定理内容: 设 X, Y 都是 B 空间, T 是线性有界算子, 且 $T(X) = Y$ (即 $D(T) = X, R(T) = Y$), 则 T 是开映射, 即把 X 中

的开集映成 Y 中的开集。

逆算子定理内容为: 设 X, Y 都是 B 空间, T 是 X 到 Y 上的一一线性有界算子。则 T^{-1} 有界。

设 $(X, \| \cdot \|_1), (Y, \| \cdot \|_2)$ 都是数域 K 上的赋范空间, 积空间 $(X \times Y, \| \cdot \|_1 + \| \cdot \|_2)$ 也是赋范空间。集 $D(T) \times R(T)$ 叫做 T 的图象, 记作 $G(T)$ 。

闭图象定理内容为: 设 $(X, \| \cdot \|_1), (Y, \| \cdot \|_2)$ 是 B 空间, T 是 $X \rightarrow Y$ 的闭算子 (G_T 是闭集) $D(T) \subset (X, \| \cdot \|_1)$ 是闭集, 则 T 是有界算子。设 (X, ρ) 为距离空间, T 为 $X \rightarrow X$ 的映射, 若 $\rho(Tx, Ty) \leq \alpha \rho(X, y), 0 < \alpha < 1$, 则称 T 为 X 上的压缩映射。

压缩映射原理 (*Banach* 不动点定理) 内容为: 设 (X, ρ) 是完备距离空间, T 是 $X \rightarrow X$ 的压缩映射, 则 T 有唯一不动点 (即有 $X^* \in X$ 使 $TX^* = X^*$)。

【算子的正则值与谱】 在应用中, 一个最基本的问题是解齐次方程

$AX = \lambda X$; 即 $(A - \lambda I)X = 0$; 以及解非齐次方程 $(A - \lambda I)X = y$, 其中 $X, y \in X$; 设 X 为赋范空间, $A: X \rightarrow X$ 为闭线性算子。称集合 $\rho(A) = \{\lambda \in \mathbb{C} \mid (\lambda I - A)^{-1} \text{ 为有界线性算子} \}$, 为 A 的预解集, $\rho(A)$ 中的 λ 叫做 A 的正则值。若 $\lambda \notin \rho(A)$, 就说 λ 属于 A 的谱集 $\delta(A)$, 称 λ 为 A 的一个谱点。从逻辑上说, 谱点有三种可能: (1) $(\lambda I - A)^{-1}$ 不存在。这时称 λ 为 A 的一个特征值 (或本征值, 或固有值), 特征值全体叫做点谱记作 $\sigma_p(A)$; (2) $(\lambda I - A)$ 存在, $R(\lambda I - A) \neq X$, 但 $R(\lambda I - A) = X$, 这时说 λ 属于 A 的连续谱 $\sigma_c(A)$; (3) $(\lambda I - A)^{-1}$ 存在, 但 $R(\lambda I - A) \neq X$; 这时说 λ 属于 A 的剩余谱, 记作 $\sigma_r(A)$ 。这三部分互不相交, 且 $\sigma(A) \approx \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$ 。

【变分法】研究求泛函极值的问题。举几个典型的问题,(1)最速降线问题,即要求一条曲线,连接不在同一铅直线上所给二点 A, B , 使当质点沿着这条曲线由 A 滑至 B 时所用时间最少。将坐标原点置于 A 点, O_x 轴放在水平位置, O_y 轴铅直向下。质点运动的速度为 $\frac{dX}{dt} = \sqrt{2gy}$ 。因此质点由 $A(0,0)$ 运动到 $B(X_1, y_1)$ 所用时间为

$$t[y(X)] = \frac{1}{\sqrt{2g}} \int_0^{X_1} \frac{\sqrt{1+y'^2}}{\sqrt{y}} dX;$$

$y(0)=0, y(X_1)=y_1$; (2)等周问题。求长度一定的封闭线 l , 使其所围面积 S 最大。将此曲线用参数 t 表示: $X=X(t), y=y(t), t \in [t_0, t_1]$, 则 $X(t_0)=X(t_1), y(t_0)=y(t_1), l=$

$$\int_{t_0}^{t_1} \sqrt{[\dot{X}(t)]^2 + [\dot{y}(t)]^2} dt = \text{常数}, \text{曲线所围面积 } S = \frac{1}{2} \int_{t_1}^{t_2} (\dot{X}y - y\dot{X})$$

dt 。问题是求 S 的极大值; (3)短程线问题。在曲面 $Q(X, y, z)=0$ 上, 给定两点 $A(X_0, y_0, z_0), B(X_1, y_1, z_1)$ 要在曲面上找一曲线, 使此曲线最短, 设曲线方程为 $y=y(X), z=z(X)$ 。它们应满足

$$\varphi(X, y(X), z(X))=0, \text{曲线 } AB \text{ 之长为 } l = \int_{X_0}^{X_1} \sqrt{1+y'^2+z'^2} dX,$$

问题是要求 l 的极小值; (4)求泛函 $v =$

$$\int_{X_0}^{X_1} F(X, y, y') dX \text{ 之极值, 但其边界点 } (X_0, y_0), (X_1, y_1) \text{ 分别允许在}$$

曲线 $y=f_0(X)$ 及 $y=f_1(X)$ 上移动。

泛函的概念, 首先在变分法中见到, 从而为泛函分析提供了这种概念。像数学分析那样, 可以定义泛函 $v[y(X)]$ 的变分 $\delta v = \frac{\partial}{\partial \alpha} v[y(X) + \alpha \delta y]$ 。如果 $v[y(X)]$ 及 $y=y_0(X)$ 上达到极值, 则在 $y=y_0(X)$ 上 $\delta v=0$ 。将此条件简化, Euler 得到泛函 $v[y(X)] =$

$\int_{X_0}^{X_1} F(X, y, y') dX$ 取极值的(必要)条件为 $F_y - \frac{d}{dX} F_{y'} = 0$, 这叫做 *Euler* 方程。(2)与(3)都是在某些条件下求极值的问题, 对于(3)可以证明: 函数 $y_1, y_2, \dots, y_n$ 使泛函 $v = \int_{X_0}^{X_1} F(X, y_1, \dots, y_n, y'_1, \dots, y'_n) dX$ 在条件 $\varphi_i(X, y_1, \dots, y_n) = 0 (i=1, \dots, m; m < n)$ 之下达到极值时, 对适当选择的 $\lambda_i(X) (i=1, 2, \dots, m)$, $y_1, \dots, y_n$ 必满足由泛函 $v^* =$

$\int_{X_0}^{X_1} \left(F + \sum_{i=1}^m \lambda_i(X) \varphi_i \right) dX = \int_{X_0}^{X_1} F^* dX$ 所做出的 *Euler* 方程, 而函数 λ_i

(X) 及 $y_j(X)$ 由 *Euler* 方程 $F_{y_j}^* - \frac{d}{dX} F_{y'_j}^* = 0 (j=1, \dots, n)$ 和方程 $\varphi_i = 0 (i=1, \dots, m)$ 确定, 由此, (3)的解可由

$\lambda \varphi_y - \frac{d}{dX} \frac{y'}{\sqrt{1+y'^2+z'^2}} = 0; \lambda Q_x - \frac{d}{dX} \frac{z'}{\sqrt{1+y'^2+z'^2}} = 0; \varphi(X, y, z) = 0$ 确定。由此变分法中求极值的问题, 变成了解微分方程的问题。

几 何 学

解析几何

【几何学】数学中最古老的一门分科,起源于古埃及尼罗河泛滥后为整修土地而产生的测量法。我国古代西周有勾股测量,即“勾三,股四,弦五”之说,后来几何学从埃及传到希腊,逐步发展成为一个比较严谨的推理几何。特别是欧几里得把当时得到的数学知识集其大成,编写出“原本”一书,对后世有很大的影响,欧几里得在“原本”中首先叙述一些定义,然后提出五个公设,其中以第五公设为最著名。至今我们把欧几里得在“原本”所述的几何学称为欧氏几何学。但是“原本”中的公理体系是不完善的,经过了许多数学家的长期努力在 19 世纪末,德国数学家希尔伯特才真正建立了严密的欧氏几何公理体系。在建立欧氏几何公理体系的长期过程中,19 世纪初罗巴切夫斯基和鲍约独立地创建了一种新几何学,它与欧氏几何的区别在于扬弃了欧氏的第五公设,这样创建的几何学称为双曲非欧几何学,后来黎曼又创建了椭圆非欧几何学。

【解析几何】数学中最基本的学科之一。解析几何的萌芽,起源于古希腊数学家对圆锥曲线所作的大量的系统研究,17 世纪初随解析几何学的创始人法国的笛卡尔和费马把变量引入数学领域,通过建立坐标,把几何中的曲线(或曲面)看成动点的轨迹,用变量 X, y (或 X, y, z) 来表示这个动点的坐标,那么,这个轨迹便可以用 X, y (或 X, y, z) 所满足的方程来表示。这样把几何学中的基本元素“点”和代数学中的研究对象“数”建立对应,并在这个基础上,建立

起曲线(或曲面)与方程的对应。这样,通过几何与代数的紧密结合,就可用代数的方法来研究几何问题,这便是解析几何的基本思想,解析几何只研究二次方程所表示的曲线或曲面。解析几何除上述坐标法外,还可通过向量的运算来讨论曲线和曲面的一些几何性质,这便是向量法,向量法对某些问题的讨论是很方便的。

【直角坐标与极坐标】为了决定平面上点的位置,选定两条互相垂直的直线称为坐标轴,两轴的交点称为坐标原点。通常把水平方向的轴叫横轴或 x 轴,且把从左到右取为正方向,垂直于水平方向的轴叫纵轴或 y 轴,且把从下到上取为正方向。最后再取一个单位长度,这样就建立了一个平面直角坐标系,也称为笛卡儿坐标系,记作 $o-Xy$ 。

设 M 为平面上任意一点, P 为 M 在 x 轴上投影(即 $MP \perp OX$, P 是垂足), Q 为 M 在 y 轴上投影,记长度 $OP = x$, $OQ = y$,分别称 x, y 为点 M 的横坐标和纵坐标,并记 $P(x, y)$ 。因此给出平面上一点 M 后,便有一对有序实数与之对应,反之亦然。

两坐标轴把平面分成四个部分,每一部分叫做象限。分别把右上、左上、左下、右下四个象限称为Ⅰ、Ⅱ、Ⅲ、Ⅳ象限。显然,这四个象限中点的坐标的符号为Ⅰ(+,+),Ⅱ(-,+),Ⅲ(-,-),Ⅳ(+,-)。

完全类似地可以建立空间直角坐标。从某定点 o 引三条互相垂直的直线 oX, oY, oZ 称为坐标轴, o 点称为坐标原点。称平面 YOZ, YOX, XOY 为坐标面,再取一个单位长度,这样就建立了一个空间直角坐标系,也称笛卡儿坐标系,记为 $o-XYZ$ 。

设 P 为空间任一点,过 P 作三平面分别与三坐标面平行,顺次交三个坐标轴于 P_X, P_Y, P_Z 三点,则有三数。

$$X=op_X, y=op_Y, z=op_Z$$

分别称为点 P 的横坐标,纵坐标和立坐标,记 $P(X, y, z)$ 。因此给出空间一点 P 后,便有三个有序实数与之对应,反之亦然。

附带说明,如果将右手的母指和食指分别指着 X 轴和 y 轴的正方向,而中指方向与 z 轴正方向相同,则称此坐标系为右手坐标系,否则称为左手坐标系。一般均采用右手坐标系。

在平面内取一定点 O , 叫做极点。由 O 引一条射线 OX 叫做极轴, 再确定一个长度和计算角度的正方向, (通常取逆时针方向为正向) 这样就建立了一个极坐标系。

设 P 是平面上任一点, 连结 OP 。设 $OP = \rho$, $\angle XOP = \theta$, 则 ρ, θ 的值就确定了。 ρ 叫做 P 点的极半径, θ 叫做 P 点的极角。 (ρ, θ) 叫做 P 点的极坐标。显然每一对有序实数 (ρ, θ) 决定一个点的位置。

又 θ 可取任何正值或负值, 同时为了方便起见, ρ 也可取负值。 ρ 的正, 负依下法规定: 若 P 点位于 θ 角终边上, 则 ρ 为正, 若 Q 点位于 θ 角终边的延长线上, 则 ρ 为负。

【直线】若直线 l 与 X 轴相交, 它们的右上方的交角称为直线 l 的倾斜角, 简称倾角, 记为 α 。当直线 l 与 X 轴平行时, 规定 $\alpha = 0$ 。因此倾角的取值范围是 $0 \leq \alpha < \pi$ 。

直线 l 倾角 α 的正切值, 即 $\tan \alpha$ 叫做这条直线的斜率, 一般用 k 来表示, k 即 $\tan \alpha$ 。当直线 l 与 X 轴垂直时, 即 $\alpha = \frac{\pi}{2}$, 它的斜率不存在。

经过 $P_0(X_0, y_0)$ 点, 斜率为 k 的直线方程为

$$y - y_0 = k(X - X_0)$$

此方程称为点斜式。如果直线平行于 y 轴, 则斜率不存在, 此时这条直线方程为 $X = X_0$ 。

直线斜率为 k , 在 y 轴上的截距为 b 的方程为

$$y = kX + b$$

此方程称为斜截式。

经过两点 $p_1(X_1, y_1)$ 与 $p_2(X_2, y_2)$ 的直线方程为

$$\frac{y - y_1}{y_2 - y_1} = \frac{X - X_1}{X_2 - X_1}$$

此方程称为两点式。如果 $y_1 = y_2$ 此时直线方程为 $y = y_1$; 如果 $X_1 = X_2$, 此时直线方程为 $X = X_1$ 。

直线在 X 轴和 y 轴上的截距分别为 a, b 时 ($a \neq 0, b \neq 0$) 直线方程为

$$\frac{X}{a} + \frac{y}{b} = 1$$

此方程称为截距式。

任何一条直线均可表示为

$$AX + By + C = 0 \quad (A, B \text{ 不同时为零}) \text{ 此方程为一般式。}$$

在空间直角坐标系, 也可以讨论空间直线的方程。设直线过定点 $P_0(X_0, y_0, z_0)$, 它的方向可用向量 $\vec{s} = (l, m, n)$ 来表示, 此直线方程可写为:

$$\begin{cases} X = X_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases} \quad (-\infty < t < \infty)$$

这组方程叫做直线的参数式, t 称为参数。

从直线参数方程可得出:

$$\frac{X-X_0}{l}=\frac{g-g_0}{m}=\frac{z-z_0}{n}$$

这组方程称为直线的对称式。

过两定点 $P_1(X_1, y_1, z_1)$ 与 $P_2(X_2, y_2, z_2)$ 的直线方程为

$$\frac{X-X_1}{X_2-X_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1}$$

这组方程叫做直线的两点式。

因为一条直线也可以看作某两个平面相交而得到,因此下述方程组也表示一条直线。

$$\begin{cases} A_1X+B_1y+C_1Z+D_1=0 \\ A_2X+B_2y+C_2z+D_2=0 \end{cases}^{\circ}$$

【椭圆、双曲线、抛物线】一动点到两定点的距离之和等于定量,动点的轨迹称为椭圆,这两定点称为焦点(此时这定量必须大于两焦点间的距离)以 $2a$ 表示这定量。把两定点 F 与 F' 的连线作为 X 轴, $F'F$ 线段的中点取为坐标原点,以 $2c$ 表示焦点间的距离,即 $F'F=2c$,则此椭圆的方程为

$$\frac{X^2}{a^2}+\frac{y^2}{b^2}=1$$

其中 $b^2 = a^2 - c^2$ ($a > c$) 此方程称为椭圆的标准方程。此时两坐标轴为椭圆的对称轴, 原点为椭圆的对称中心, 椭圆与两对称轴的交点称为顶点。在取标准方程时, 椭圆四个顶点的坐标为

$$A(a, 0), A'(-a, 0), B(0, b), B'(0, -b)$$

连接椭圆的相对的顶点的线段 AA' 和 BB' 的长度为 $2a$ 和 $2b$, 分别称为椭圆的长轴和短轴。

量 $\frac{c}{a}$ 称为椭圆的离心率, 记为 e , 即 $e = \frac{c}{a}$ 。显然椭圆的离心率 $e < 1$ 。

两条垂直于椭圆长轴且与中心距离为 $\frac{a}{e}$ 的直线称为椭圆的准线, 准线的方程为 $X = \frac{a}{e}$ 与 $X = -\frac{a}{e}$ 。准线有下述性质: 椭圆上任意一点到焦点和对应的准线距离之比等于离心率 e 。

一动点到两定点的距离之差等于一定量, 动点的轨迹称为双曲线。这两定点称为焦点。用 $2a$ 表示这定量。把两定点 F 与 F' 边的连线为 X 轴, $F'F$ 线段的中点取为坐标原点, 以 $2c$ 表示焦点间距离, 即 $FF' = 2c$, 则此双曲线方程为

$$\frac{X^2}{a^2} - \frac{y^2}{b^2} = 1$$

其中 $b^2 = c^2 - a^2$ ($c > a$) 此方程称为双曲线的标准方程。此时两坐标轴为双曲线的对称轴, 原点为双曲线的对称中心。双曲线与对称轴的交点称为顶点, 在取标准方程时, 双曲线的两个顶点坐标为

$$A(a, 0), A'(-a, 0)$$

因为双曲线与 y 轴的交点坐标为虚数 $(0, \pm bi)$ 即 y 轴与双曲线不相交, 因此把与双曲线相交的对称轴称为实对称轴, 与双曲线不相交的对称轴称为虚对称轴。连接双曲线顶点 $A'A$ 的线段称为双曲线的实轴。若从虚对称轴上从中心 O 截取线段 OB 与 OB' , 使其长度等于 b , 则线段 BB' 称为双曲线虚轴。

过双曲线中心的直线 $y = \frac{b}{a}X$ 与 $y = -\frac{b}{a}X$, 称为双曲线的渐近线。它有这样一个性质: 对于双曲线上任意点 P 沿双曲线趋于无穷远时, 它与直线 $y = \frac{b}{a}X$ 或 $y = -\frac{b}{a}X$ 的距离会趋于零。即 P 点会无限地靠近此直线。

量 $\frac{c}{a}$ 称为双曲线的离心率, 记为 e , 即 $e = \frac{c}{a}$ 。显然双曲线的离心率 $e > 1$ 。

两直线 $X = \pm \frac{a}{e}$ 称为双曲线的准线, 它垂直于双曲线的实轴, 且到中心的距离等于 $\frac{a}{e}$ 。因为 $e > 1$, 所以 $\frac{a}{e} < a$, 因此两准线在两顶点之间, 准线有下面性质: 双曲线上任意点到一焦点与对应的准线的距离之比等于离心率 e 。

与一定点和一定直线等距离的点的轨迹称为抛物线, 定点称为焦点, 定直线称为准线。(假设定点不在定直线上) 通过焦点 F 引准线的垂线(从准线到焦点的方向取为正向) 把它取为 X 轴, 从点 F 到定直线的线段中点作为坐标原点, 且以 p 表示 F

到定直线的线段的长,则焦点 F 坐标为 $(\frac{p}{2}, 0)$, 此时抛物线方程为

$$y^2 = 2pX$$

称为抛物线的标准方程。此时, X 轴为抛物线的对称轴, 抛物线与对称轴的交点为顶点, 抛物线只有一个顶点。

我们把椭圆、双曲线、抛物线统称为圆锥曲线。

【向量】空间中的一个有向线段叫做向量或矢量。用 $\overrightarrow{AB}$ 来表示, A 表示向量的起点, B 表示向量的终点。有时也可用一个字母 $\vec{a}$ 来表示。因此向量有两个要素: (1) 向量的长度, 也叫做模, 用 $|\overrightarrow{AB}|$ 或 $|\vec{a}|$ 来表示。如果确定了一个单位长度后, 向量的模是一个非负实数。(2) 向量的方向, 向量 $\overrightarrow{AB}$ 的方向是指起点 A 到终点 B 的方向。

如果两个向量的长度与方向均相同, 则称这两个向量相等。如果把相等向量看作是起点在不同位置的同一向量, 这样的向量称为自由向量。向量的自由性表现在向量 $\overrightarrow{AB}$ 的起点可以随着 $\overrightarrow{AB}$ 方向平行移动而改变。

在空间直角坐标系下任意一点 P 的位置可以用一个向量 $\overrightarrow{OP}$ 来表示。(O 点是坐标原点) 因此向量 $\overrightarrow{OP}$ 称为 P 点的位置向量, 设 P 点的坐标为 (X, y, z) 则向量 $\overrightarrow{OP}$ 与 P 点可建立 1-1 对应。因此我们也可以把 (X, y, z) 作为向量 $\overrightarrow{OP}$ 的坐标, 并记为: $\overrightarrow{OP} = (X, y, z)$ 。长度为零的向量称为零向量。

两个自由向量 $\vec{a}$ 和 $\vec{b}$ 可以经过平移后使起点落在同一点 O 上, 以这两个向量为邻边所作平行四边形之对角线, 起点为 O 所表示的向量称为该两向量之和, 记作 $\vec{a} + \vec{b}$ 。这种运算叫做向量的加法。向量的加法也可以用如下方式来完成: 由第一个向量的终点引出第二个向量, 则以第一向量的起点为起点, 第二个向量的终点为终点的向量叫做这两个向量之和。

两个向量的差是一个向量, 它与所减向量之和就是被减向量, 这种运算叫向量的减法, 例如, $\vec{a} - \vec{b} = \vec{c}$, 即 $\vec{c} + \vec{b} = \vec{a}$ 。

实数 λ 与向量 $\vec{a}$ 的乘积 $\lambda \vec{a}$ 是一个向量, 当 $\lambda = 0$ 时或 $\vec{a} = 0$ 时, 它是一个零向量; 当 $\lambda \neq 0$, 且 $\vec{a} \neq 0$ 时, 若 $\lambda > 0$, 则 $\lambda \vec{a}$ 的模为向量 $\vec{a}$ 的模的 λ 倍且方向与 $\vec{a}$ 相同; 若 $\lambda < 0$, 则 $\lambda \vec{a}$ 的模是向量 $\vec{a}$ 的模

的 $|\lambda|$ 倍,且方向与 $\vec{a}$ 相反。

与已知向量同向且模为1的向量叫做这个向量的单位向量。

显然若 $\vec{a} \neq 0$,则 $\vec{a}_0 = \frac{\vec{a}}{|\vec{a}|}$ 就是向量 $\vec{a}$ 的单位向量。

在空间直角坐标系中,在 X 轴, y 轴, z 轴的正方向上各取一个单位向量 $\vec{i}, \vec{j}, \vec{k}$ 。容易看出,若 $\vec{OP}$ 向量的坐标为 $\vec{OP} = (X, y, z)$,则 $\vec{OP} = X\vec{i} + y\vec{j} + z\vec{k}$ 。换言之,任意一个向量均可仿上述方式进行分解,因此我们把 (X, y, z) 也叫做向量的分量。

向量 $\vec{a}$ 与 $\vec{b}$ 的内积(或称数量积、点积)是一个数量,它的大小是它们的模和它们夹角 $\varphi (0 \leq \varphi \leq \pi)$ 的余弦的乘积,即若用 $\vec{a} \cdot \vec{b}$ 表示向量 $\vec{a}$ 与 $\vec{b}$ 的内积,则

$$\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}| \cos \varphi$$

向量的内积满足以下运算律:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a};$$

$$(\lambda \vec{a}) \cdot \vec{b} = \lambda (\vec{a} \cdot \vec{b});$$

$$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$$

容易看出, $\vec{i} \cdot \vec{i} = \vec{j} \cdot \vec{j} = \vec{k} \cdot \vec{k} = 1; \vec{i} \cdot \vec{j} = \vec{j} \cdot \vec{k} = \vec{k} \cdot \vec{i} = 0$

如果 $\vec{a} = (a_1, a_2, a_3), \vec{b} = (b_1, b_2, b_3)$,则 $\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 +$

$a_3 b_3$,因此向量 $\vec{a}$ 与 $\vec{b}$ 之间夹角可以用下述公式来计算

$$\cos \varphi = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{b_1^2 + b_2^2 + b_3^2}}$$

向量 $\vec{a}$ 与 $\vec{b}$ 的外积(或称向量积、叉积)是一个向量,记作 $\vec{a} \times \vec{b}$,它的模是 $\vec{a}$ 和 $\vec{b}$ 的模与其夹角 φ 的正弦的乘积,它的方向垂直于 $\vec{a}, \vec{b}$ 所在平面,且 $\vec{a}, \vec{b}$ 与 $\vec{a} \times \vec{b}$ 构成右手系,即

$$\vec{a} \times \vec{b} = |\vec{a}| |\vec{b}| \sin \varphi \cdot \vec{e}$$

其中 $\vec{e}$ 是垂直于 $\vec{a}, \vec{b}$ 所在平面的单位向量, 且 $\vec{a}, \vec{b}, \vec{e}$ 成右手系。

向量外积满足如下运算规律:

$$\begin{aligned}(\lambda \vec{a}) \times \vec{b} &= \lambda(\vec{a} \times \vec{b}); \\(\vec{a} + \vec{b}) \times \vec{c} &= \vec{a} \times \vec{c} + \vec{b} \times \vec{c}; \\\vec{a} \times \vec{b} &= -\vec{b} \times \vec{a}\end{aligned}$$

容易验证, $\vec{i} \times \vec{i} = \vec{j} \times \vec{j} = \vec{k} \times \vec{k} = 0$, $\vec{i} \times \vec{j} = \vec{k}$, $\vec{j} \times \vec{k} = \vec{i}$, $\vec{k} \times \vec{i} = \vec{j}$ 。

设 $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$, 则

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

两个向量的外积与第三个向量再作内积, 所得的数量称为三个向量的混合积。若三个向量, $\vec{a}, \vec{b}, \vec{c}$ 的混合积记作 $(\vec{a}, \vec{b}, \vec{c})$, 即有

$$(\vec{a}, \vec{b}, \vec{c}) = (\vec{a} \times \vec{b}) \cdot \vec{c}$$

设 $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$, $\vec{c} = (c_1, c_2, c_3)$ 则

$$(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}。$$

【平面】与一个平面垂直的非零向量叫做此平面的法向量。通过定点 $P_0(X_0, y_0, z_0)$ 以 $\vec{N} = (A, B, C)$ 为法向量的平面方程为 $A(X - X_0) + B(y - y_0) + C(z - z_0) = 0$ 这方程称为平面的点法式。

在空间直角坐标系中三元一次方程 $AX + By + Cz + D = 0$ 表示一个平面, 反之一个平面的方程是一个三元一次方程。

通过三个不共线的点 $P_i(X_i, y_i, z_i) (i=1, 2, 3)$ 的平面方程为

$$\begin{vmatrix} X-X_1 & y-y_1 & z-z_1 \\ X_2-X_1 & y_2-y_1 & z_2-z_1 \\ X_3-X_1 & y_3-y_1 & z_3-z_1 \end{vmatrix} = 0$$

设平面在 X, y, z 轴上的截距分别为 a, b, c 则此平面方程为

$$\frac{X}{a} + \frac{y}{b} + \frac{z}{c} = 1$$

两个平面 π_1 与 π_2 相交时, 形成四个二面角, 即两对相对的二面角。两平面 π_1 与 π_2 的夹角是指相邻两个中的一个。它的大小用二面角的平面角来度量, 因此两平面之间夹角也可以用这两个平面的法向量之间夹角来度量。

设平面 π_i 的方程为 $A_iX + B_iy + C_iZ + D_i = 0 (i=1, 2)$ 则它们的法向量为 $\vec{N}_i = (A_i, B_i, C_i)$, 故 π_1 与 π_2 之间的夹角 θ 为

$$\begin{aligned} \cos\theta &= \frac{\vec{N}_1 \cdot \vec{N}_2}{|\vec{N}_1| |\vec{N}_2|} \\ &= \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2} \sqrt{A_2^2 + B_2^2 + C_2^2}}. \end{aligned}$$

【椭圆面、双曲面、抛物面】在空间选取一个直角坐标系 $o - Xyz$ 后, 方程

$$\frac{X^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

所表示的曲面叫做椭圆面或称椭球面。上述方程称为椭圆面的标准方程。椭圆面关于三个坐标面, 三个坐标轴和原点都是对称的, 我们把这三个对称面, 三条对称轴和原点分别称为椭圆面的主平面, 主轴和中心。椭圆面与三对称轴有六个交点, 这三对交点称为

椭圆面的顶点。把三对顶点之间距离,即 $2a, 2b, 2c$ 叫做主轴长。如果用三个坐标面去截椭圆面,而得到的截口均为椭圆。

方程

$$\frac{X^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

所表示的曲面叫做单叶双曲面。上述方程称为单叶双曲面的标准方程。单叶双曲面关于三个坐标面,三条坐标轴和原点都对称。我们把这三个对称面,三条坐标轴及原点分别称为单叶双曲面的主平面,主轴和中心。 z 轴与 y 轴与单叶双曲面有四个交点,这两对交点称为单叶双曲面的顶点。这两对顶点之间距离,即 $2a, 2b$ 叫做主轴长。如果用坐标面 yoZ 或 zoX 去截单叶双曲面得出的截口均为双曲线。用 Xoy 坐标面去截,得出的截口是椭圆。

方程

$$\frac{X^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

所表示的曲面叫做双叶双曲面。上述方程称为双叶双曲面的标准方程。双叶双曲面关于三个坐标面,三条坐标轴及原点都是对称的。我们把这三个对称面,三条对称轴及原点分别称为双叶双曲面的主平面,主轴和中心。 X 轴与双叶双曲面有两个交点,这一对交点称为双叶双曲面的顶点,这一对顶点之间距离即 $2c$ 称为主轴长。如果用 yoZ 和 zoX 坐标面去截双叶双曲面得出的截口均为双曲线。

方程

$$\frac{X^2}{b^2} + \frac{y^2}{b^2} = \frac{2z}{c}$$

所表示的曲面叫做椭圆抛物面,上述方程称为椭圆抛物面的标准方程。椭圆抛物面关于 yoZ 及 zoX 坐标面是对称的,这两个对称面

称为主平面,椭圆抛物面关于 z 轴对称, z 轴称为主轴。除原点外椭圆抛物面与坐标轴无其它交点,我们把原点称为顶点,如果用 yOz 及 zOx 坐标面去截椭圆抛物面得出的截面均为抛物线。

方程

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{2z}{c}$$

所表示曲面叫做双曲抛物面,又称马鞍面。上述方程称为双曲抛物面的标准方程,它关于 yOz 及 zOx 坐标面对称。这两个对称面叫做主平面,曲面关于 z 轴对称, z 轴称为主轴,除原点外曲面与坐标轴无其它交点,把原点称为顶点。如果用 yOz 及 zOx 坐标面去截双曲抛物面得出的截面是抛物线,用 xOy 坐标面去截得出的截面是一对相交直线。

几何基础

【几何基础】欧几里德的“原本”中所提出的欧氏几何公理体系是很不完善的。许多概念缺乏公理基础,为了建立一套完善的欧氏几何公理体系,在 1899 年德国数学家希尔伯特出版了“几何基础”一书。他提出了一整套接近于欧几里得且最受人们欢迎的公理体系,而全部欧氏几何的内容都可由这套公理体系按逻辑推理展开而得到。这就为欧氏几何学奠定了坚实的基础。而且希尔伯特的公理体系还有一个美妙的特点,即如果用罗巴切夫斯基的平行公理代替欧氏平行公理,其余公理不变便可立即得到双曲非欧几何的公理体系。

【第五公设问题】欧几里德在《原本》中提出了如下一条公设：在一平面上，若一直线与两直线相交，且若同侧所交两内角之和小于两直角，则两直线无限延长后必相交于该侧的一点。这个公设称为欧几里德第五公设。它在欧氏几何体系中占有特殊的地位。首先它较晚地显示出它的作用，在《原本》中前 28 个命题的证明中均未用到它。其次它看起来较其它所有公设均复杂。这就很自然地引起人们的一种想法，以为这条公设是否是多余的，即是否可以作为定理，利用其它公理来加以证明。事实上，欧几里德以后的许多数学家都曾作过企图证明第五公设的尝试。

关于某一命题的数学证明，其意义就是要仅根据所采用的公理体系纯逻辑地推导出此命题的结论。（只能用此系统中的公理及用这些公理已证明过的定理作为依据）但是《原本》中欧几里德并没有提出完善的公理系统，因此这些企图证明第五公设者都不能精确地提出问题。而所有这些第五公设的证明，从今天的观点来看都是不对的。因为他们都自觉或不自觉，或明或暗地不加证明地运用了某个与第五公设等价的命题。因此两千年来对第五公设的证明都失败了。事实上，在 19 世纪末以前，几何学现代公理法出现之前对于如何辨别几何学证明的严密性是没有完全清楚的准绳。因此在一定程度上，第五公设的每个证明者都自以为他的假设是合理的，或他所利用的直观是合理的。而且认为他已经证明了第五公设。但是直到今天才知道这些假设，或所利用的直观都是站不住脚的，实际情况是这些假设或直观都隐含了一个与第五公设等价的命题。

虽然企图证明第五公设都失败了，但在这各种各样的证明中积累了丰富的几何知识。而正是从这些试证第五公设的过程中才

产生了非欧几何的初步,以致引起人们对几何学观点的根本变化。

【希尔伯特公理体系】欧几里德《原本》有逻辑的缺点,这不仅表现在公理不够,而且许多重要的几何概念,如顺序、运动、连续等都没有定义。事实上,这些概念在 19 世纪以前都没有完全明确地建立起来。德国数学家希尔伯特(*D. Hilbert*)在总结了前人成果的基础上于 1899 年出版了《几何基础》一书。希尔伯特在此书中成功地建立了几何学的公理系统,并非常自然地划分公理,使得几何学的逻辑结构变得非常清楚,非常严密。希尔伯特所提出的欧氏几何这套完整而简单的公理系统,使欧氏几何公理体系问题原则上得到解决。

在希尔伯特的系统里讨论了三种几何对象:“点”、“直线”和“平面”,以及这些对象之间的三种关系:“属于”、(或说“关联”,“在…之上”等等)“介于”、“合同于”。(或说“全等于”)这些就是基本概念,所有其余的概念都可以在列举的三个基本概念的基础上给以直接的定义。

以上所说的三种基本对象和三种基本关系必须满足以下五组公理(共 20 条,见附录)

I₁₋₈ 关联公理(或称结合公理,从属公理)

II₁₋₄ 顺序公理(或称次序公理)

III₁₋₅ 合同公理(或称全合公理,全等公理)

IV 平行公理

V₁₋₂ 连续公理

根据这五组公理可以逻辑地导出全部欧氏几何。希尔伯特《几何基础》一书,已被世界上一些数学家公认为一本经典著作,希尔伯特的公理体系是欧氏几何的一套最完美的公理体系。

附录:希尔伯特公理体系

第一组公理:关联公理

I_1 对于两点 A 和 B , 恒有一直线 a , 它同 A 和 B 这两点的每一点相关系。

I_2 对于两点 A 和 B , 至多有一直线, 它同 A 和 B 这两点的每一点相关联。

I_3 一直线上恒至少有两点, 至少有三点不在同一直线上。

I_4 对于不在同一直线上的任意三点 A, B 和 C , 恒有一平面 α , 它同 A, B 和 C 这三点的每一点相关联, 对于任一平面, 恒有一点同这平面相关联。

I_5 对于不在同一直线上的三点 A, B 和 C , 至多有一平面, 它同 A, B 和 C 这三点的每一点相关联。

I_6 若一直线 α 的两点 A 和 B 在一平面 α 上, 则 α 的每一点都在平面 α 上。

I_7 若两平面 α 和 β 有一公共点 A , 则它们至少还有一公共点 B 。

I_8 至少有四点不在同一平面上。

第二组公理:顺序公理

II_1 若一点 B 在一点 A 和一点 C 之间, 则 A, B 和 C 是一直线上的不同的三点, 这时, B 也在 C 和 A 之间。

II_2 对于两点 A 和 C , 在线 AC 上恒至少有一点 B , 使得 C 在 A 和 B 之间。

II_3 一直线的任意三点中, 至多有一点在其他两点之间。

II_4 设 A, B 和 C 是不在同一直线上的三点: 设 a 是平面 ABC 的一直线, 但不通过 A, B, C 这三点中的任一点, 若直线 a 通过线段 AB 的一点, 则它必定也通过线段 AC 的一点, 或线段 BC 的一

点。

第三组公理:合同公理

Ⅲ₁ 设 A 和 B 是一直线 a 上的两点, A' 是这直线或另一直线 a' 上的一点, 而且给定了直线 a' 上 A' 的一侧, 则在直线 a' 上 A' 的这一侧, 恒有一点 B' , 使得线段 AB 和线段 $A'B'$ 合同或相等; 用记号表示, 即 $AB \equiv A'B'$ 。

Ⅲ₂ 若两线段 $A'B'$ 和 $A''B''$ 都和另一线段 AB 合同, 则这两线段 $A'B'$ 和 $A''B''$ 也合同; 简言之, 若两线段都和第三线段合同, 则它们彼此也合同。

Ⅲ₃ 设两线段 AB 和 BC 在同一直线 a 上, 无公共点, 而且两线段 $A'B'$ 和 $B'C'$ 在这直线或另一直线 a' 上亦无公共点, 若 $AB \equiv A'B'$, 而且 $BC \equiv B'C'$, 则 $AC \equiv A'C'$ 。

Ⅲ₄ 设给定了一平面 α 上的一个角 $\angle(h, k)$, 一平面 α' 上的一直线 a' , 和在 α' 上 a' 的一侧。设 h' 是 a' 上的, 从一点 o' 起始的一条射线, 则平面 α' 上恰有一条射线 k' , 使 $\angle(h, k)$ 与 $\angle(h', k')$ 合同或相等, 而且使 $\angle(h', k')$ 的内部在 a' 的这给定了的一侧; 用记号表示, 即 $\angle(h, k) \equiv \angle(h', k')$ 每一个角和它自己合同, 即 $\angle(h, k) \equiv \angle(h, k)$ 。

Ⅲ₅ 若两个三角形 ABC 和 $A'B'C'$ 有下列合同式

$$AB \equiv A'B', AC \equiv A'C', \angle BAC \equiv \angle B'A'C'$$

则也恒有合同式 $\angle ABC \equiv \angle A'B'C'$ 。

第四组公理: 平行公理。 IV (欧几里德公理) 设 a 是任一直线, A 是 a 外的任一点。在 a 和 A 所决定的平面上, 至多有一条直线通过 A , 而且不和 a 相交。

第五组公理: 连续公理

V₁ (度量公理或阿基米德公理) 若 AB 和 CD 是任意两线段, 则必

存在一个数 n , 使得沿 A 到 B 的射线上, 自 A 作首尾相接的 n 个线段 CD , 必将越过 B 点。

V_2 (直线完备公理) 一直线上的点集连同其顺序关系与合同关系不可能再这样扩充, 使得这直线上原来元素之间所具有的关系, 从公理 $I \sim III$ 所推出的直线顺序与合同的基本性质以及公理 V_1 都仍保持。

【绝对几何、欧氏几何与罗氏几何】在希尔伯特公理体系中, 去掉欧氏平行公理 IV , 即利用 I_{1-8} , II_{1-4} , III_{1-5} , V_{1-2} 这四组公理建立起来的几何学叫做绝对几何。因此在绝对几何中包含了欧氏几何也包含了罗氏几何。(即欧氏平行公理 IV 不成立的几何, 有关罗氏几何后面再详述) 换言之, 在绝对几何中成立的定理, 在欧氏几何与罗氏几何中都成立。例如: “在三角形中一个外角大于其任一不相邻的内角。”、“等腰三角形的底角相等。”等这些都是绝对几何中的定理。

由希尔伯特公理体系 (即 I_{1-8} , II_{1-4} , III_{1-5} , IV , V_{1-2}) 所建立起来的几何学称为欧氏几何, 如果在希尔伯特公理体系中, 我们去掉欧氏平行公理 IV , 并用它的反面命题 IV^* (即罗氏平行公理): 有这样的直线 a 和不在其上的点 A , 过 A 至少有两条直线与 a 共面不相交。来代替, 由这样的公理体系, 即 I_{1-8} , II_{1-4} , III_{1-5} , IV^* , V 所建立起来的几何学称为罗巴切夫斯基几何, 简称罗氏几何, 有时也称非欧几何。

从直观上看, 人们很难相信公理 IV^* 能够能立。因此罗氏几何的产生有一段漫长的历史。从欧几里德提出第五公设以后 (第五公设与欧氏平行公理是等价的) 就有一些数学家企图去“证明”它, 但均未成功。17 世纪以后有些数学家从假设第五公设不成立的情况

下着手,追究它能否得出与已知定理相矛盾的结果。如果得不出矛盾,它又会产生怎样的事实。实际上,这样的思想方法已开辟了一条通向非欧几何的道路。例如 18 世纪瑞士数学家 *Lambert* 就认识到,假如欧氏平行公理不成立的话,三角形的“内角和”就小于平角。他把平角与“内角和”之差叫做这个三角形的角亏。他证明了三角形的面积与角亏成正比。他还说过这样一段颇有深意的话:“我几乎可以推断,平行公理不成立的那种几何应该可以发生在半径为虚数的球面上!”

以后法国数学家 *Legendre* 证明了:“若有一个三角形的内角和是平角,则一切三角形的内角和都是平角;若有一个三角形的内角和小于平角,则一切三角形的内角和都小于平角。”并证明了在存在大小不同的相似三角形的假设下,可以证明第五公设。这些事实的发现对第五公设的本质认识深入了一大步。

到了 17 世纪,高斯(*Gauss*)首先承认了第五公设是不可能证明的,亦即它与其它公理不相依赖。因此高斯已预感到有一种新的几何存在的可能性。

匈牙利数学家 *J. Bolyai* 在他父亲的一本书中以附录形式于 1832 年出版了题为“绝对空间的科学”一文。在此篇论文中 *Bolyai* 建立起非欧几何学,并且相信这种新几何是无矛盾的。

俄国数学家罗巴切夫斯基在 1826 年于喀山大学的一次会议上宣读了题为“关于几何原理的议论”的报告,第一次提出了他关于非欧几何的思想,在 1829 年他正式发表了题为“论几何基础”的论文,叙述了他关于平行公理的研讨和他对新几何的探索。这种几何以后就称为罗氏几何。

【公理体系的相容性、独立性、完备性】若由公理系统 Σ 不可能导

出两个互相矛盾的命题,则称 Σ 是相容的或无矛盾的。

一个公理体系 Σ 是否有意义,其首要问题就是 Σ 必须是相容的,否则它在逻辑上就站不住脚。因此更谈不上有什么实际价值了。如何来确认一个公理体系 Σ 是相容的呢?显然单靠演绎推理的方法是不行的,因为无论推出多少个定理没有发现矛盾也不能保证继续推下去永远不会发现矛盾,因此数学上要证明公理体系是相容的必须采用另外的方法,即必须采用作模型(或说解释)的方法。

对于一个公理体系 Σ ,若有一组其性质已知的具体的事物 M (或理论体系),在规定 Σ 中每一基本概念指 M 中某一具体事物后,可验证 Σ 的每个公理在 M 中都成立,则称 M 为公理体系 Σ 的一个解释或模型。只要公理体系 Σ 有一个模型 M 存在,则 Σ 就是相容的。因为 Σ 中的命题与 M 中的事物之间已建立起 1—1 对应关系,如果由 Σ 能推出两个互相矛盾的命题,这矛盾必然同样会出现在模型 M 的事物中,但 M 是已知存在的事物(或理论体系)不可能在逻辑上有矛盾。因此利用模型来证明 Σ 相容时,我们只证明了公理体系 Σ 在 M 存在的条件下是无矛盾的,即 Σ 相容是有条件的。例如,我们可以在承认实数理论 M 存在的条件下,来证明欧氏几何体系 Σ 是相容的。而且在承认欧氏几何是相容的前提下证明罗氏几何 Σ^* 是相容的。

对公理体系 Σ 的独立性的要求就是确承在 Σ 中的每个公理都是必要的,不是多余的。事实上,若其中某个公理能由 Σ 中其它公理推出,则应把此公理列为定理,即由 Σ 中去掉它,而对 Σ 的推论不会发生影响。所以独立性问题也就是在保留同样多推论的前提下, Σ 中公理最少个数的问题。

如果相容公理体系 $\Sigma = \{A_1, A_2, \dots, A_n\}$ 中某个公理 $A_i (1 \leq i$

$\leq n$)不能由另外 $(n-1)$ 个公理导出,则称 A_i 在 Σ 中是独立的。因此如果相容公理体系 Σ 中一个公理 A_i 换成它的反面命题后所得到的新系统仍是无矛盾的,则 A_i 在 Σ 中就是独立的。所以证明某个公理 A_i 在 Σ 中是独立的,仍需用作模型的方法。

有时我们希望一个公理体系 Σ 中的公理个数是足够的,也就是说 Σ 所刻划的几何是唯一的。怎样来理解 Σ 中公理个数是足够的呢?如果把 Σ 中的任一命题 A (A 中不引入新的基本概念)作为公理加入 Σ 体系中,得到的结果或 A 是 Σ 的推论,即 A 在 Σ 中不独立;或 A 与 Σ 组成的公理体系是矛盾的,即 A 与 Σ 是不相容的。那么我们称公理体系 Σ 是完备的,即 Σ 中公理个数是足够的了。

一个公理体系 Σ 总是有可能有许多具体模型,若公理体系 Σ 的两个模型 M_1 和 M_2 所有的同名对象之间可以建立这样的 1-1 对应,使得对应的对象之间恒有同样的关系,则称 M_1 和 M_2 是同构的模型。利用同构的概念,我们也可以这样来理解公理的完备性,即若一公理体系 Σ 的所有模型都是同构的,则称 Σ 是完备的。

仿射几何

【仿射几何】射影空间上凡保持某一给定直线不变的直射多控叫做关于该直线的仿射变换。仿射几何就是研究图形在仿射变换下保持不变的性质。这些仿射变换的全体对变换的乘法形成一个群,叫仿射变换群。仿射几何中的主要结果是:仿射变换保持直线的平行性不变;保持平行线段的长度比不变;保持三个点的共线不变。

【平行射影与仿射变换】设有相交的两个平面 π 与 π' , 它们的交线为 l , o 与 o' 分别是 π 与 π' 上两点, 通过 π 内各点 $A, B, C, \dots$ 引平行于 oo' 的直线交 π' 于 $A', B', C', \dots$, 这样平面 π 内的点与 π' 内的点

与 π' 内的点建立一种一一对应, 这种对应叫做 π 到 π' 的平行射影或叫做透视仿射对应, 显然它把 π 上直线变成 π' 内的直线。

设有 $n+1$ 个平面 $\pi, \alpha_1, \dots, \alpha_{n-1}, \pi'$ 。如果在平面偶 $(\pi, \alpha_1), (\alpha_1, \alpha_2), \dots, (\alpha_{n-1}, \pi')$ 之间都存在着平行射影, 这样平面 π 与 π' 的点之间就建立了一种一一对应, 这种对应叫做平面 π 到 π' 的仿射对应。如果平面 π 与 π' 重合, 则 π 与 π' 之间的仿射对应叫做平面 π 上的仿射变换。

一个笛卡儿直角坐标系在仿射对应下的象叫做仿射坐标系。显然仿射坐标系中的两个坐标轴, 即 X 轴和 y 轴不一定是互相垂直的, 且 X 轴和 y 轴上的度量单位也不一定相等。

设 π 上的一个仿射坐标系 $o-Xy$, 那么在 π 上的一个仿射变换可以用下述代数式来表示

$$\begin{cases} X' = a_{11}X + a_{12}y + a_{13} \\ y' = a_{21}X + a_{22}y + a_{23} \end{cases} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$$

其中点 $P(X, y)$ 经过仿射变换后的象为 $P'(X', y')$ 。显然平面上的平移, 旋转变换在直角坐标系下的表达式为

$$\begin{cases} X' = X \cos \theta - y \sin \theta + a \\ y' = X \sin \theta + y \cos \theta + b \end{cases}$$

平面上的相似变换表达式为

$$\begin{cases} X' = kX + a \\ y' = ky + b \end{cases}$$

这些均是仿射变换的特殊情形。

【单比】 设 P_1, P_2 是有向直线上两定点, P 为此直线上任意一点, 把量 $\frac{P_1P}{P_2P}$ (这里 P_1P, P_2P 均为有向线段) 叫做三点 P_1, P_2, P 的单比或简比。记为 (P_1, P_2, P) , 即

$$(P_1, P_2, P) = \frac{P_1P}{P_2P}$$

其中 P_1, P_2 称为基点, P 叫做分点。

单比是仿射变换下的基本不变量, 即一条直线上三点 P_1, P_2, P 经过仿射变换后的象为 P'_1, P'_2, P' , 则 $(P_1, P_2, P) = (P'_1, P'_2, P')$ 。

射影几何

【射影几何】研究图形的射影性质,即图形经过中心射影不变的性质。在经典几何中,射影几何处于一种特殊的地位,通过它可以把它一些几何联系起来,如仿射几何学可以看成射影几何学的子几何。利用射影度量的概念可以把欧氏几何,双曲非欧几何,椭圆非欧几何统一起来。射影几何的主要不变量是交比。

【中心射影与无穷远元素】设 l 与 l' 是同一平面内的两条不同的直线, O 在此平面内,但不在 l 也不在 l' 上的点。 P 是 l 上任一点,若 OP 与 l' 相交,则交点 P' 叫做 P 从 O 投射到 l' 上的中心射影, O 叫做投影中心。

若 l 上一点 V 与 O 连线 OV 与 l' 平行,则 v 的象 v' 不存在,即 v 无中心射影。同样在 l' 上有一点 w' 不存在 l 上的点中心射影,因此中心射影不是直线 l 到 l' 上点之间的一一对应。我们把 v 点叫做直线 l 上的影消点。

设 π 与 π' 是两个不同的平面, o 是不在此两平面上的点, p 是 π 内任一点, p 和 o 的连线 op 交平面 π' 于 p' 点,则称 p' 叫做 p 在 π' 内的中心投影, o 叫做投影中心。不难看出,在中心投影下平面 π 内的直线对应 π' 内的一条直线。

通过 o 与平面 π' 平行的平面与 π 的交线若为 v ,则直线 v 的中心射影不存在,直线 v 叫做 π 上的影消线。因此中心射影不是平面 π 到 π' 上点之间的一一对应。

中心射影不能建立一一对应的原因是两条平行直线没有交点,为了建立一一对应。可以设想在每条直线上引进一个新点(或称理想点)且规定:

(1)平面内任何一组平行直线引入唯一个点叫做无穷远点。换言之,两平行直线相交于无穷远点。

(2)一平面内所有无穷远点的集合是一条直线,叫做无穷远直线。

由上述规定可以看出空间中一组平行平面相交于无穷远直线。

显然,引入无穷远点和无穷远直线后,直线 l 到 l' 的中心射影,平面 π 到 π' 的中心射影都是一一对应了。

【笛沙格定理】平面内不共线三点与其每两点连线所组成的图形叫做三点形。平面内不共点三直线与其每两直线交点所组成图形叫做三线形,笛沙格定理是指:

如果两个三点形对应顶点的连线交于一点,则对应边的交点必共线。

笛沙格的逆定理也成立,即:如果两个三点形对应边的交点在一直线上,则对应顶点的连线交于一点。

如果两个三点形对应边的交点共线,则所在直线叫做它们的透视轴,如果两个三点形对应顶点连线共点,则公共点叫做它们的透视中心。笛沙格定理与逆定理说明:两个三点形如果有透视中心,则必有透视轴。反之,如果有透视轴,则必有透视中心。

【齐次坐标】设 p 点的笛卡儿坐标为 (X, y) , 则称满足 $\frac{X_1}{X_3} = X, \frac{X_2}{X_3} = y$ 的三个序数 X_1, X_2, X_3 (其中 $X_3 \neq 0$) 叫做点 p 的齐次坐标, 记为 $p(X_1, X_2, X_3)$ 由此可知, 平面内一点的齐次坐标有无穷多组。 $(\rho X_1, \rho X_2, \rho X_3)$ 其中 $\rho \neq 0$ 与 (X_1, X_2, X_3) 都是同一点的齐次坐标。为了区别起见把 (X, y) 叫做点 p 的非齐次坐标。

设直线 l 的方程为 $y = \lambda X + b, \lambda = \tan \theta$ 为直线的斜率, θ 为直线

与 X 轴夹角。如果把直线方程中的 b 看作参数, λ 看作常数(此时 θ 也就是常数) b 变动时, 此方程便表示一组平行直线。因为直线 l 一点 p 的非齐次坐标为 $(X, \lambda X + b)$ 故其齐次坐标可取为 $(X_1, \lambda X + b, 1)$ 或 $(1, \lambda + \frac{b}{X}, \frac{1}{X})$, 当 p 点沿直线 l 两个方向趋于无穷远时, 即当 $X \rightarrow +\infty$ 或 $X \rightarrow -\infty$ 时, 点 p 的齐次坐标的极限为 $(1, \lambda, 0)$ 。这是与 b 无关的一组数, 因此可以规定由 λ 决定方向的无穷远点的齐次坐标为 $(\rho, \rho\lambda, 0) (\rho \neq 0)$ 即:

任何三个有序实数 $(X_1, X_2, 0)$ 其中 $(X_1 \neq 0) \frac{X_2}{X_1} = \lambda$, 规定为以 λ 为斜率所决定方向的无穷远点的齐次坐标, $(0, X_2, 0) (X_2 \neq 0)$ 或 $(0, 1, 0)$ 规定为 y 轴方向的无穷远点的齐次坐标。

注意, 在齐次坐标中没有以 $(0, 0, 0)$ 为坐标的点。

根据齐次坐标的定义, 不难看出设有一直线 l , 在笛卡儿坐标系下的方程为 $aX + by + c = 0$ 则此直线方程可写为 $aX_1 + bX_2 + cX_3 = 0$, 这是关于 X_1, X_2, X_3 的一个齐次方程, 故我们把点的坐标 (X_1, X_2, X_3) 称为齐次坐标的原因。另外无穷远点组成的直线即无穷远线的方程为 $X_3 = 0$ 。

【对偶原则】在射影平面内, 我们把“点”和“直线”叫做对偶元素, “通过一点作一直线”与“在一直线上取一点”叫做对偶运算。设有一个“点与直线”组成的图形 F , 将此图形中的各元素改为它的对偶元素, 各运算改为它的对偶运算, 其结果形成另一个由“直线与点”组成的图形 F' , 则这两个图形叫做对偶图形。例如: 点列与线束是对偶图形; 三点形 F 的对偶图形是三线形 F' , 因为三点形与三线形是相同的图形(通常称为三角形)即 $F = F'$, 此时把这种图形称为自对偶图形。

在射影平面里,设有点、直线所构成的一个命题(此命题中只涉及点、直线之间的结合关系)将此命题中的各元素改为它的对偶元素,各运算改为它的对偶运算,其结果形成另一个命题,则这两个命题叫做对偶命题。

根据射影几何的公理体系可以证明:

在射影平面内,如果一个命题成立,则其对偶命题必然成立,这叫做对偶原则。

例如:笛沙格定理,由于三点形与三线形是自对偶图形,笛沙格定理的对偶命题就是它的逆定理,因此根据对偶原则,我们只要证明其一,另一个必然成立。

【交比与调和比】设 p_1, p_2, p_3, p_4 是四个共线点,它们的交比定义为两个单比 (p_1, p_2, p_3) 与 (p_1, p_2, p_4) 之比,记为 $(p_1 p_2, p_3 p_4)$ 即

$$(p_1 p_2, p_3 p_4) = \frac{(p_1, p_2, p_3)}{(p_1, p_2, p_4)}$$

其中 p_1, p_2 叫做基点偶, p_3, p_4 叫做分点偶。

如果 $(p_1 p_2, p_3 p_4) = -1$, 则称点偶 p_1, p_2 和 p_3, p_4 成调和分离(或成调和共轭), 也称 p_4 为 p_1, p_2, p_3 的第四调和点, 交比值 -1 叫做调和比。

交比有如下性质:

$$(1) (p_1 p_2, p_3 p_4) = (p_1 p_2, p_3 p_4) = \frac{1}{(p_1 p_2, p_3 p_4)}$$

$$(2) (p_2 p_1, p_4 p_3) = (p_1 p_2, p_3 p_4)$$

$$(3) (p_1 p_3, p_2 p_4) = (p_4 p_2, p_3 p_1) = 1 - (p_1 p_2, p_3 p_4)$$

共线四点 p_1, p_2, p_3, p_4 可构成 $4! = 24$ 种排列, 因此有 24 个交比, 设 $(p_1 p_2, p_3 p_4) = r$, 根据上述三个性质可知, 这 24 个交比可

分成六类,即 $r, \frac{1}{r}, 1-r, 1-\frac{1}{r}, \frac{1}{1-r}, \frac{r}{1-r}$ 。

根据调和比的定义可以知道:若 $(p_1p_2, p_3p_4) = -1$ 即 p_1, p_2 与 p_3, p_4 成调和共轭,当且仅当 p_3 是线段 p_1p_2 的中点时, p_4 是直线上的无穷远点。

【二阶曲线】设有点集,它的齐次点坐标为 (X_1, X_2, X_3) 满足下列三元二次方程

$$\sum_{i,j=1}^3 a_{ij} X_i X_j = 0 \quad (a_{ij} = a_{ji})$$

其中 a_{ij} 是实数,且至少有一个不为零,则此点集叫做二阶曲线。系数行列式 $|a_{ij}| \neq 0$ 的二阶曲线称为常态的二阶曲线, $|a_{ij}| = 0$ 的二阶曲线称为变态二阶曲线(或退化的二阶曲线)。二阶曲线的方程是二次(齐次)方程。因此也称为二次曲线。

在一条二阶曲线上任取相异两点 A, B , 令 X 为二阶曲线上的动点,于是得到两个线束,可以证明过 A 点的线束与过 B 点的线束之间成射影对应。反之,任意给定两个线束,如果它们之间可以建立一个射影对应,则其对应直线交点的集合必为一条二阶曲线,这个定理称为旋泰纳(*Steiner*)定理,有了这个结果,可以对二阶曲线下一个射影定义:

在射影平面内,成射影对应的两个线束的对应直线交点的集合叫做二阶曲线。

【巴斯加定理】如果一个 n 点形的顶点都在一个非退化二阶曲线上,则叫做该二阶曲线的内接 n 点形,巴斯加定理是指:内接于一个非退化二阶曲线的简单六点形的三对对边的交点在同一直线上,这条直线称为巴斯加直线,这个定理称为巴斯加定理。

巴斯加定理的逆定理也成立,即若简单六点形的三对对边交点在一直线上,则此六点形内接于一个二阶曲线。

【配极原则】给定一条二阶曲线 $c: \sum_{i,j=1}^3 a_{ij}X_iX_j=0$ ($a_{ij}=a_{ji}$) 以及不在 c 上一点 p ,过 p 点作直线与二阶曲线交于 M_1, M_2 两点(可以是虚交点)如果 Q 是直线 PM_1 上一点,且使交比 $(M_1M_2, PQ)=-1$,则称点 P 与 Q 关于二阶曲线 c 成调和共轭,或称 Q 与 P 关于 c 互为共轭点。可以证明定点 P 关于二阶曲线 c 的调和共轭点的轨迹是一条直线,此直线称为 P 点关于二阶曲线 c 的极线,而 P 点叫做该直线的极点。因此对于一条常态的二阶曲线来说,平面上极点与极线形成一一对应。关于一条二阶曲线 c 的极点与极线之间有如下关系:

如果 P 点关于二阶曲线 c 的极线通过 Q ,则 Q 点关于 c 的极线必通过 P 点。

上述定理就称为配极原则,由配极原则立刻可知:两点连线的极点是此二点极线的交点,两直线交点的极线是此二直线极点的连线。共线点的极线必共点,共点线的极点必共线。

【变换群与几何学】用公理法可以建立起各种几何学,这可以说是用静的观点来研究几何学,我们也可以用变换群来研究所对应的几何学,这也可以说是用动的观点来研究几何学,我们以欧氏几何为例。

欧氏几何所研究图形 F 的性质是指在不同地点与 F 合同的一切图形所共同具有的性质,这就是说,合同的图形本质并无区别。现在我们改用下述观点来说明这个事实。

欧氏空间中正交变换构成群,也就是说,

- (1)恒等变换是正交变换。
- (2)正交变换的逆变换仍是正交变换。
- (3)两个正交变换的乘积仍是正交变换。

于是,我们可以用正交变换来建立起合同的概念,即从一个图形经过正交变换所得到的另一个图形二者是合同的,因此可以用正交变换使在不同地点的图形合同,这样由上述正交变换的三个性质可以推出合同具有下列三个性质:

- (1)反身性,图形 F 与其本身合同,即 $F \cong F$ 。
- (2)对称性,如果 F 合同于图形 F' ,则图形 F' 必合同于 F ,即若 $F \cong F'$,则 $F' \cong F$ 。
- (3)传递性,如果图形 F 合同于图形 F' , F' 又合同于图形 F'' ,则 F 必合同于 F'' ,即

若 $F \cong F', F' \cong F''$,则 $F \cong F''$ 。

因此,合同关系是一个等价关系,它可以把欧氏空间中所有图形分类:凡是合同的图形则属于同一等价类。欧氏几何既然是研究不同地点互相合同图形的性质,也可以说它是研究等价类里一切图形所共有的性质。于是同一类里一切图形所共有的性质必是正交群下的不变性质。反之,倒过来说也可以。因此关于正交群不变性质所构成的命题系统就是欧氏几何学。

将上述讨论加以抽象化。设给出任意元素集合 S 和它的一个变换群 G ,对于 S 中两个子集 A 和 B ,如果在群 G 中有一个变换将 A 变为 B ,则称 A 与 B 等价,可以证明这样规定的等价概念有以下性质:

- (1)任何子集 A 必与自己等价;
- (2)若子集 A 与子集 B 等价,则子集 B 也将与子集 A 等价;

(3)若子集 A 与子集 B 等价,子集 B 与子集 C 等价,则子集 A 将与子集 C 等价。

因此,上述规定的等价概念是一个等价关系,所以它可以决定集合 S 的一个分类方法,亦即将集合 S 分成若干个子集,使得 S 的每一个元素属于而且只属于一个子集。我们把集合 S 的分类方法看作 S 的一种结构,把 S 叫做空间,它的元素叫做点,它的子集叫做图形,于是关于这变换群 G 的不变性质可以作为同一等价类里一切图形所共有的几何性质。因此我们可以用变换群观点去研究相应的几何学。

积分几何

【积分几何】通过各种积分考察图形性质的一门学科。积分几何本质上属于整体微分几何的范畴。它起源于几何概率的研究,其发展也始终与几何概率联系在一起。几何概率的研究是以有关图形的测度为基础,其中的度量往往表为积分,因此自然就导致了积分几何的建立。最早的几何概率是布丰(*Buffon*)提出的投针问题:设在平面上有一组平行线,其行距都等于 D ,把一根长 $l < D$ 的针随机地投上去,则这根针与一条直线相交的概率为 $\frac{2l}{\pi D}$ 。后来英国数学家克劳夫顿(*W. Crofton*)对这类问题作了深入的研究,找到了这些积分间的许多巧妙的关系。法国数学家庞加莱(*H. Poincare*)高瞻远瞩引进了运动密度的概念,把积分几何与变换群相结合。1935年德国数学家布拉施克(*W. Blaschke*)等人在“积分几何”这个总

标题下发表了一系列的论文,积分几何就开始作为几何的一个独立分支得到了系统而深入的发展。以后,陈省身教授与韦伊(*A. Weil*)建立了齐性空间的积分几何理论,为积分几何作出了卓越的贡献。50年代以来西班牙-阿根廷数学家桑塔洛(*L. A. Santalo*)教授对积分几何的研究一直处于领先的地位,他所著的“积分几何与几何概率”是迄今为止这方面最完备的专著。

由于积分几何是和概率论及数理统计紧密联系着的。因此它在许多学科,如生物学、医学、矿物学,以至物理学、天文学等中均有广泛的应用,随着电子计算机的发展,这种应用正方兴未艾。

曲线与曲面 的微分几何

【曲线与曲面的微分几何学】经典微分几何学。主要研究三维欧氏空间 E_3 中的曲线和曲面在一点邻近的性质。经典微分几何学差不多与微积分学同时起源于17世纪。单变量函数的几何形象就是一条曲线,而函数的导数就是曲线的切线之斜率,函数的积分在几何上则可解释为一曲线下的面积等等。这种把微积分应用于曲线、曲面的研究,实质上就是微分几何学的开端。欧拉、蒙日、拉格朗日和哥西等著名数学家都曾为微分几何学的发展做出过重要贡献。但使微分几何独立成为一门学科并使之系统化的当归功于德国数学家高斯。

按照克莱因变换群几何的观点,经典微分几何学应属于运动

群几何。所以也称运动几何学或初等微分几何学。

经典微分几何学的研究工具大部分是微积分学。而力学、物理学与天文学以及技术和工业的日益增长的要求则是微分几何学发展的重要因素。尽管经典微分几何学主要研究曲线、曲面的局部性质,但它形成了现代微分几何学的基础则是毋庸置疑的。

【曲线】区间 $[a, b]$ 到 R^3 的单值连续映射的象称为 R^3 中的曲线。

设 $\{0, X^1, X^2, X^3\}$ 为 R^3 中的笛卡尔直角坐标系,则 R^3 中的曲线 Γ 定义为由 $\vec{r} = \vec{r}(t) = \{X^1(t), X^2(t), X^3(t)\}, a \leq t \leq b$ 所表示的点集。 $X^i(t), (i=1, 2, 3)$ 为连续函数。其中 $\vec{r}(t)$ 表示曲线上点的位置矢量。

曲线上满足 $X^i(t_1) = X^i(t_2), t_1 \neq t_2 (i=1, 2, 3)$ 的点称为重点。无重点的曲线叫简单曲线。若曲线的两端点重合,除端点外再无其他重点的曲线称为简单闭曲线。

坐标函数 $X^i(t) \in C^k (i=1, 2, 3)$,则称曲线为 C^k 类的。 $k \geq 1$ 的曲线是微分几何研究的对象。

曲线上使得 $\frac{d\vec{r}}{dt} \neq \vec{0}$ 的点称为正则点,每点都是正则点的曲线叫正则曲线。

设 $\Gamma: \vec{r} = \vec{r}(t), a \leq t \leq b$,为连续曲线。它的端点分别为 A, B 。在 A, B 之间任取 $n-1$ 个点: $p_1, p_2, \dots, p_{n-1}$,为方便计把 A 写成 p_0, B 写成 p_n ,它们将 Γ 分成 n 段。设各点对应的参数依次为 $a = t_0, t_1, \dots, t_{n-1}, t_n = b$ 。用直线段连结相邻的点,得到一折线形,它的长:

$$\sigma_n = \sum_{v=1}^n \overline{P_{v-1}P_v}$$

当分点无限增加时,若 σ_n 趋于一个与分法的选择无关的确定极

限。则此极限称为曲线段 $\widehat{P_0P_n}$ 的弧长。此时曲线 Γ 为可求长的。

曲线 $\vec{r}=\vec{r}(t)=\{X^1(t), X^2(t), X^3(t)\}$ 有长度的重要条件是坐标函数

$X^i(t)$ 为有界变差函数。特别微分几何中考虑的曲线都有长度,其长度:

$$s(t)=\int_{t_0}^t |\dot{\vec{r}}(t)| dt$$

弧长称为曲线的自然参数,是曲线在刚体运动下的不变量。在取自然参数时, $\Gamma: \vec{r}=\vec{r}(s)$,则 $|\dot{\vec{r}}(s)|=1$ ($\dot{\vec{r}}(s)$ 表示对弧长的导矢),反之,若 $|\vec{r}'(t)|=1$,则 t 可视为曲线从某点量起的弧长。

设 $\Gamma: \vec{r}=\vec{r}(s)$, s 为弧长, Γ 为 C^3 类正则曲线。 $\vec{\alpha}=\dot{\vec{r}}(s)$ 为单位切向矢。 $\vec{r}=\vec{O}$ 的点称为曲线上的逗留点。在非逗留点处,单位矢量 $\vec{\beta}=\frac{\ddot{\vec{r}}}{|\ddot{\vec{r}}|}$ 为主法矢。在弧立的逗留点处,我们取曲线在该点邻近

处的主法矢的极限(如果存在)作为该点的主法矢。 $\vec{\gamma}=\vec{\alpha}\times\vec{\beta}$ 称为副法矢, $\vec{\alpha}, \vec{\beta}, \vec{\gamma}$ 三单位矢相互垂直,且成右旋系。通常称

$\{\vec{r}(s); \vec{\alpha}(s), \vec{\beta}(s), \vec{\gamma}(s)\}$ 为曲线的弗朗内(Frenet)标架。

过曲线上 $\vec{r}(s_0)$ 点的三条直线: $\vec{\rho}=\vec{r}(s_0)+\lambda\vec{\alpha}(s_0)$, $\vec{\rho}=\vec{r}(s_0)+\lambda\vec{\beta}(s_0)$, $\vec{\rho}=\vec{r}(s_0)+\lambda\vec{\gamma}(s_0)$ 分别叫作在 $\vec{r}(s_0)$ 点的切线,主法线和副法线。过 $\vec{r}(s_0)$ 的三个平面: $\vec{\alpha}(s_0)\cdot(\vec{\rho}-\vec{r}(s_0))=0$, $\vec{\beta}(s_0)\cdot(\vec{\rho}-\vec{r}(s_0))=0$, $\vec{\gamma}(s_0)\cdot(\vec{\rho}-\vec{r}(s_0))=0$ 分别叫作在 $\vec{r}(s_0)$ 点的法平面、从切面和密切面。

在 Γ 的每一非逗留点处都有唯一确定的弗朗内标架。当点沿 Γ 移动时,弗朗内标架作为一个刚体来运动。设 Γ 上两邻近点对应的弧长参数分别为 $s, s+\Delta s$,对应这两点的弗朗内标架依次为 $\{\vec{r}$

$(s); \vec{\alpha}(s), \vec{\beta}(s), \vec{\gamma}(s)\}$ 及 $\{\vec{r}(s+\Delta s); \vec{\alpha}(s+\Delta s), \vec{\beta}(s+\Delta s), \vec{\gamma}(s+\Delta s)\}$, 当 $\Delta s \rightarrow 0$ 时, 两组标架有如下关系:

$$\begin{cases} \vec{\alpha} = k\vec{\beta} \\ \vec{\beta} = -k\vec{\alpha} + \tau\vec{\gamma} \\ \vec{\gamma} = -\tau\vec{\beta} \end{cases}$$

其中 $k(s), \tau(s)$ 分别称作曲线 Γ 的曲率、挠率。这一公式是曲线论的基本公式(或称 *Frenet* 公式)。

【曲率与挠率】在弗朗内公式中, $k(s) = |\ddot{\vec{r}}(s)| = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta \theta}{\Delta s} \right|$, 表示曲线上参数为 $s, s+\Delta s$ 的两邻近点的切向 $\vec{\alpha}(s), \vec{\alpha}(s+\Delta s)$ 之间的夹角 $\Delta \theta$ 与 Δs 之比, 在 $\Delta s \rightarrow 0$ 时的变化情况, 它度量了曲线上两邻近点切向量夹角关于弧长的变化率, 反映了曲线在该点弯曲程度。当 $\kappa(s) \neq 0$ 时, 其倒数 $\rho(s) = \frac{1}{k(s)}$ 称为曲线 $\vec{r}(s)$ 在 s 处的曲率半径。 $\kappa(s) \equiv 0$ 的曲线为直线。由于角度、弧长均是运动不变量, 因此曲率也是运动不变量, 并且与曲线的参数选择无关。在一般参数下, 曲率的公式是

$$\kappa = \frac{|\vec{r}' \times \vec{r}''|}{|\vec{r}'|^3}$$

挠率 $\tau = \vec{\beta} \cdot \dot{\vec{\gamma}} = -\vec{\gamma} \cdot \dot{\vec{\beta}}$, 其绝对值 $|\tau| = |\dot{\vec{\gamma}}| = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta \theta}{\Delta s} \right|$, 此处 $\Delta \theta$ 表示 Γ 上参数分别为 $s, s+\Delta s$ 的两邻近点的副法矢 $\vec{\gamma}(s)$ 与 $\vec{\gamma}(s+\Delta s)$ 之间的夹角。 $\tau(s) > 0$ 表示曲线在 S 处是右旋的, $\tau(s) < 0$, 表示曲线在 s 处是左旋的。当曲线正向颠倒时, $\tau(s)$ 不改变符号。象曲线的曲率一样, 挠率 $\tau(s)$ 也是运动不变量, 且与参数选取无关。 $\tau(s) \equiv 0$ 的曲线为平面曲线。在一般参数下, 挠率计算

公式是

$$\tau = \frac{(\vec{r}', \vec{r}'', \vec{r}''')}{(\vec{r}' \times \vec{r}'')^2}。$$

已知曲线在一点 P_0 的曲率和挠率,就可以大致推知曲线在该点的邻近结构。更一般地,有两条不含逗留点的空间曲线经过运动可重合的充要条件是:在适当地选择自然参数 s 后,二者有相同的曲率和挠率。这一断言,不仅说明曲线的形状取决于它的曲率和挠率,而且曲率、挠率还决定曲线的形状。曲线的一切度量性质和运动不变量都可由 κ, τ 来表达。

对任意给定的一对连续函数

$\kappa(s), \tau(s), s \in [s_0, s_1]$ 其中 $\kappa(s) > 0$, 就一定存在某些曲线, 分别以 $\kappa(s), \tau(s)$ 为曲率、挠率, 以 s 为弧长。这些曲线除在空间的位置不同外, 形状完全相同。通常称它为曲线存在定理。

【平面曲线】假设曲线 Γ 为有向平面 Xy 面上的 C^2 类曲线。选取法向量 $\vec{n}$, 使得切向量 $\vec{\alpha}$ 与法向量 $\vec{n}$ 的定向与普通直角坐标系的 X, y 轴的定向相同。此时弗朗内公式为:

$$\begin{cases} \dot{\vec{\alpha}} = \kappa_r \vec{n} \\ \dot{\vec{n}} = -\kappa_r \vec{\alpha} \end{cases}$$

κ_r 为平面曲线的相对曲率, 它可取正值, 也可取负值。在曲线的拐点处, $\kappa_r = 0$ 。

设 Γ 上一点 P 处的切向与 X 轴正向的有向夹角为 φ , 则相对曲率 $\kappa_r = \frac{d\varphi}{ds}$ 。曲线“向左转”的地方, $\kappa_r > 0$, “向右转”的地方 $\kappa_r < 0$ 。相对曲率是曲线的运动不变量。当取自然参数时, $\kappa_r = \ddot{X}\dot{y} - \ddot{Y}\dot{X}$, 在一般参数下,

$\kappa_r = \frac{X' y'' - X'' y'}{(X'^2 + y'^2)^{3/2}}$, 特别曲线由显函数 $y = y(X)$ 给出时, $K_r = \frac{d^2 y}{dX^2} / [1 + (\frac{dy}{dX})^2]^{3/2}$. $\rho = 1/|K_r|$ 称为曲线的曲率半径。以 $\vec{r}(s) + \vec{\rho n}(s)$ 为中心, ρ 为半径的圆称为 $\vec{r}(s)$ 处的密切圆, 它与曲线的切触阶最高。密切圆的圆心叫曲率中心。曲线 Γ 的曲率中心轨迹 Γ' 叫 Γ 的渐屈线, 而 Γ 称为 Γ' 的渐伸线。

【曲面】 设 D 是 R^2 的一个区域, (u, v) 表示 D 中点的坐标。 D 到 R^3 的一可微映射:

$$(u, v) \rightarrow \{X(u, v), y(u, v), z(u, v)\}, (u, v) \in D$$

此映射下, D 的像集构成了 R^3 中的一张曲面 Σ 。

(u, v) 称为曲面 Σ 上的点的曲线坐标或参数。固定 $v = v_0$, 得到 Σ 上一条 u 曲线, 固定 $u = u_0$ 得到 Σ 上一条 v 曲线, 它们统称为参数曲线。

设 $P_0(u_0, v_0) \in \Sigma$, 若在 P_0 点, $\vec{r}_u|_{u=u_0, v=v_0} \times \vec{r}_v|_{u=u_0, v=v_0} \neq \vec{0}$ 就称 P_0 为 Σ 上的正则点, 由正则点构成的曲面为正则曲面。在正则曲面的每点处, 由 $\vec{r}_u, \vec{r}_v$ 作成一组基的二维线性空间称为该点的切面, 其单位法向用 $\vec{n} = \vec{r}_u \times \vec{r}_v / |\vec{r}_u \times \vec{r}_v|$ 表示。

当参数变换时, $u = u(\bar{u}, \bar{v}), v = v(\bar{u}, \bar{v}), \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} \neq 0$, 有 $\vec{r}_{\bar{u}} \times \vec{r}_{\bar{v}} = \vec{r}_u \times \vec{r}_v \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})}$ 。因此, 当参数变换时, 曲面的正则点仍变为正则点。

【曲面的第一基本形式】 设 $\Sigma: \vec{r} = \vec{r}(u, v), (u, v) \in D$ 是 C^1 类正则曲面。 $P(u, v)$ 及 $P'(u + \Delta u, v + \Delta v)$ 是 Σ 上两邻近点, 其位置矢

量分别为 $\vec{r}(u, v), \vec{r}(u + \Delta u, v + \Delta v)$, 由矢函数的泰勒展开式, $P\vec{P}' = \Delta \vec{r} = \vec{r}(u + \Delta u, v + \Delta v) - \vec{r}(u, v) = \vec{r}_u \Delta u + \vec{r}_v \Delta v + \dots$ 。其中省略的部分是关于 $\Delta u, \Delta v$ 的二阶和二阶以上的小量。当 P' 点无限接近 P 点时 (即 $\Delta u \rightarrow 0, \Delta v \rightarrow 0$) 把 $P\vec{P}'$ 在 R^3 中的长度的主要部分 ds 定义为曲面 Σ 上这两个无限接近的点之间的距离, 即

$$\begin{aligned}\varphi_1 = ds^2 &= |\vec{dr}|^2 = (\vec{r}_u du + \vec{r}_v dv)^2 = \vec{r}_u^2 du^2 + 2\vec{r}_u \cdot \vec{r}_v dudv + \vec{r}_v^2 dv^2 \\ &= Edu^2 + 2Fdudv + Gdv^2\end{aligned}$$

其中 $E = \vec{r}_u^2, F = \vec{r}_u \cdot \vec{r}_v, G = \vec{r}_v^2$ 且 $E > 0, G > 0, EG - F^2 > 0$ 。称为曲面 $\vec{r} =$

$\vec{r}(u, v)$ 的第一基本形式, 或称为曲面的线素。它是关于 du, dv 的正定二次形式。第一基本形式 φ_1 决定曲面的内在性质, 因此只要给出 φ_1 , 无须知道曲面的方程或形式, 就可以计算内在几何量, 诸如曲面上的曲线的弧长及测地曲率, 两曲线的夹角以及一块曲面的面积等。

设曲面 $\Sigma: \vec{r} = \vec{r}(u, v)$ 上的曲线 Γ 的方程: $\vec{r} = \vec{r}(u(t), v(t)), a \leq t \leq b$ 。则该段曲线的弧长为:

$$L = \int_a^b \sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\frac{dudv}{dt dt} + G\left(\frac{dv}{dt}\right)^2} dt$$

若 Γ_1, Γ_2 为 Σ 上过 P 点的两条曲线, $du, dv; \delta u, \delta v$ 分别为沿 Γ_1, Γ_2 的微分, Γ_1, Γ_2 在 P 点的交角为 φ , 则

$$\cos \varphi = [Edu\delta u + F(du\delta v + dv\delta u) + Gdv\delta v] / [\pm \sqrt{Edu^2 + 2Fdudv + Gdv^2} \sqrt{E\delta u^2 + 2F\delta u\delta v + G\delta v^2}]$$

曲面 $\Sigma: \vec{r} = \vec{r}(u, v), (u, v) \in D$ 的面积用 σ 表示, 则

$$\sigma = \iint_D \sqrt{EG - F^2} dudv.$$

如果两曲面 $\Sigma: \vec{r} = \vec{r}(u, v)$,
 $(u, v) \in D, \Sigma^*: \vec{r}^* = \vec{r}^*(u^*, v^*)$,
 $(u^*, v^*) \in D^*$ 之间存在一一对应:

$$\begin{cases} u^* = u^*(u, v) \\ v^* = v^*(u, v) \end{cases} \quad \frac{\partial(u^*, v^*)}{\partial(u, v)} \neq 0,$$

在此对应下, 它们的第一基本形式相等: $ds^{*2} = ds^2$ 则称这两曲面是等距的。例如悬链面和正螺面彼此等距。等距变换下不变的几何量称为曲面的等距不变量, 等距变换下不变的有关曲面的几何性质, 称为曲面的内在(内蕴)性质。

假若两曲面之间存在一一对应, 在此对应下它们的第一基本形式成比例: $ds^{*2} = \rho^2 ds^2$, ($\rho \neq 0$) 则称这两曲面为共形的。两共形曲面, 其对应的角度保持不变, 因此共形对应又称保角对应。特别任何 C^3 类正则曲面(在局部范围内)必与平面共形, 即存在参数 u, v , 使得 $\varphi_1 = \rho^2(du^2 + dv^2)$, 此时 u, v 称为等温参数。人们绘制麦卡托地图的方法就利用了球面与平面共形的原理。

【曲面的第二基本形式】 设 $\Sigma: \vec{r} = \vec{r}(u, v)$ 是 C^2 类正则曲面。 $p(u, v)$,

$p'(u + \Delta u, v + \Delta v)$ 是 Σ 上两邻近点, π 为 Σ 在 P 点的切平面, 应用泰勒公式有:

$$\begin{aligned} P\vec{P}' &= \vec{r}(u + \Delta u, v + \Delta v) - \vec{r}(u, v) \\ 0 &= \vec{r}_u \Delta u + \vec{r}_v \Delta v + \frac{1}{2} [\vec{r}_{uu} (\Delta u)^2 + 2\vec{r}_{uv} (\Delta u \Delta v) + \vec{r}_{vv} (\Delta v)^2] + \dots \end{aligned}$$

略去的部分为 $\Delta u, \Delta v$ 的三阶及三阶以上小量。

设 P^1 到平面 π 的有向距离为 δ , 则

$$2\delta = P\vec{P}' \cdot \vec{n} = \vec{r}_{uu} \Delta u^2 + 2\vec{r}_{uv} \Delta u \Delta v + \vec{r}_{vv} \Delta v^2 + \dots \cdot \vec{n} \text{ 当 } P' \text{ 无限}$$

接近 P 时,称其主要部分为曲面 Σ 的第二基形式 $\varphi_2 = Ldu^2 + 2Mdudv + Ndvd^2$ 。其中,

$$L = \vec{r}_{uu} \cdot \vec{n} = -\vec{r}_u \cdot \vec{n}_u = (\vec{r}_u, \vec{r}_v, \vec{r}_{uu}) / |\vec{r}_u \times \vec{r}_v|$$

$$M = \vec{r}_{uv} \cdot \vec{n} = -\vec{r}_u \cdot \vec{n}_v = -\vec{r}_v \cdot \vec{n}_u = (\vec{r}_u, \vec{r}_v, \vec{r}_{uv}) / |\vec{r}_u \times \vec{r}_v|$$

$$N = \vec{r}_{vv} \cdot \vec{n} = -\vec{r}_v \cdot \vec{n}_v = (\vec{r}_u, \vec{r}_v, \vec{r}_{vv}) / |\vec{r}_u \times \vec{r}_v|。因此, \varphi_2 = -\vec{dr} \cdot \vec{dn}。$$

曲面的第二基本形式刻画了曲面离开切平面弯曲程度,即曲面在空间 R^3 中的弯曲程度。第一、二基本形式除了曲面在空间的位置外完全决定了曲面的形状。

【曲面的基本公式及基本方程】设曲面 $\Sigma: \vec{r} = \vec{r}(u^1, u^2)$ 为定义在 D 上的 C^2 类正则曲面。将第一基本形式写成 $\varphi_1 = g_{ij}du^i du^j$, ($g_{ij} = g_{ji}$), 第二基本形式写成 $\varphi_2 = b_{ij}du^i du^j$, ($b_{ij} = b_{ji}$)。($i, j = 1, 2$, 且对指标 i, j 取和)

φ_1 的系数矩阵 (g_{ij}) 之逆矩阵用 (g^{ij}) 表示, $g = \det(g_{ij})$ 。由 φ_1 的系数 g_{ij} 确定的第一、二类克氏(Christoffel)记号分别为:

$$\Gamma_{ikj} = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{ik}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k} \right)$$

$$\Gamma_{ij}^k = g^{kl} \Gamma_{ilj}$$

它们都关于 i, j 对称。

$$\text{设 } \vec{r}_i = \frac{\partial \vec{r}}{\partial u^i}, \vec{r}_{ij} = \frac{\partial^2 \vec{r}}{\partial u^i \partial u^j}$$

在 Σ 的每点处,存在由 $\vec{r}_1, \vec{r}_2, \vec{n}$ 构成的基标架,则

$$\begin{cases} \vec{r}_{ij} = \Gamma_{ij}^k \vec{r}_k + b_{ij} \vec{n} \\ \vec{n}_i = -b_{il} g^{lk} \vec{r}_k \end{cases}$$

分别称为 *Gauss*(高斯)公式及 *Weingarten*(魏加吞)公式。视它们为偏微分方程组,其可积条件(实际上仅是 *Gauss* 公式的可积条件)为:

$$\begin{cases} \frac{\partial \Gamma_{ij}^l}{\partial u_k} - \frac{\partial \Gamma_{ik}^l}{\partial u_j} + (\Gamma_{ij}^m \Gamma_{mk}^l - \Gamma_{ik}^m \Gamma_{mj}^l) \\ \quad + b_{ik} g^{lm} b_{mj} - b_{ij} g^{lm} b_{mk} = 0 \\ \frac{\partial b_{ij}}{\partial u^k} - \frac{\partial b_{ik}}{\partial u^j} + \Gamma_{ij}^m b_{mk} - \Gamma_{ik}^m b_{mj} = 0 \end{cases}$$

前者称为高斯方程,后者称为柯达齐方程。并分别称

$$\begin{aligned} R_{ijk}^l &= \frac{\partial \Gamma_{ik}^l}{\partial u^j} - \frac{\partial \Gamma_{ij}^l}{\partial u^k} + \Gamma_{ik}^m \Gamma_{mj}^l - \Gamma_{ij}^m \Gamma_{mk}^l \\ R_{hijk} &= g_{hl} R_{ijk}^l = \frac{\partial \Gamma_{ikh}}{\partial u^j} - \frac{\partial \Gamma_{ijh}}{\partial u^k} \\ &\quad + \Gamma_{ij}^m \Gamma_{hkm} - \Gamma_{ik}^m \Gamma_{hjm} \end{aligned}$$

为曲面的第二、第一黎曼曲率张量。

若 D 为 R^2 上的单连通区域,在 D 上定义了 C^2 类正定对称矩阵 (g_{ij}) ,及 C^1 类对称矩阵 (b_{ij}) ,且六个二元函数 $g_{ij}, b_{ij} (i, j=1, 2)$, 满足高斯-柯达齐方程,则存在以 g_{ij} 为第一基本量, b_{ij} 为第二基本量的曲面。如果在 R^3 中给定初始点以及初始基标架,则这样的曲面片是唯一的。

【法曲率】设 $\Gamma: \vec{r} = \vec{r}(u(s), v(s))$ 是曲面 Σ 上过 P 点的曲线。 s 为其弧长。 Γ 的曲率矢 $K\vec{\beta}$ 在曲面 Σ 的单位法向 $\vec{n}$ 上的投影 $K\vec{\beta} \cdot \vec{n} = \frac{\varphi_2}{\varphi_1} = Ldu^2 + 2Mdudv + Ndv^2 / Edu^2 + 2Fdudv + Gdv^2$, 通常记作 $K_n = K \cos \theta$ 。称为曲面在 P 点沿 Γ 之切向 $du:dv$ 的法曲率。此处 θ 是曲线 Γ 的主法矢 $\vec{\beta}$ 与曲面法向 $\vec{n}$ 之间的夹角。

在 P 点取一固定的非渐近方向 (亦即 $\varphi_2 \neq 0$ 的方向), 选择曲面法向 $\vec{n}$, 使 $\varphi_2 > 0$, 经过所取的固定方向及法向 $\vec{n}$ 作一平面, 该平面与曲面 Σ 的交线 Γ 称为曲面的一条法截线。对于法截线 Γ , 有 $\vec{\beta} = \vec{n}$, $K_n = \bar{K} \vec{\beta} \cdot \vec{n}$ 。此处 $\vec{\beta}$, $\bar{K}$, 分别是法截线 Γ 的主法矢, 曲率。由法曲率定义知 $\bar{K} = K \cos \theta$, 或写成 $R = \bar{R} \cos \theta$ 。其中 $R, \bar{R}$ 分别表示 $\Gamma, \bar{\Gamma}$ 的曲率半径。其几何意义是: 若曲面 Σ 上一条曲线 Γ 在 P 点的方向不是渐近方向, 则 Γ 对应于 P 点的曲率中心 C 是沿方向的法截线 $\bar{\Gamma}$ 的曲率中心 $\bar{C}$ 在 Γ 的密切平面上的投影。这一结果称为默尼埃定理。

【高斯曲率】在曲面上一点 P , 法曲率 K_n 可以看作在 P 点切平面上的单位圆周上的连续函数, 它必存在取最大、最小值的方向:

$$\begin{vmatrix} \text{Edu} + \text{Fdv} & \text{Fdu} + \text{Gdv} \\ \text{Ldu} + \text{Mdv} & \text{Mdu} + \text{Ndv} \end{vmatrix} = 0$$

该方程决定的两个方向称为曲面在 P 点的两个主方向, 沿主方向的法曲率叫主曲率。主曲率满足下面的二次方程:

$$(\text{EG} - \text{F}^2)K_n^2 - (\text{LG} - 2\text{MF} + \text{NE})K_n + \text{LN} - \text{M}^2 = 0$$

在曲面的脐点 (φ_1, φ_2 成比例) 处, 任何方向都是主方向; 在非脐点处总存在两个不同的主曲率 K_1, K_2 以及对应的两个相互垂直的主方向。设 K_1, K_2 是曲面的两个主曲率, 定义:

$$K = K_1 \cdot K_2 = \frac{\text{LN} - \text{M}^2}{\text{EG} - \text{F}^2}$$

$$H = \frac{K_1 + K_2}{2} = \frac{\text{LG} - 2\text{MF} + \text{NE}}{2(\text{EG} - \text{F}^2)}$$

则 K 叫做曲面的高斯曲率, H 叫中曲率。

由 *Gauss* 方程知 $K = \frac{R_{1212}}{g}$, 因此高斯曲率只依赖第一基本形式, 这就是著名的高斯定理。

设 P 为 Σ 上的非脐点, $\vec{e}_1, \vec{e}_2$ 为 P 点处的两个单位主方向。设 $\vec{\alpha}$ 为过 P 点的任意方向, φ 为 $\vec{\alpha}$ 和 $\vec{e}_1$ 夹角(有向或无向), 则有:

$$K_n = K_1 \cos^2 \varphi + K_2 \sin^2 \varphi.$$

其中 K_1, K_2 分别为对应主方向 $\vec{e}_1, \vec{e}_2$ 的主曲率, K_n 为沿 $\vec{\alpha}$ 方向的法曲率。这一公式称为 *Euler*(欧拉)公式。

【曲率线】曲面 Σ 上一条曲线 Γ , 在每点的切线总沿着在该点的一个主方向, 则称 Γ 为 Σ 上的曲率线。曲率线的微分方程是:

$$\begin{vmatrix} Edu + Fdv & Fdu + Gdv \\ Ldu + Mdv & Mdu + Ndv \end{vmatrix} = 0$$

在全脐点曲面(如平面、球面)上, 任何一条曲线都是曲率线。在曲面的非脐点处总有两条相互垂直的曲率线经过。在不含脐点的曲面上, 参数曲线为曲率线的充要条件是 $F = M = 0$ 。就是说, 将曲面的第一、二基本形式同时化为只含平方项的“典型式”。

曲面 Σ 上一条曲线 Γ 为曲率线的充要条件是该曲线的切向 $d\vec{r}$ 平行于 $d\vec{n}$ 这一结果称为罗德里克(*Rodrigues*)方程。由罗德里克方程容易得出曲率线的几何特征: 曲线 Γ 是曲面 Σ 上的曲率线充要条件是: Σ 沿 Γ 的法线构成的曲面为可展曲面。

【特殊曲面】由单参数直线族生成的曲面叫直纹面, 族中直线称为母线。单叶双曲面、双曲抛物面、柱面、锥面和切线曲面等都是直纹面。假若沿每条母线上的点有共同的切平面, 这种直纹面称为可展

曲面。

在直纹面上取一曲线 $\Gamma: \vec{r} = \vec{\rho}(u), u_0 \leq u \leq u_1$, 和所有母线相交, $\vec{\tau}(u)$ 为经过 $\vec{\rho}(u)$ 点的母线上的非零矢量, 则直纹面方程可写成:

$$\vec{r} = \vec{\rho}(u) + v \vec{\tau}(u) \quad u_0 \leq u \leq u_1, -\infty < v < +\infty$$

直纹面为可展的充要条件为: $\vec{\rho}', \vec{\tau}, \vec{\tau}' = 0$

可展曲面只有柱面、锥面和切线曲面。

高斯曲率 $K \equiv 0$ 是可展曲面的另一特征。

柱面、锥面和切线曲面均可与平面建立等距对应(和平面贴合), 反之因高斯曲率是内蕴几何量, 因此 $K \equiv 0$ 的曲面一定能与平面建立等距对应。可与平面贴合是可展曲面的特征之一。

由包络理论, 一个单参数平面族的包络面必为可展曲面, 反之可展曲面总是一个单参平面族的包络面。

中曲率 $H \equiv 0$ 的 C^2 类曲面叫做极小曲面。以知闭曲线为边界, 且有相对最小面积的解析曲面则 $H \equiv 0$ 。若曲面是 C^2 类的且 $H \equiv 0$, 则曲面必为解析的(Plateau 问题)。在仅由双曲点构成的曲面片为极小曲面的充要条件是存在两族直交渐近线网。极小的旋转面只能是悬链面, 极小的直纹面必为正螺面。

$K = \text{const}$ 的曲面称为常曲率曲面。具有同一常数的两常曲率曲面是等距的。伪球面是负常曲率曲面的最简单的例子。

【测地线】 设 $\Sigma: \vec{r} = \vec{r}(u, v)$ 为 C^2 类曲面, Γ 为 Σ 上 C^2 类曲线。 Γ 的曲率为 K , 付法矢 $\vec{\gamma}$ 。称 $K_g = K \vec{\gamma} \cdot \vec{n}$ 为 Γ 在一点的测地曲率。它是内蕴几何量。当曲面上取直交纲时,

$$K_g = \frac{d\varphi}{ds} - \frac{1}{2\sqrt{G}} \frac{\partial \ln E}{\partial v} \cos \varphi + \frac{1}{2\sqrt{E}} \frac{\partial \ln G}{\partial u} \sin \varphi$$

其中 φ 表示 Γ 的切向和 u 线方向的有向夹角(见图)。这一公式称为柳维尔公式。

测地曲率为零的曲线称为测地线。其微分方程为二阶常微分方程组:

$$\frac{d^2 u^k}{ds^2} + \Gamma_{ij}^k \frac{du^i}{ds} \frac{du^j}{ds} = 0$$

$$(u^1 = u, u^2 = v \quad i, j, = 1, 2)$$

这里的 s 为 Γ 的弧长参数或弧长的线性函数。

在正交纲下,测地线微分方程可表为:

$$\begin{cases} \frac{dv}{du} = \sqrt{\frac{E}{G}} \operatorname{tg} \varphi \\ \frac{d\varphi}{du} = \frac{1}{2} \sqrt{\frac{E}{G}} \frac{\partial \ln E}{\partial v} - \frac{1}{2} \frac{\partial \ln G}{\partial u} \operatorname{tg} \varphi \end{cases}$$

在充分小的曲面片上,测地线是自身上任意两点最短连线。而展开为直线,表明了测地线的局部最短性质。

【列维-基维它平移】设曲面 $\Sigma: \vec{r} = \vec{r}(u^1, u^2)$ 上的曲线 $\Gamma: u^1 = u^1(t), u^2 = u^2(t)$ 。沿 Γ 上分布一向量场 $\vec{a}(t) = a^i(t) \vec{r}_i$ 。该向量场关于变量 t 求普通导数:

$$\frac{d\vec{a}}{dt} = \frac{da^i(t)}{dt} \vec{r}_i + a^i(t) \frac{d\vec{r}_i}{dt}$$

由 Gauss 公式知:

$$\frac{d\vec{a}}{dt} = \left[\frac{da^k(t)}{dt} + a^i(t) \Gamma_{ij}^k \frac{du^j}{dt} \right] \vec{r}_k + b_{ij} a^i \frac{du^j}{dt} \vec{n}。$$

将 $\frac{d\vec{a}}{dt}$ 投影到切平面上,得到:

$$\frac{D\vec{a}}{dt} = \left[\frac{da^k}{dt} + \Gamma_{ij}^k a^i \frac{du^j}{dt} \right] \vec{r}_k$$

称 $\frac{D\vec{a}}{dt}$ 为向量场 $\vec{a}$ 沿 Γ 的协变导数, 而 $D\vec{a}$ 为协变微分或绝对微分。

若 $D\vec{a}/dt = \vec{0}$, 亦即

$$\frac{da^k}{dt} + \Gamma_{ij}^k a^i \frac{du^j}{dt} = 0, \text{ 则称向量场沿 } \Gamma \text{ 是平行移动的, 这种平移}$$

为列维-基维它平移。

列维-基维它平移完全由曲面的第一基本形式决定, 因此属曲面内在几何范畴。两个沿 Γ 做列维-基维它意义下平移的向量场, 保持内积不变。当曲面 Σ 是平面时, *levi-Civita* 平移就是通常欧氏意义下的平移。因此前者可以看作是欧氏平移的推广。

由测地线的微分方程知, 测地线的切向总是沿测地线本身平移的。所以测地线是自平移曲线。

一般说来, 列维-基维它平移与曲面上的道路有关, 但在单连通可展曲面上, 平移与道路无关, 这个条件也是必要的。

【高斯-波内公式】考虑 C^2 类曲面 Σ 上一条 C^2 类简单光滑闭曲线 Γ , 它所限定的部分曲面 Σ_1 为单连通的, 则有 $\oint_{\Gamma} K_g ds + \iint_{\Sigma_1} K d\sigma = 2\pi$ 。

其中 K_g 为 Γ 的测地曲率, K 为 Σ 的高斯曲率, $d\sigma$ 为曲面面积元素, s 为 Γ 的弧长参数。上述公式称为曲面关于光滑闭曲线的高斯-波内公式。

若 Γ 是由逐段光滑曲线段连接而成的简单闭曲线, 它限定单连通有界区域 Σ_1 , 设各弧连接点处的外角为 $\alpha_i (i=1, 2, \dots, n)$ 则

$$\oint_{\Gamma} K_g ds + \sum_{i=1}^n \alpha_i + \iint_{\Sigma_1} K d\sigma = 2\pi$$

特别当各曲线为测地线时, 有

$$\sum_{i=1}^n \alpha_i + \iint K d\sigma = 2\pi.$$

这一公式是欧几里德几何学中“三角形内角之和等于 π ”的命题的推广。

当 Σ 为定向闭曲面时,则有

$$\iint_{\Sigma} K d\sigma = 2\pi \chi(\Sigma)$$

其中 $\chi(\Sigma)$ 表示曲面 Σ 的欧拉示性数。上述公式称为闭曲面的高斯-波内公式。公式的左边出现的高斯曲率 K 依赖于曲面的度量,而右边的欧拉示性数 $\chi(\Sigma)$ 却是拓扑不变量。它建立了曲面的局部性质与整体性质的联系。陈省身首先给出这一著名公式在高维流形上的推广的简单、内蕴证明。高斯-波内公式是大范围微分几何学的一个经典结果。

微分流形

【微分流形】一类重要的拓扑空间,具有通常的拓扑结构,也具有所谓的微分结构。它是微分几何学和微分拓扑学研究的对象。

设 M 是豪斯道夫空间, M 的每点 p 都有与 R^n 中的开集同胚的邻域 (U, φ) , $\varphi: U \rightarrow \varphi(U) \subset R^n$ 。称这样的拓扑空间 M 为 n 维拓扑流形。

假若 (U, φ) 是包含 p 点的邻域,设 $\varphi(p)$ 在 R^n 中的坐标为 $(X^1(p), \dots, X^n(p))$, 当点在 U 中变动时, $X^1, \dots, X^n$ 是 U 上的实值连续函数。函数组 $(X^1, \dots, X^n)$ 叫做在坐标邻域 (U, φ) 中 M 的局部坐

标系。 $(X^1(p), \dots, X^n(p))$ 称为 p 点的局部坐标。 $u = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ 叫做 M 的坐标邻域族(或图册)。

设 $(U_\alpha, \varphi_\alpha), (U_\beta, \varphi_\beta)$ 是 M 的坐标邻域族 U 中的两个坐标邻域, $U_\alpha \cap U_\beta \neq \emptyset$, 其局部坐标分别为 $(X^1, \dots, X^n), (y^1, \dots, y^n)$ 。如果同胚映射:

$\varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$ 给出坐标变换: $y^i = h^i(X^1, \dots, X^n) \in C^r, (i=1, \dots, n)$ 。类似地 $\varphi_\alpha \circ \varphi_\beta^{-1}: \varphi_\beta(U_\alpha \cap U_\beta) \rightarrow \varphi_\alpha(U_\alpha \cap U_\beta)$ 给出坐标变换: $X^i = g^i(y^1, \dots, y^n) \in C^r, (i=1, \dots, n)$ 。就说坐标邻域 $(U_\alpha, \varphi_\alpha)$ 与 (U_β, φ_β) 为 C^r 相容。同样地可定义两坐标邻域的 C^∞ 相容, C^ω 相容(实解析相容)性。

设 M 为 n 维拓扑流形, M 上的一个 C^r 微分结构(或 C^r 结构), 指的是具以下性质的坐标邻域族 $U = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$: (i) $\{U_\alpha\}$ 覆盖 M , (ii) $\forall \alpha, \beta \in A$ 则 $(U_\alpha, \varphi_\alpha), (U_\beta, \varphi_\beta)$ 为 C^r 相容。若 $U_\alpha \cap U_\beta = \emptyset$, 则认为 $(U_\alpha, \varphi_\alpha), (U_\beta, \varphi_\beta)$ 是 C^r 相容的。(iii) 任何与 U 中的每一个坐标邻域 $(U_\alpha, \varphi_\alpha)$ 都 C^r 相容的坐标邻域 (U, φ) , 它本身也属于 U 。亦即 U 是满足(i), (ii)的最大坐标邻域族。

具有 C^r 微分结构的 n 维拓扑流形称为 n 维 C^r 类微分流形。当 $r=\omega$ 时称为实解析流形。

存在不具备任何微分结构的拓扑流形。首先给出这种例子的是 $M \cdot Kervaire$ 。另方面米尔诺于 1956 年给出著名的例子: 在七维球面 S^7 上存在 28 个不同的微分结构, 得到的微分流形彼此不微分同胚。因此微分结构有独立于拓扑结构的意义。

如果微分流形 M 是仿紧或紧致拓扑空间, 则称 M 为仿紧或紧致微分流形。假若在微分流形 M 上可选取坐标邻域族使坐标变换时的雅可比行列式都大于零, 则称 M 是可定向的。例如 $S^2 \subset \mathbb{R}^3$ 是 2 维紧致可定向的解析流形, 而莫比乌斯(Möbius)带是 2 维仿

紧不可定向的。

在微分流形上可定义可微函数、切向量、向量场、张量场等概念以及两个微分流形间的可微映射等,并建立相应的分析学。

仿紧微分流形上总存在黎曼结构而成为黎曼流形。黎曼流形是微分几何学的主要研究对象。

n 维流形一般概念是由德国著名数学家黎曼于 19 世纪首先提出的。1913 年外尔在一条代数曲线的特殊场合下,用坐标册和坐标变换第一次给出了流形的精确定义。微分流形的现代定义应归于维布伦和怀特黑德。

黎曼几何

【黎曼几何学】黎曼流形上的几何学。德国数学家黎曼于 19 世纪中期提出的几何学理论。1854 年,黎曼在哥廷根大学发表的题为“论做为几何学基础的假设”的就职演说。在这篇演说中,黎曼将曲面本身看作一个独立的几何实体,而不仅仅把它看成是欧氏空间中的一个几何实体。他首先发展了空间的概念。并提出这种空间上的几何学应基于无限接近的两点 (X^i) 和 $(X^i + dX^i)$, $(i=1, \dots, n)$ 之间的距离,在无穷小的意义下,这种距离仍然满足勾股定理。换言之,用微分弧长度的平方所确定的正定二次型理解度量。亦即 $ds^2 = \sum_{i,j=1}^n g_{ij} dX^i dX^j$, (g_{ij}) 是由函数构成的正定对称矩阵,这便是后来所说的黎曼度量。从黎曼度量出发而定义的几何学,即为黎曼几何学。

黎曼认识到度量只是加到流形上的一种结构,并且在同一流形上可以有许多不同的黎曼度量。黎曼以前的数学家仅知道三维欧氏空间 E_3 中的曲面 S 上存在诱导度量 $ds^2 = Edu^2 + 2Fdudv + Gdv^2$, 即第一基本形式。而未认识到曲面 S 还可以独立于 E_3 而定义,可以独立地赋予度量结构。黎曼意识到区分诱导度量与独立的黎曼度量的重要性,摆脱了经典微分几何曲面局限于诱导度量的束缚。创立了黎曼几何学,为近代数学和物理学的发展做出了杰出贡献。

黎曼几何以种种非欧几何作为特例。例如:定义度量 $ds^2 =$

$$\frac{\sqrt{\sum_{i=1}^n (dX^i)^2}}{1 + \frac{\alpha}{4} \sum_{i=1}^n (X^i)^2} \quad (\alpha \text{ 是常数}),$$

则当 $\alpha = 0$ 时是普通的欧氏几何,而当 $\alpha > 0$ 时为椭圆几何,当 $\alpha < 0$ 时就是双曲几何。

黎曼几何中的一个基本问题是微分形式的等价性问题。该问题 1869 年由克里斯托费尔和李普希兹解决。克氏的解中包含了以他的名字命名的两类记号,即克氏记号和协变微分概念。在此基础上李奇发展了张量分析方法。这在广义相对论中起了基本作用。他们进一步发展了黎曼几何学。

但在黎曼的时代,李群以及拓扑学还没有发展起来,因此黎曼几何只限于小范围的理论。大约在 1925 年霍普夫才开始对黎曼空间的微分结构与拓扑结构的关系进行了研究。随着微分流形精确概念的确立,特别嘉当在 20 世纪 20 年代开创并发展了外微分形式与活动标架法,建立起李群与黎曼几何之间的联系,从而为黎曼几何的发展奠定重要基础并开辟了广阔园地,影响极其深远。并由此发展了线性联络及纤维丛方面的研究。1944 年陈省身给出 n 维

流形 *Gaun-Bonnet* 公式的内蕴证明以及差不多同时他关于埃米尔特流形的示性类的研究,引进了后来通称的陈示性类,为大范围微分几何研究提供了不可缺少的工具,并为复流形的微分几何与拓扑的研究开创了先河。

1915 年,爱因斯坦运用黎曼几何和张量分析工具创立了新的引力理论——广义相对论。使黎曼几何研究获得巨大动力。而广义相对论近年的发展则受到整体微分几何的强烈影响。例如矢量丛和联络论构成规范场(杨-米尔斯场)的数学基础。

半个世纪以来,黎曼几何的研究已从局部发展到整体。产生了许多深刻的结果。随着 20 世纪 60 年代以来大范围分析的发展,黎曼几何与偏微分方程,多复变函数论,代数拓扑学等学科互相渗透,互相影响,在现代数学和物理学中取得了辉煌的成就。

拓 扑 学

一般拓扑学

【一般拓扑学】用点集的方法研究拓扑空间自身结构及拓扑不变量的学科。按克莱茵在厄兰格纲领中的观点,可用变换群对几何学进行分类,拓扑学是研究在拓扑变换群作用下不变的性质的学科。拓扑变换是指连续满单射且逆映射也连续。

一般拓扑学也称为点集拓扑学,一方面是因为 19 世纪数学奠基工作的需要,另一方面受 19 世纪以来几何基础以及更一般的公理化方法的影响而产生的。19 世纪末,康托建立了集合论,且进一步定义了欧氏空间的开集、闭集、导集、闭包等,获得了欧氏空间拓扑结构的重要结果。函数空间的研究,泛函分析中函数空间的列紧性问题,希尔伯特空间的研究,都为一般拓扑学理论的形成准备了大量具体材料,且使理论建立成为必要和可能。1906 年弗雷歇创立了广义分析这个具有高度抽象性的方向,在一定意义上可以说是数学分析基础部分的公理化,也是拓扑空间理论建立的开始。

数学分析研究的中心问题是极限,而收敛与连续是极限的基本问题。为把收敛和连续的概念推广到一般的集合上,需要在集合中描述与某点“邻近”点的概念。如何描述“邻近”,可以用距离,但距离与“邻近”并无必然联系。在 1913 年,大约同时弗雷歇和豪斯道夫开始考虑用“邻域”来定义拓扑,即一个集合 X ,规定每点有一个包含此点的 X 的子集簇,满足一组邻域公理,这就给出了 X 的一个拓扑结构,带有拓扑结构的集合 X 就称为一个拓扑空间, X 的每点均有一些邻域,就可按照微积分的方法定义两个拓扑空间之间的连续映射。总之,在 1900—1915 年间,经过弗雷歇、豪斯道夫等人的工作,奠定了一般拓扑学的基础。以后进一步系统研究拓扑映射下不变的性质。弗雷歇最早看出紧致与列紧的密切关系。豪斯道夫研究 T_2 空间等分离公理。亚历山大洛夫和乌里松对紧致空间进行了系统的研究,且对拓扑空间可否度量化问题作出了卓越的贡献。1937 年嘉当引进了滤子的概念,使收敛的更本质的属性刻画了出来,因为“邻近”只能刻画一点的“邻近”,而不能表达“两点很近”,因而无法刻画“一致性”的概念。维数的问题是嘉当在研究约当曲线时提出的。1912 年庞加莱给出了归纳定义,乌里松

等人的改进提出了局部维数的定义。在维数问题的研究中,布劳威尔第一次成功地建立了组合拓扑与一般拓扑学之间的结合。20世纪60年代,科恩提出了连续统假设的统一理论。此外近年来模糊拓扑的发展,也与点集拓扑研究有关。

【拓扑空间】集合 X 的一个子集簇 T , 满足(1)任意 $u_\alpha \in T, \alpha \in P$, 则 $\bigcup_{\alpha \in P} u_\alpha \in T$; (2)任意 $u_1, u_2, \dots, u_{n-1}, u_n \in T$, 则 $\bigcap_{i=1}^n u_i \in T$; (3) $X, \Phi \in T$, (其中 Φ 表示不含任何元素的集合) 则 (X, T) 称为一个拓扑空间, 有时简记为拓扑空间 X , 称 T 为 X 的一个拓扑, T 中的元称为开集。任意 $X \in X$, 包含 X 的开集 u 称为 X 的邻域, 记作 $u = u(X)$ 。 A 是 X 的子集, 若 $CA = X - A$ 是 X 的开集, 则称 A 是 X 的闭集。

定义拓扑结构的方法很多, 除以上方法外, 还有利用邻域公理或闭包公理、内部公理定义拓扑。

X 是集合, $T = \{X, \Phi\}$, 则称 (X, T) 为平庸拓扑空间, T 为平庸拓扑。 $T_1 = \{A \mid \text{任意 } A \subset X\}$, 则称 (X, T_1) 为离散拓扑空间。实数集 $R^1, T = \{G \subset R^1 \mid \text{任意 } X \in G, \text{存在 } \left. \begin{matrix} (X - \epsilon, X + \epsilon) \subset G \\ \epsilon > 0 \end{matrix} \right\} (\epsilon > 0)$, 则称 (R^1, T) 为一维欧氏空间。类似地可定义 $R^n = R^1 \times \dots \times R^1 = \{(X_1, X_2, \dots, X_n) \mid X_i \in R^1, i = 1, 2, \dots, n\}$ 中的欧氏拓扑。

若 $A \subset X, (X, T)$ 是拓扑空间, $T_A = \{A \cap u \mid u \in T\}$, 则称 (A, T_A) 是 (X, T) 的子空间。

若 $(X_i, T_i)_{i=1,2,\dots,n}$ 为拓扑空间簇, 在 $X_1 \times X_2 \times \dots \times X_n$ 上给拓扑 $T = \left\{ \bigcup_{\alpha \in P} (u_1^\alpha \times u_2^\alpha \times \dots \times u_n^\alpha) \mid \begin{matrix} u_i^\alpha \in T_i \\ i = 1, 2, \dots, n \end{matrix} \right\}$, 则称 $(X_1 \times X_2 \times \dots \times X_n, T)$ 为积空间, 注意此处 n 是有限数。

X 是拓扑空间, $\sim$ 是 X 上的等价关系, $X/\sim$ 是商集, $P: X \rightarrow X/$

$\sim$ 是商投影, 定义为 $P(X) = [X]$, 对任意 $X \in X$, 其中 $[X]$ 表 X 所在的等价类。在 $X/\sim$ 上定义拓扑 $T_P = \{u \subset X/\sim \mid P^{-1}(u) \subset X \text{ 开集}\}$, 称为 P 所诱导的商拓扑。 $(X/\sim, T_P)$ 称为商空间, 其中 $P^{-1}(u) = \{X \in X \mid P(X) \in u\}$ 。

(X, T) 是拓扑空间, B 是 T 的子簇, 满足 (1) $X = \bigcup_{u \in B} u$; (2) 任意 $G \in T$, 存在 $B_1 \subset B$, 使得 $G = \bigcup_{u \in B_1} u$, 则称 B 是 T 的拓扑基或拓扑空间 X 的拓扑基。

【连续映射】 $(X, T_1), (Y, T_2)$ 是拓扑空间, $f: X \rightarrow Y$ 是映射, 若对任意 $u \in T_2$, 有 $f^{-1}(u) \in T_1$, 则称 f 是连续映射。连续映射还有很多等价命题, 如: (1) $f: X \rightarrow Y$ 是映射, X 和 Y 是拓扑空间, 任意 $X \in X$, 任意 $f(X)$ 的邻域 W , 存在 X 的邻域 V , 使得 $f(V) \subset W$, 则称 f 是连续映射; (2) 任意 Y 中闭集 A , 有 $f^{-1}(A)$ 是 X 中的闭集, 则称 f 是连续映射。

若 $f: X \rightarrow Y$ 是连续映射, 且是满单射, 逆映射 $f^{-1}: X \rightarrow Y$ 也是连续映射, 则称 f 是同胚映射, X 和 Y 称为同胚的拓扑空间。

证明两个拓扑空间同胚, 只要找到它们之间的同胚映射。但区分不同胚的空间, 需要找一些拓扑不变性 (即在同胚映射下不变的性质)。

【连通空间】 X 是拓扑空间, 存在 U 和 V 是 X 的非空不交的两个开集, $X = U \cup V$, 则称 X 是不连通的拓扑空间。如果不存在这样的两个开集, 则称 X 是连通的。连通性是拓扑不变性。连通空间有一个最重要的性质是介值性, 若 $f: X \rightarrow R^1$ 是连续映射, X 是连通空间, R^1 是一维欧氏空间, 则任意 $X_1, X_2 \in X, c \in (f(X_1), f(X_2))$, 存

在 $\xi \in X$, 使 $f(\xi) = c$ 。这是微积分中闭区间上连续函数的介值定理的推广。

若 X 是拓扑空间, 任意 $X_1, X_2 \in X$, 存在 $f: [0, 1] \rightarrow X$ 是连续映射, 满足 $f(0) = X_1, f(1) = X_2$, 则称 X 为道路连通空间, f 为连结 X_1 和 X_2 的道路。道路连通空间一定是连通空间, 反之不一定。但在欧氏空间 R^n 中的开子集 G , G 作为子空间连通则路连通。道路连通性也是拓扑不变性。

【分离公理】 X 是拓扑空间, 任意 X_1 和 X_2 属于 X , 存在 X_1 的邻域 u_1 和 X_2 的邻域 u_2 , 满足 $u_1 \cap u_2 = \emptyset$, 则称 X 是豪斯道夫空间(或 T_2 空间); 如果 X 是 T_2 空间, 任意 X 的闭集 A 和任意 $X \in X - A$, 存在 A 的邻域 w (即包含 A 的开集), 及 X 的邻域 u , 使 $u \cap w = \emptyset$, 则称 X 是正则空间; X 是 T_2 空间, 任意 X 的两个不相交的闭集 A, B , 存在各自的邻域 W_1 和 W_2 , 满足 $W_1 \cap W_2 = \emptyset$, 则称 X 是正规空间。以上 T_2 空间、正则空间、正规空间均在同胚映射下不变。正规空间有两个重要特征:

(1) 乌里松引理: X 是 T_2 空间, 则 X 是正规空间, 当且仅当任意 X 的不相交的闭子集 A 和 B , 存在连续映射 $f: X \rightarrow [0, 1]$, 满足 $f(A) = \{0\}, f(B) = \{1\}$ 。此连续映射称为乌里松函数。

(2) 铁兹扩张定理 X 是 T_2 空间, 则 X 是正规空间, 当且仅当任意 X 的闭子集 A , 任意连续映射 $f: A \rightarrow R^1$, 存在连续扩张 $F: X \rightarrow R^1$, 即 F 连续, 且 $F|_A = f$ 。

【第一可数空间和第二可数空间】 X 是拓扑空间, 任意 $X \in X$, 存在至多可数个 X 的邻域作成的集簇 $\{u_n(X) | n = 1, 2, \dots, u_n(X) \text{ 是 } X \text{ 的邻域}\}$, 满足对任意 X 的邻域 $W(X)$, 存在某 $u_n(X) \subset W(X)$, 对

某 $n \in \mathbb{N}$ (自然数集), 则称 X 是第一可数空间; X 是拓扑空间, 如果存在至多可数个开集组成的拓扑基, 则称 X 是第二可数空间。第一可数性和第二可数性均是拓扑不变性。

【度量空间】 X 是集合, $d: X \times X \rightarrow \mathbb{R}^1$ 是映射, 满足 (1) $d(X, y) \geq 0$, $d(X, y) = 0 \Leftrightarrow X = y$, 对任意 $X, y \in X$, (2) $d(X, y) = d(y, X)$, 对 $\forall X, y \in X$, (3) 任意 $X, y, z \in X$ $d(X, z) = d(X, y) + d(y, z)$ 。则称 (X, d) 是度量空间, d 称为度量。令 $\beta_d(X, \epsilon) = \{y \in X \mid d(X, y) < \epsilon\}$, $\epsilon > 0$, 则 $T_d = \{\bigcup_{\alpha \in \rho} \beta_d(X_\alpha, \epsilon_\alpha) \mid \forall X_\alpha \in X, \forall \epsilon_\alpha > 0, \alpha \in \Gamma\}$ 称为距离 d 所诱导的拓扑, (X, J_α) 称为 d 所诱导的拓扑空间。

若拓扑空间 (X, T) 存在 X 的一个度量 d , 使 $T_d = T$, 则称 (X, T) 可度量化。若 (X, T) 是第二可数空间, 则 X 是正则空间当且仅当 X 可度量化。

【紧致空间】拓扑空间 X , 任意开子集簇 A , 满足 $X = \bigcup_{A \in A} A$, 则称 A 是 X 的一个开覆盖。如果能从 A 中选出有限个集合 $\{A_1, A_2, \dots, A_n\}$ 仍然作成 X 的一个开覆盖, 则称 $\{A_1, A_2, \dots, A_n\}$ 是 A 关于 X 的有限子覆盖。 X 是拓扑空间, 如果 X 的任意开覆盖均存在有限子覆盖, 则称 X 是紧致空间。微积分中闭区间 $[a, b]$ 有有限覆盖定理, 则在一维欧氏空间 $\mathbb{R}^1$ 中, $[a, b]$ 作为 $\mathbb{R}^1$ 的子空间, 是紧致空间。紧致空间有以下重要性质, 若 $f: X \rightarrow \mathbb{R}^1$ 是连续映射, X 是紧致拓扑空间, 则 $f(X)$ 是有界闭集; $f(X)$ 有最大和最小值; 若 (X, d) 是紧度量空间, $f: X \rightarrow \mathbb{R}^1$ 是连续映射, 则 f 一致连续, 即任意 $\epsilon > 0$, 存在 $\delta > 0$, 对任意 $X_1, X_2 \in X$, 当 $d(X_1, X_2) < \delta$ 时, 有 $|f(X_1) - f(X_2)| < \epsilon$ 。下面几个与紧致性有密切关系的概念。

X 是拓扑空间。 A 是 X 的子集, $x \in X$, 满足任意 x 的邻域 $U(x)$, 则 $U(x) \cap (A - \{x\}) \neq \emptyset$, 则称 x 为 A 的聚点。

X 是拓扑空间。如果 X 的任意无限子集 A 必有聚点, 则称 X 为列紧空间。 X 是拓扑空间, X 中的任意序列 $\{X_n\}_{n=1,2,\dots}$, 如果存在 $X_0 \in X$, 对于 X_0 的任意邻域 $U(X_0)$, 均存在自然数 N , 当 $n > N$ 时, $X_n \in U(X_0)$, 则称 $\{X_n\}_{n=1,2,\dots}$ 收敛于 X_0 。若 X 中的任何序列 $\{X_n\}_{n=1,2,\dots}$, 均存在收敛子列, 则称 X 是序列紧致空间。列紧性与序列紧性均是拓扑不变性。对于度量空间, 紧致性、列紧性、序列紧性等价。对于欧氏空间 R^n , $A \subset R^n$ 作为子空间, 下列命题等价: (1) A 是 R^n 的有界闭集; (2) A 是紧致空间; (3) A 是列紧空间; (4) A 是序列紧空间。

代数拓扑

【代数拓扑】拓扑学中用代数工具解决问题的分支, 研究拓扑空间的代数不变量的拓扑学的分支。与一般拓扑学渊源于集合论及泛函分析的研究不同, 它主要起源于几何, 前身是组合拓扑。最初的定理是 1736 年欧拉发表的研究著名的哥尼斯七桥问题的解答, 且给出了连通网络一笔画的充要条件。接着是欧拉-庞加莱公式, 即与二维球面 S^2 同胚的闭多面体的顶点数、棱数和面数分别用 α_0 , α_1 和 α_2 表示, 则 $\alpha_0 - \alpha_1 + \alpha_2 = 2$ 。继欧拉之后, 1833 年前后高斯研究了空间中各种各样的纽结。1855 年数学大师黎曼进行了黎曼面的研究。庞加莱是公认的组合拓扑、代数拓扑的奠基人。1895 年他以“关于位置解析”为题发表了数篇论文, 对拓扑空间进行代数研

究,是数学发展的一个里程碑。他建立了单纯同调群,即可三角剖分的流形(即多面体)的同调群,引进了重要的拓扑不变量贝蒂数及挠系数。同时庞加莱还引进了多面体的基本群。1904年,他给出了著名的庞加莱猜想,即每个单连通的闭的可定向的三维流形同胚于三维球面 S^3 。后来这个著名的猜想又被推广为每个单连通的闭的 n 维流形, $n \geq 2$, 如果具有 n 维球 S^n 的贝蒂数和挠系数,则必同胚于 S^n 。庞加莱猜想至今还没有证明。推广了的庞加莱猜想也仅对 $n \geq 5$ 的情况被斯梅尔(*S. Smale*) 在 1960 年证明了。庞加莱企图用同调群和基本群对三维流形进行拓扑分类,但亚历山大在 1919 年指出存在不同胚的三维流形,它们有同构的同调群和基本群。但他没有解决庞加莱猜想,亚历山大在 1915 年证明了贝蒂数和挠系数是同胚不变量,且用两个多面体之间的连续映射的单纯逼近(即一种线性逼近)证明了同调群是拓扑不变量。20 世纪初,布劳威尔扩张了若当曲线定理。20 世纪 20 年代莱夫谢茨和亚历山大发表了同调论,给出了霍普夫(*Hopf*)不变量,证明了莱夫谢茨不动点定理,亚历山大对偶定理。1930 年以后,拓扑加速发展,有惠特尼关于微分形式的工作,霍普夫关于李群的工作,拓扑学使微分几何焕发了青春,且促进了代数中同调代数这一新分支的发展,使现代代数学出现了一次小变革,这些工作是艾伦伯格作的。20 世纪初引进了一般拓扑空间的同调群,1932 年切克上同调产生,1940 年苏联的数学家亚历山大等人建立了上同调群、上同调环。1944 年艾伦伯格定义了奇异同调群。且用艾伦伯格-斯廷罗德公理把各种同调群统一起来。在同伦论方面,赫尔维茨(*W. Hurewicz*) 定义了同伦群,怀特赫德把研究对象推到 CW 复形。1947 年斯廷罗德在障碍理论中定义了斯廷罗德平方运算,1951 年塞尔对纤维丛引进了谱序列,在同伦群的计算方面取得了不少进

展。此外纽结问题也进一步发展成为高维合痕和嵌入问题。

【同调群】拓扑空间在同胚映射下不变的群的序列。同调的直观思想解释如下：

左图平环中的曲线 C_1 是所围区域的边界,则称 C_1 为零调的一维闭链。 C_2 和 C_3 一起是共同所围区域的边界,则称一维闭链 C_2 和 C_3 同调。类似地可定义各维闭链之间的同调关系。可证同调是一个等价关系。用它对各维闭链进行等价分类,所得的商群称为同调群。用它可刻划拓扑空间包含各维洞的“个数”及扭曲的情况。

R^n 中的点 $\sigma_0 = \langle a_0 \rangle$ 称为零维单形;线段 $\sigma_1 = \langle a_0, a_1 \rangle$ 称为一维单纯形(或 1-单形),不共线三点为顶点作成的三角形 $\sigma_2 = \langle a_0, a_1, a_2 \rangle$ 称为二维单纯形(或 2-单形);若 $\{a_0, a_1, \dots, a_k\}$ 是 R^n 中的 $k+1$ 个点,满足 $a_1 - a_0, a_2 - a_0, \dots, a_k - a_0$ 线性无关,则包含 $a_0, a_1, \dots, a_k$ 的最小凸集 $\sigma_k = \langle a_0, a_1, \dots, a_k \rangle$ 称为 k 维单形,简称 k -单形。

$\langle a_0, a_1, \dots, a_k \rangle$ 是 k -单形,若 $0 \leq i_0 < i_1 < \dots < i_l \leq k$ 则 $\langle a_{i_0}, a_{i_1}, \dots, a_{i_l} \rangle$ 称为 $\langle a_0, a_1, \dots, a_k \rangle$ 的 l 维面单形。 m 维单纯复形 K (简称为

m 维复形)是有限个适当连结的单形的集合。满足若 $\sigma_p \in K$, 则 σ_p 的所有面单形也属于 K ; 若 $\sigma_p, \sigma_q \in K$, 且 $\sigma_p \cap \sigma_q \neq \emptyset$, 则 $\sigma_p \cap \sigma_q$ 是 σ_p 与 σ_q 的公共面单形, 此时称 σ_p 与 σ_q 适当连结。

K 是单纯复形, K 中单形的并称为联系于 K 的多面体, 记作 $|K|$ 。

若 X 是拓扑空间, 存在单纯复形 K , 使多面体 $|K|$ 同胚于 X , 则称 X 可三角剖分, 称 K 为 X 的一个三角剖分。

单形给一定向称为定向单形。例如 $\sigma_1 = \langle a_0, a_1 \rangle$ 若给一个顶点的顺序 a_0, a_1 , 或 a_1, a_0 , 则称为一个定向单形, 如给 a_0, a_1 为正向, 则记定向单形为 $+\sigma_1 = +\langle a_0, a_1 \rangle$; 如给 a_1, a_0 为正向, 则记定向单形为 $+\sigma_1 = +\langle a_1, a_0 \rangle$ 。又如单形 $\langle a_0, a_1, \dots, a_k \rangle$ 给 $a_0 a_1 a_2 \dots a_k$ 的奇排列或偶排列, 就可得到两个不同定向的单形。如把 $a_0, a_1, \dots, a_k$ 的偶排列作为正向, 则记 $+\sigma_k = +\langle a_0, a_1, \dots, a_k \rangle$ $-\sigma_k = \langle a_0, a_2, \dots, a_k \rangle$; 若把 $a_0, a_1, \dots, a_k$ 的奇排列作为正向, 则记 $+\sigma_k = +\langle a_1, a_0, a_2, \dots, a_k \rangle$, $-\sigma_k = -\langle a_0, a_1, \dots, a_k \rangle$ 。若一个复形 K 的各单形均给定了定向, 则称 K 是一个定向复形。

若 $\sigma_k = \langle a_0, a_1, \dots, a_k \rangle$ 是 σ_{k+1} 的面单形, V 是 σ_{k+1} 的不在 σ_k 中的顶点 $+\sigma_{k+1} = +\langle v, a_0, a_1, \dots, a_k \rangle$, $+\sigma_k = +\langle a_0, a_1, \dots, a_k \rangle$ 称为 $+\sigma_{k+1}$ 的顺向面; 而 $-\sigma_k = +\langle a_1, a_0, \dots, a_k \rangle$ 称为 $+\sigma_{k+1}$ 的逆向面。

K 是定向复形, 若 $\sigma_k, \sigma_{k+1} \in K$, 当 $+\sigma_k$ 是 $+\sigma_{k+1}$ 的顺向面, 则定义关联系数 $[+\sigma_{k+1}, +\sigma_k] = 1$; 若 $+\sigma_k$ 是 $+\sigma_{k+1}$ 的逆向面, 则定义关联系数 $[+\sigma_{k+1}, +\sigma_k] = -1$; 若 σ_k 不是 σ_{k+1} 的面单形, 则定义关联系数 $[+\sigma_{k+1}, +\sigma_k] = 0$, 不难证明若 σ_{m-2} 是 σ_m 的面单形, 它们都是复形 K 的单形, 全体 K 的 $m-1$ 维单形为 $\sigma_{m-1}^i, i = 1, 2 \dots$

α_{m-1} , 则 $\sum_{i=1}^{\alpha_{m-1}} [+\sigma_m, +\sigma_{m-1}^i] [+\sigma_{m-1}^i + \sigma_{m-2}] = 0$ (*)。

K 是 m 维定向单纯复形, $0 \leq p \leq m$, K 的 P 维链是形如 $C_p = \lambda_1(+\sigma_p^1) + \cdots + \lambda_{\alpha_p}(+\sigma_p^{\alpha_p})$ 的线性组合, 其中 $\sigma_p^i \in K, i=1, 2, \dots, \alpha_p, \lambda_i$ 是整数。两个 P 维链相加之和是一个 P 维链, 其中每个 P 维定向正单形前的系数是相加的两个 P 维链相应单形前的系数之和。在此加法之下, 所有 K 的 P 维链的集合 $C_P(K)$ 作成一群, 称为 K 的 P 维链群。

$$\partial_p: C_P(K) \rightarrow C_{P-1}(K) \text{ 是一个同态, 定义为 } \partial(+\sigma_p) = \sum_{i=1}^{\alpha_{p-1}} [+ \sigma_p, +\sigma_{p-1}^i](+\sigma_{p-1}^1, \partial(\lambda_1(+\sigma_p^1) + \cdots + \lambda_{\alpha_p}(+\sigma_p^{\alpha_p})) = \sum_{i=1}^{\alpha_p} \lambda_i \partial(+\sigma_p^i)。$$

K 的 P 维闭链群 $Z_P(K) = \text{Ker } \partial_p$

K 的 P 维边缘链群 $B_P(K) = \text{Im } \partial_{P+1}$ 。由 $(*)$ 可证 $\partial_{P-1} \circ \partial_P = 0$, 则 $B_P(K)$ 是 $Z_P(K)$ 的子群, 则是正规子群。

K 的 P 维同调群 $H_P(K) = Z_P(K) / B_P(K)$ 。

利用多面体之间的连续映射的单纯逼近可诱导同调群的同态, 可证明这诱导同态具函子性质, 由此可证明单纯同调群是多面体 $|K|$ 的同胚不变量。则我们称多面体 $|K|$ 的同调群为 $H_p(|K|) = H_p(K)$ 。如果拓扑空间 X 的三角剖分为 K , 则记 $H_p(X) = H_p(K)$ 。根据代数知识, 有限生成群可分解为自由交换群和有限循环群(挠群)的直和, 则 $H_p(K)$ 同构于 $Z \oplus Z \oplus \cdots \oplus Z \oplus Z_{\theta_p^1} \oplus \cdots \oplus Z_{\theta_p^{R_p}}$, 其中 Z 是整数加群 Z_i 是整数加群模 i 的同余类, R_p, Z_p 是整数, 且 $Z_p > 0$ 时, θ_p^i 整除 θ_p^{i+1} 。 R_p 称为 K 的 P 维贝蒂数, θ_p^i 称为 p 维挠系数 $i=1, 2, \dots, z_p$ 。零维贝蒂数表示 $|K|$ 的连通分支的个数, p 维贝蒂数表示同调无关的 p 维闭链的个数。如果某个 p 维非边缘的闭链的 l 倍是边缘链, 且是使此闭链变为边缘链的最小数, 则称 l 是 P 维挠系数。 m 维单纯复形 K 无零维和 m 维挠系数。因此可用复形 K 的

p 维同调群研究多面体 $|K|$ 或与它同胚的拓扑空间 X 的 p 维洞的个数, 及扭曲的情况。

K 是单纯复形, G 是交换群, 称同态群 $Hom(C_p(K), G) = \{C^p: C_p(K) \rightarrow G \text{ 是同态}\}$ 为复形 K 的以 G 为系数群的 p 维上链群, 记作 $C^p(K; G)$ 。

上边缘同态 $\delta^p: C^p(K; G) \rightarrow C^{p+1}(K; G)$ 定义如下: 任意 $c^p \in C^p(K; G)$, $\delta^p c^p \in C^{p+1}(K; G)$ 满足对任意 $d_{p+1} \in C_{p+1}(K)$ 有 $\delta^p c^p(d_{p+1}) = c^p(\partial_{p+1})$ 。不难证明 $\delta^{p+1} \circ \delta^p = 0$ 。

$Z^p(K; G) = Ker \delta^p$ 称为复形 K 的上闭链群。

$B^p(K; G) = Im \delta^{p-1}$ 称为复形 K 的上边缘链群。 $B^p(K; G)$ 是 $Z^p(K; G)$ 的子群。

$H^p(K; G) = Z^p(K; G) / B^p(K; G)$ 称为 K 的 p 维上同调群。上同调群仍然是拓扑不变量。它不仅有加法, 还可以定义乘法, 使其作成环。上同调环也是拓扑不变量。

上、下同调群均满足相应的七条公理、反映函子性质的恒同公理、合成公理、自然变换性质的变换公理以及同伦公理、切除公理、恰当公理、维数公理。艾伦伯格等人证明了其它的同调群如奇异同调群也满足此七条公理, 因而可抽去单纯同调群的具体几何含意, 定义满足七组公理的同调理论, 可用恰当序列来研究同调群。

以后数学家们又发现了满足除维数公理以外六组公理的广义同调理论, 它的具体模型如研究纤维丛引出的 K 同调群; 研究微分流形协边理论引出的协边同调群; 研究球面同伦群引入的 BP 同调群等。20 世纪 60 年代证明了满足一定条件的空间的广义同调群, 自然同构于此空间到某空间的同伦类组成的群, 这就是所谓表现定理。由此将同调论与同伦论统一起来。同调论在数学各分支均有很广泛的应用。

【基本群】拓扑空间所相应的在同胚映射下不变的群。它是庞加莱为研究三维以及更高维流形的拓扑分类时引入的拓扑不变量。其引入对研究覆盖空间、若当曲线定理及代数拓扑的很多问题起了很大的作用。拓扑空间 X 的以 X_0 为始点终点的闭道路 α_1 和 α_2 如果能够保持始点终点 X_0 不动连续地把 α_1 变为 α_2 , 则称 α_1 与 α_2 保持基点 X_0 同伦, 记作 $\alpha_1 \sim_{X_0} \alpha_2$ 。在 X 中以 X_0 为基点的闭道路的集合 $\Omega(X, X_0)$ 中 $\sim_{X_0}$ 是一个等价关系。记在等价关系 $\sim_{X_0}$ 之下, 闭道路 α 所在的等价类为 $[\alpha]$ 。两个以 X_0 为基点 (即以 X_0 为始点终点) 的闭道路 α_1 和 α_2 相乘 $\alpha_1 * \alpha_2$ 仍是一个以 X_0 为基点的闭道路, 定义为

$$\alpha_1 * \alpha_2(t) = \begin{cases} \alpha_1(2t) & 0 \leq t \leq \frac{1}{2} \\ \alpha_2(2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases},$$

即前 $\frac{1}{2}$ “时间”走 α_1 , 后 $\frac{1}{2}$ 时间走 α_2 。 $[\alpha_1] \circ [\alpha_2] \stackrel{\text{定义}}{=} [\alpha_1 * \alpha_2]$, 可以证明在乘法 $\circ$ 之下, $\Omega(X, X_0) / \sim$ 作成一群, 记作 $\pi_1(X, X_0)$, 称为空间 X 的以 X_0 为基点的基本群。用它可刻划 X 中 X_0 所在路连通支中一维洞的个数。由此, 单连通区域是基本群为零且路连通的区域, 而多连通平面区域的基本群不为零。

【同伦论】代数拓扑的重要分支之一, 是研究连续映射与其形变有关的问题。两个连续映射 $f, g: X \rightarrow Y$, 如果 f 能连续地形变到 g , 则称 f 与 g 同伦。在拓扑空间 X 到 Y 的全体连续映射的集合 Y^X 中, 同伦是等价关系, 它们的等价类的集合记作 $[X; Y]$ 。同伦论主要研究下面几方面的问题:

(1)同伦扩张问题 $f:A \subset Y$ 连续, $A \subset X$ 是子空间, 如果存在连续映射 $F:X \rightarrow Y$, 满足 $F|_A = f$, 则称 F 是 f 的连续扩张。同伦扩张问题是指若 $f, g:A \rightarrow Y$ 是连续映射, f 与 g 同伦, 若 f 可连续扩张为 $F:X \rightarrow Y$, 问 g 是否可连续扩张到 X 上。如果每两个 A 到 Y 的同伦的连续映射, 一个能连续扩张到 X 上, 则另一个也能, 我们就称 (X, A) 关于 Y 有同伦扩张性。如果 (X, A) 关于任何空间 Y 均有同伦扩张性, 则称包含映射 $i:A \rightarrow X$ 为上纤维化, 此处 $i(a) = a$, 对任意 $a \in A$ 。

(2)同伦提升性 $P:E \rightarrow B$ 是连续映射, X 是给定的拓扑空间, 对任意连续映射 $f:X \rightarrow E$, 任意连续映射 $H:X \times I \rightarrow B$, 满足 $p \circ f = H|_{X \times \{0\}}$, 如果均存在连续映射 $\tilde{H}:X \times I \rightarrow E$, 满足 $P \circ \tilde{H} = H$, $\tilde{H}|_{X \times \{0\}} = f$, 则称 $P:E \rightarrow B$ 关于 X 有同伦提升性。如果 $P:E \rightarrow B$ 关于任何空间 X 均有同伦提升性, 则称 P 为一个纤维化。

(3)同伦分类问题, $[X; Y]$ 的分类问题。主要讨论同伦群 $\Pi_n(X, X_0)$ 。 X 是拓扑空间。 $X_0 \in X$, n 维球 $S^n = \{t \in \mathbb{R}^{n+1} \mid \|t\| = 1\}$, $\bar{I} = (1, \overbrace{0, 0, \dots, 0}^{n \uparrow}) \in S^n$, 对于 $\alpha, \beta(S^n, T) \rightarrow (X, X_0)$ 是连续映射, 存在 $H:S^n \times I \rightarrow X$ 是连续映射, 满足 $H(t, 0) = \alpha(t)$, $H(t, 1) = \beta(t)$ 对任意 $t \in S^n$, $H(\bar{1}, s) = X_0$, 对任意 $s \in I = [0, 1]$ 。则称 α 与 β 保持基点 X_0 同伦, 记作 $\alpha \sim_{X_0} \beta$ 。用 $[S^n, \bar{1}; X, X_0]$ 记 $(S^n, \bar{1})$ 到 (X, X_0) 的连续映射的保持基点同伦类的集合, 可定义一个乘法“ $\circ$ ”, 我们可证在此乘法下 $[S^n, \bar{1}; X, X_0]$ 作成一群, 记作 $\Pi_n(X, X_0)$ 称为 X 的以 X_0 为基点的 n 维同伦群。同伦群定义简单, 但计算复杂, 现在甚至关于 S^n 的同伦群都不全会称。我们现已知 $\Pi_n(S^n, \bar{1})$ 同构于 $\mathbb{Z}$ (整数加群); $\Pi_k(S^n, \bar{1}) = \{0\}$, 当 $k < n$ 时; $\Pi_k(S^1, \bar{1}) = \{0\}$, 当 $k > 1$ 时;

$\Pi_k(S^n, \bar{1})$ 同构于 $\pi_{k+1}(S^{n+1}, \bar{1})$ 当 $k < 2n-1$ 时。同伦群是代数拓扑的重要课题。1960 年以后, 利用谱序列, 同伦群的计算方面的研究有了很大的进展。

利用同伦群可刻画空间的“连通性”。早在 1851 年, 伟大数学家黎曼(*G. F. B. Riemaun*) 在黎曼面中引进了“ n 连通”的概念, 他定义: “一个曲面 F (具有边界) 上能画 n 条闭曲线 $\alpha_1, \alpha_2, \dots, \alpha_n$, 它们各自独立或集体均不能包围曲面 F 的一部分, 但它们连同另外一条曲线就能包围曲面或它的一部分, 则称 F 是 n 连通的。可以用同伦群定义 n 连通, 如果 X 连通, $\Pi_k(X, X_o) = \{0\}, k \geq n$, 但 $\Pi_{n-1}(X, X_o) \neq \{0\}$, 则称 X 是 n 连通的。例如 $\overline{B^2(0, 1)} = \{X \in \mathbb{R}^2 \mid \|X\| \leq 1\}, \Pi_n(\overline{B^2(0, 1)}, X_o) = \{0\}, n = 1, 2, \dots$, 所以 $B^2(0, 1)$ 单连通。又如 $\Pi_k(S^1, i) = \{0\} k \geq 2$, 而 $\Pi(i, s_1) \neq \{0\}$, 则 S^1 2-连通。

同调群与同伦群有着密切的关系, 如果 K 是连通的单纯复形, 若 $n \geq 2$, 当 $k \leq n-1$ 时 $\Pi_k(|K|, X_o) = \{0\}$, 则 $\Pi_n(|K|, X_o)$ 同构于 $H_n(K)$ 。

微分拓扑

【微分拓扑】研究微分流形在微分同胚映射下不变的性质的学科。它的研究对象是微分流形及这些流形之间的可微映射。它首先研究微分流形的拓扑结构、组合结构与微分结构之间的关系。把研究可微映射放在非常重要的地位, 主要研究的问题有(1)微分同胚问

题:两个微分流形判断它们是否同胚;(2)配边(协边)问题:给一个无边的紧致微分流形,判断它们是否为某个有边微分流形的边界,或两个无边的微分流形,是否一起作成第三个有边微分流形的边界;(3)嵌入问题:两个微分流形 M 和 M' , M 是否可以光滑地嵌入 M' 中;(4)微分动力体系:研究单参微分同胚群;(5)奇点理论:可微映射的奇芽的分类问题;(6)突变论等。

微分流形从直观上看是局部欧氏空间,各局部之间用光滑的同胚映射(局部坐标映射)将它们粘在一起。局部坐标映射的全体称为微分流形的微分结构。一个拓扑流形(即局部坐标变换不要求光滑)可以有不同的微分结构,例如 S^7 存在多个相异的微分结构,称为怪球。

$f:M \rightarrow N$ 是微分流形之间的连续映射,如果在每个点的邻域, f 合成局部坐标映射后,是欧氏空间的 C^∞ 光滑的映射,则称 f 是 C^∞ 光滑的映射。如果 $f:M \rightarrow \mathbb{R}^1$ 是光滑映射,其中 M 是微分流形, $\mathbb{R}^1$ 是一维欧氏空间,在点 $p \in M$ 的局部坐标映射为 φ ,映射 $f \circ \varphi^{-1}:\mathbb{R}^n \rightarrow \mathbb{R}$ 的雅可比矩阵在 $\varphi(p)$ 点的秩为零,则 p 称为 M 的临界点。又若矩阵 $\left[\frac{\partial^2 f \circ \varphi^{-1}}{\partial X_i \partial X_j}(\varphi(p)) \right]$ 非退化,则 p 称为非退化的临界点。上述矩阵的负特征值的个数称为 f 在 p 点的指数。若 $f:M \rightarrow \mathbb{R}$ 是 C^2 光滑映射,具有非退化的临界点,则称 f 是莫尔斯函数,莫尔斯理论告诉我们,若 M 是 n 维微分流形, $f:M \rightarrow \mathbb{R}^1$ 是 C^2 光滑映射, $p \in M$ 是非退化的临界点,则存在 p 点的局部坐标函数 $\varphi:V(p) \rightarrow \varphi(V(p)) \subset \mathbb{R}^n$ 中,满足 $f \circ \varphi^{-1}(X_1, X_2, \dots, X_n)$

$=f(p) - \sum_{i=1}^k X_i^2 + \sum_{j=k+1}^n X_j^2$ 。即 f 在 p 的邻域是 $f(p)$ 加上一个二次型。莫尔斯函数不仅可以研究微分流形的局部性质,而且还提供了

定义在其上的微分流形的整体性质的一些信息。如果 $f: M \rightarrow \mathbb{R}^1$ 是莫尔斯函数, 则微分流形 M 具有 CW 复形的伦型, 对每个 f 的指数为 k 的非退化的临界点, 这个 CW 复形对应一个 k 胞腔。此处 k 胞腔 l_k 是指内部同胚于 k 圆盘 $B^k(0, 1)$ 。而 CW 复形是由有限个各维胞腔适当地粘贴而成。两个拓扑空间 X 和 Y 称为伦型一样, 是指存在连续映射 $f: X \rightarrow Y$ 连续, $g: Y \rightarrow X$ 连续, 满足 $g \circ f$ 同伦于 X 上的恒等映射 Id_X , $f \circ g$ 同伦于 Y 上的恒等映射 Id_Y 。莫尔斯理论应用到大范围的变分问题上, 可利用它刻划测地线的“条数”。

光滑映射的奇点理论其渊源是莫尔斯理论及 20 世纪 40 年代惠特尼的与微分嵌入有关的奇点问题方面的工作, 但直到 1955 年, 惠特尼发表了关于平面到平面的可微映射的奇点的分类问题方面的工作, 才标志着它作为一个独立的数学分支登上了数学舞台。1956 年托姆题为“可微映射奇点”论文的发表, 对整个奇点理论的发展提出了一个纲领性的描述。对微分流形 M 到 N 的光滑映射的全体作成的集合 $C^\infty(M, N)$ 中的映射的芽进行等价分类, 是比对 $\mathbb{R}^n$ 中的全体闭集进行等价分类还要广泛得多的问题, 是不能实现的。因而要找一个等价关系能对 $C^\infty(M, N)$ 的一个稠密子集进行等价分类, 而类别是有限个或可数个, 就称此分类问题有完全解。莫尔斯理论指出: 在 $C^\infty(M, \mathbb{R}^1)$ 中, 全体莫尔斯函数作成它的一个稠密子集, 且若 f 是莫尔斯函数, 则 f 在每点 $p \in M$ 的邻域上等价于映射芽 $\tilde{g}: (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^1, 0)$, 其中 $\tilde{g}(X_1, X_2, \dots, X_n) = X_1$ 或 $\tilde{g}(X_1, X_2, \dots, X_n) = X_1^2 + X_2^2 + \dots + X_{n-k}^2 - X_{n-k+1}^2 - \dots - X_n^2, k = 1, 2, \dots, n$ 其中之一。当 M^2 是二维 C^∞ 光滑流形, 马瑟证明了 $C^\infty(M^2, \mathbb{R}^2)$ 中的稳定映射作成它的一个稠密子集, 且 $f: M^2 \rightarrow \mathbb{R}^2$ 在 $P \in M$ 稳定的充要条件是 P 点的芽局部表示为下列三者之一: (1)

$$\left\{ \begin{array}{l} u=X \\ v=y \end{array} \right. \text{(正则点)}, (2) \left\{ \begin{array}{l} u=X \\ v=y^2 \end{array} \right. \text{(折叠点)}, (3) \left\{ \begin{array}{l} u=X \\ v=Xy-\frac{1}{3}y^3 \end{array} \right. \text{(尖点)}。马$$

瑟证明了对于 m 维微分流形 M 到 n 维微分流形 N 的全体 C^∞ 光滑映射的集合, 它的稳定映射作成其稠密子集的条件是 m, n 满足下列条件: (1) $n < 7(n-m) + 8$, 当 $n-m \geq 4$ 时; (2) $n < 7(n-m) + 9$, 当 $0 \leq n-m \leq 3$ 时; (3) $n < 8$, 当 $n-m = -1$ 时; (4) $n < 6$, 当 $n-m = -2$ 时; (5) $n < 7$, 当 $n-m \leq -3$ 时。满足上述条件的 (n, m) 称为“好维数”, 对好维数, 问题有完全解。

突变论是 20 世纪 60 年代末, 托姆提出的。马瑟证明了突变论的主要定理, 以后经过阿尔诺德、齐曼等人的介绍文章及大量地在各学科的应用, 使突变论轰动了科学界。对于连续变化引起不连续突变的现象, 在自然界乃至社会科学中是常见而又十分重要的, 例如火山爆发、地震、桥梁突然倒塌等突变现象, 过去处理起来很棘手, 突变论帮助我们提供找数学模型的途径。找一个系统的势函数, 要做大量试验, 并不指望试验条件一样, 结果一样, 而希望条件近似时结果性质不变。用数学的语言即所求势函数结构稳定, 当状态变量作微小扰动时, 函数的性质不变(仅相差一个微分同胚)。势函数 $f(X_1, X_2, \dots, X_n, u_1, \dots, u_m)$ 的状态变量为 $(X_1, \dots, X_n)$ 控制变量为 $(u_1, u_2, \dots, u_m)$ 。“本质性”的状态变量的个数可以小于 n 。例如 $f(X_1, X_2, \dots, X_n) = X_1^2 + \dots + X_k^2 +$

$g(X_{k+1}, \dots, X_n)$, “本质性”的(即决定 f 性质)的状态变量个数 $\leq n - k$ 。莫尔斯理论告诉我们, 当黑森矩阵 $\left[\frac{\partial^2 f}{\partial X_i \partial X_j}(0) \right]$ 秩 $= n - r$ 时, 此时称 f 的余秩 $\text{rank} f = r$, 此时存在微分同胚 φ , 使 $f \circ \varphi(y_1 \dots y_n) = \pm y_{r+1}^2 + \dots + y_n^2 + g(y_1, \dots, y_r)$, 此时 f “本质性”的状态变量仅有 r 个。此时突变论告诉我们, 当 f 的控制变量的个数 $\text{Codim} f \geq$

$C_{\gamma+1}^2$ 。由此,当 f 的控制变量个数 ≤ 4 时,“本质性”的状态变量的个数为 1 或 2。而当 $\text{Codim}f \leq 4$ 时, f 与它的 6 次泰勒多项式等价(即相差一个微分同胚)。突变论还进一步证明,此时 $f(X_1, X_2, \dots, X_n, u, v, w, t)$ 等价于 $\pm X_1^2 \pm \dots \pm X_{n-2}^2 + g(X, y, u, v, w, t)$, 其中 $g(X, y, u, v, w, t)$ 为下列七种突变模型之一: 折叠 $X^3 + uX$; 尖点 $X^4 + uX^2 + vX$; 燕尾 $X^5 + uX^3 + vX^2 + wX$; 蝴蝶 $X^6 + uX^4 + vX^3 + wX^2 + tX$; 椭圆脐点 $X^3 - XY^2 + uX + vy + w(X^2 + y^2)$; 双曲脐点 $X^3 + y^3 + uX + vy + wXy$; 抛物脐点 $y^4 + X^2y + uX + vy + wX^2 + ty^2$ 。这七种突变模型结构稳定(即满足横截条件),我们研究此七种模型的特征,可在实际问题中根据特征找模型,也可给出模型,确定它属于七类中那一类,由此得知其特征。例如尖点突变 $g(X, u, v) = X^4 + uX^2 + vX$, $g'_X = 4X^3 + 2uX + v$, $g''_{XX}(X, u, v) = 12X^2 + 2u$

则平衡曲面为 $4X^3 + 2uX + v = 0$ (*), 判别式 $\Delta = 8u^3 + 27v^2$, 当 $\Delta \leq 0$ 时(*)有三个实根, $\Delta > 0$ 时有一个实根, $\Delta = 0$ 时或两实根相等, 或三个实根相等。不难算出 $\Delta =$

$$\left\{ (u, v) \left| \begin{array}{l} 4X^3 + 2uX + v = 0 \\ 12X^2 + 2u = 0 \end{array} \right. \right\} = \{ (u, v) \mid 8u^3 + 27v^2 = 0 \}, \text{由此可求出}$$

不同 (u, v) 点, $g(X, u, v)$ 的极值点, 可将函数图像简单画出平衡曲面 $4X^3 + 2uX + v = 0$ 及 (u, v) 面上控制曲线

可分析出尖点突变具下列特性; (1) 突跳性: 当控制平面内一条曲线穿越尖点区域 $8u^3 + 27v^2 = 0$ 时, 相应点发生突跳; (2) 滞后性: 同一条曲线穿越尖点区域时, 从左边穿越与右边穿越, 发生突跳的点不同, 具有滞后性; (3) 双模性: 尖点区域内有两个极小值点(即稳定平衡点); (4) 不可达性: 尖点区域内有一个极大值点(不稳平衡点); (5) 发散性: 尖点区域外一点出发两条很邻近的曲线 C_1, C_2 通过原点附近进入尖点区域, 相应平衡点位置不同, 一个在上叶,

一个在下叶。(见图)。

概 率 论

【概率论】研究随机现象数量规律的数学分支。随机现象是指这样的客观现象:当人们在一定条件下进行试验或观察时,不能预先确定会出现什么结果,只知出现的是多个可能结果中的一个。在自然界和人类社会中广泛存在着这类现象。例如,射手向目标射击一次,可能命中或未命中目标;在相同条件下对一物体进行测量,每次测量结果往往不尽相同;在同一生产条件下,生产出的产品的尺寸大小总是参差不齐等。这些都是随机现象。

呈现出随机现象的试验或观察统称随机试验,简称试验。它有以下特点:(1)试验可在相同条件下重复进行;(2)试验可能出现的结果至少有两个,而且事先已知;(3)在每次试验前,无法预先确定会出现可能结果中的哪一个。上述的射击、测量和生产产品等例子均为随机试验。

对于随机试验,就一次试验而言,它会时而出现这样的结果,时而出现那样的结果,呈现出一种不确定性,这是随机试验偶然性的一面;另外,人们还发现,在重复进行了大量试验后,出现的结果呈现出某种规律性,人们称之为“统计规律性”,这是随机试验必然性的一面。概率论就是从量的方面去揭示和研究这种规律性。

对于不同 u, v 的 $g(X, u, v)$ 的值

概率论起源于赌博和机会游戏其历史可追溯至 16 世纪。从 16 世纪到 17 世纪是概率论发展的早期阶段。这时,基于排列组合的方法,局限于讨论赌博中的如“分赌注”、“输光”等问题,但由此却结晶出了概率论中的概率和数学期望等重要概念。17 世纪末至 19 世纪,随着误差理论、射击问题、人口统计问题以及保险业等的迫

切需要,概率论的发展进入到一个新的阶段。这时,在概率论中开始引用更先进的分析工具,奠定了概率论的基础。当时,法国数学家拉普拉斯著的《概率论的分析理论》是一部内容丰富的奠基性的代表作。20 世纪初,由于抽象测度论与积分理论的建立,苏联数学家柯尔莫哥洛夫著的《概率论基础》一书提出了概率的测度式的定义,使概率论建立在严谨的数学公理化的基础上,从而为近代概率论的发展打下了坚实的基础。

概率论是数学中一个有特色的分支,它有自己独特的概念与方法,又与其它数学分支有密切的联系;它既有严谨的理论,又有丰富的直观背景。当前,概率论除自身在理论和应用方面进一步发展,形成了当今非常活跃的随机过程和数理统计两门学科外,它的理论与方法还渗透到许多学科领域中,而且还和多种学科结合,形成许多有生命力的边缘学科。

【随机事件】在一个试验中,可能出现也可能不出现的事情,亦简称事件。例如,在掷两颗骰子的试验中,“两颗骰子都出现 3 点”;“出现总点数为 6”等都是随机事件。在一个试验中,每一个可能出现的结果均为事件,称之为基本事件;此外,还有许许多多其它事件,称之为复合事件。在上例中,前者为基本事件,后者为复合事件。特

别,在每次试验中,一定出现的事情称为必然事件,如在上例中,“出现总点数至少为 2”是必然事件;在每次试验中,一定不出现的事情称为不可能事件,如在上例中,“出现总点数多于 12”是不可能事件。

在试验中,如将每一个可能结果视为一个样本点,则由所有样本点组成的集合称之为样本空间或基本空间,记为 Ω 。这时,事件除如上面用语言叙述来表示外,还可用 Ω 中的某子集 A 来确切地表示。如上例中,试验结果用样本点 (m,n) 表示,其中 m 和 n 分别表示第一、二颗骰子出现的点数,此时,样本空间 $\Omega = \{m,n\}; 1 \leq m \leq 6, 1 \leq n \leq 6\}$ 。于是前面叙述的基本事件可用 Ω 中含一个样本点的子集 $A = \{(3,3)\}$ 表示。复合事件可用 Ω 中含 5 个样本点的子集 $B = \{(1,5), (5,1), (2,4), (4,2), (3,3)\}$ 表示。特别,必然事件用集合 Ω 表示;不可能事件用空集 $\emptyset$ 表示。

在一个试验中,有许许多多事件,分别用 $A, B, C, \dots$ 表示,它们之间有着各种各样的关系:(1)若 A 出现则 B 必然出现称为 B 包含 A ,记为 $A \subseteq B$;(2)若 A, B 相互包含,则称 A, B 等价,记为 $A = B$;(3)“ A, B 中至少有一出现”是一事件,称为和事件或 A 与 B 之和,记为 $A \cup B$;(4)“ A, B 同时出现”是一事件,称为积事件或 A 与 B 之积,记为 $A \cap B$ 或 AB ;(5)“ A 出现但 B 不出现”是一事件,称为差事件或 A 与 B 之差,记为 $A - B$;(6)“ A 不出现”是一事件,称为 A 的对立事件,记为 $\bar{A}$;(7) A 与 B 不能同时出现,称 A, B 互不相容,记为 $A \cap B = \emptyset$ 。如果事件用 Ω 的子集表示,则上述事件之间的关系与集合的相应关系对应。

【概率】表示随机事件出现的可能性大小或机会大小的一个量。

概率又称或然率、几率。随机事件 A 的概率记为 $P(A)$,它表

示在一次试验中,事件 A 出现的可能性有多大。如果 $P(A)=0.9$, 则表明在一次试验中,事件 A 有 90% 的可能性出现。值得注意的是,事件 A 的概率只说明 A 出现的可能性而已,并不能准确无误地预言实际的未来结果,不过在试验重复进行次数愈多的情况下,实际上,事件 A 出现的次数所占的比例愈易接近概率。在实际生活中遇到的如“次品率”、“命中率”等都是概率。

概率有四种定义方式:

(1) 概率的古典定义 只适用于试验的所有可能结果为有限个且出现的可能性均等的情形。这种试验称为古典概型。这时,样本空间 Ω 由有限个样本点组成,其个数记为 n 。若事件 A 含 m 个样本点,则定义 A 的概率 $P(A)=\frac{m}{n}$ 。这个定义在 19 世纪由法国数学家拉普拉斯明确地给出。

(2) 概率的几何定义 只适于试验的所有可能结果能与一个测度不为 0 的几何区域 G 中的点建立一一对应且出现的可能性均等的情形,其中测度一词在一维空间指长度、二维空间指面积、三维空间指体积等等。这种试验称为几何概型。这时,如果事件 A 可用 G 中的子区域 g 表示,则定义 $P(A)=\frac{\mu(g)}{\mu(G)}$,其中如 $\mu(G)$ 表示 G 的测度,显然,事件的概率只与表示事件的子区域的测度有关,而与其形状和位置无关。这定义在 1777 年由法国数学家蒲丰给出并用它解决了最典型的“投针问题”。

(3) 概率的统计定义 对同一个试验重复地进行 n 次,用 μ_A 表示事件 A 在 n 次试验中出现的次数,称为 A 的频数。如果当 n 充分大时, A 出现的频率 μ_A/n 常稳定地在某一数值 p 附近摆动,而且随着试验次数 n 的增加,摆动的幅度常常变小,与 p 偏离大的情况越来越稀少,这称之为“频率的稳定性”。这时,则定义事件 A

的概率为 p , 即 $P(A)=p$, 此定义基于任何事件出现的频率都具有稳定性, 瑞士数学家伯努利从理论上给出了证明(参见大数定律)。应用这个定义时, 虽对试验没什么限制, 但求概率的方法却不够严谨, 通常只能在重复做大量试验后, 以事件出现的频率作为其概率的近似值。

(4) 概率的公理化定义 上述概率的定义都各有局限性。1933年, 苏联数学家柯尔莫哥洛夫建立了概率的测度式的定义, 它既建立在严谨的数学公理化的基础上又概括了概率的种种定义。

设样本空间 Ω 是一个任意给定的非空集合, 其中元素用 ω 表示。设 F 是由 Ω 的某些子集构成的集类, 如果 F 满足条件: ① $\Omega \in F$; ② 如果 $A \in F$, 则 $\bar{A} \in F$; ③ 如果 $A_i \in F, i=1, 2, \dots$, 则 $\bigcup_{n=1}^{\infty} A_n \in F$, 那么, 称 F 为 σ 域或 σ 代数, 亦称事件域。凡属 F 的 Ω 子集 A 均称为事件, 不属于 F 的 Ω 子集不是事件。一般称 (Ω, F) 为可测空间, 在它的基础上给出下面概率的公理化定义:

设 P 是定义在 F 上的实值函数, 即对每一 $A \in F$, 有唯一实数 $P(A)$ 与之对应, 且满足公理: ① 非负性 对一切 $A \in F, P(A) \geq 0$; ② 规范性 $P(\Omega)=1$; ③ 可列可加性(σ 可加性) 若 $A_i \in F, i=1, 2, \dots$, 且两两互不相容, 则有 $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ 成立, 则称 P 为 (Ω, F) 上的概率测度, 简称概率。特别, 对任何 $A \in F$, 称 $P(A)$ 为事件 A 的概率。

称三元总体 (Ω, F, P) 为概率空间。对任何一个具体的试验都可构造出一个概率空间。现代概率论的发展建立在概率空间的基础上。

概率的基本性质有: (1) 事件 A 的概率为介于 0 与 1 之间的一个数, 即 $0 \leq P(A) \leq 1$; (2) 必然事件的概率定为 1, 即 $P(\Omega)=1$ 。

不可能事件的概率定为 0, 即 $P(\emptyset)=0$; (3) 概率具有可加性(加法公式), 即若 $A_1, A_2, \dots, A_n$ 两两互不相容, 则 $P(\bigcup_{k=1}^n A_k) = \sum_{k=1}^n P(A_k)$; (4) $P(A) = 1 - P(\bar{A})$; (5) 若 $A \subseteq B$, 则 $P(B-A) = P(B) - P(A)$; (6) 一般加法公式:

$$P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} A_{i_2}) + \dots + (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} A_{i_2} A_{i_k}) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n);$$

(7) 概率的连续性 上连续性: 若 $A_1 \supseteq A_2 \supseteq \dots$, 则 $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$ 。下连续性: 若 $A_1 \subseteq A_2 \subseteq \dots$, 则

$$P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)。$$

【条件概率】任何一个试验都是在一定条件 S 下进行的, 通常都不提及。如果在除条件 S 外, 再附加“事件 A 已发生”($P(A) > 0$) 的条件下进行试验时, 若考虑某事件 B 的概率, 则称这概率为事件 B 在“事件 A 已发生”的条件下的条件概率, 简称事件 B 关于 A 的条件概率, 记为 $P(B|A)$ 。通常事件 B 的概率 $P(B)$ 与 $P(B|A)$ 不同, 必要时称 $P(B)$ 为无条件概率。在一般条件下, 条件概率与无条件概率有如下关系: $P(B|A) = P(AB)/P(A)$ (由上例可验证此关系)。因此, 人们以此关系作为条件概率的定义: 在概率空间 (Ω, F, P) 上, 若 $P(A) > 0$, 对一切 $B \in F$, 令 $P(B|A) = P(AB)/P(A)$, 则如此定义的 F 上的函数, 记为 $P(\cdot | A)$, 它满足概率的公理化定义中的三个条件, 故 $P(\cdot | A)$ 是概率, 并称之为在“事件 A 已发

生”的条件下的条件概率,简称关于 A 的条件概率。若 $B \in F$, 则称 $P(B|A)$ 为事件 B 关于 A 的条件概率。

关于条件概率有三个重要公式:

(1) 概率的一般乘法公式 若

$P(A_1 A_2 \cdots A_{n-1}) > 0, n \geq 2$, 则 $P(A_1 A_2 \cdots A_{n-1} A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_2 A_1) \cdots P(A_n | A_1 A_2 \cdots A_{n-1})$ 。

(2) 全概率公式 若 $A_1, A_2, \cdots, A_n$ 两两互不相容, $P(\bigcup_{k=1}^n A_k) =$

1 且 $P(A_k) > 0$ 对一切 k 成立, 则对任一事件 B , 有 $P(B) = \sum_{k=1}^n P(A_k) P(B|A_k)$ 成立。

(3) 贝叶斯公式 在(2)的条件下, 对满足 $P(B) > 0$ 的事件 B 有

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{m=1}^n P(A_m)P(B|A_m)}.$$

(2), (3) 公式可推广到 $A_1, A_2, \cdots, A_n \cdots$ 的情形。

【事件的独立性】两个事件 A, B 相互独立, 直观地讲, 是指两事件出现与否相互之间无影响。严格地讲, 是指其中一事件出现对另一事件出现的概率无影响, 即 $P(B|A) = P(B)$ 或 $P(A|B) = P(A)$ 。更一般地, 定义为: 如果 $P(AB) = P(A)P(B)$, 则称事件 A, B 相互独立。例如, 在甲、乙两射手对同一目标各进行一次射击的试验中, 事件“甲击中目标”和事件“乙击中目标”是相互独立的, 因为甲、乙击中目标与否相互无影响, 完全由他们各自的命中率决定。

对于多个事件 $A_1, A_2, \cdots, A_n (n \geq 3)$ 相互独立, 直观地讲, 也是指它们之间发生与否相互无影响。严格地讲, 如果从 $A_1, A_2, \cdots$,

A_n 中任取一组事件 $A_{i_1}, A_{i_2}, \dots, A_{i_k} (k \geq 2)$, 皆有 $P(A_{i_1}, A_{i_2}, \dots, A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$, 则称事件 $A_1, A_2, \dots, A_n$ 相互独立。事件个数越多, 它们相互独立的条件要求上述等式成立的个数越多。例如, 三个事件相互独立不能由它们中两两事件相互独立得到。

独立性的概念是概率论中特有的概念, 它的引进大大推动了概率论的发展, 概率论中有许多结论是在独立性的假设下获得的。在解决实际问题时, 常常凭直观经验来判断一组事件 $A_1, A_2, \dots, A_n$ 是否独立, 因为如果独立的话, 求这一组事件乘积的概率就可等于各自概率的乘积, 即 $P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2) \cdots P(A_n)$ 。

【 n 重伯努利试验】相当于一个简单试验 E 。独立地重复进行 n 次 (各次试验出现什么结果相互无影响), 试验 E_0 只有两个可能结果, 分别记为 A 和 $\bar{A}$, 已知 $P(A) = p, P(\bar{A}) = 1 - p = q$, 其中 E_0 被称为伯努利试验。

在这种试验中, 每一个可能出现的结果均可表为 $A'_1 A'_2 \cdots A'_n$, 其中 A'_i 表示第 i 次 E_0 试验出现的结果, 故 A'_i 为 A 或 $\bar{A}$ 。显然, $A'_1, A'_2, \dots, A'_n$ 相互独立, 因此, $P(A'_1 A'_2 \cdots A'_n) = P(A'_1)P(A'_2) \cdots P(A'_n) = p^k q^{n-k}$, 其中 k 为 $A'_1 A'_2 \cdots A'_n$ 中 A 的个数。由此, 事件“在 n 次试验中 A 恰好出现 k 次”的概率 $P_n(k) = \binom{n}{k} p^k q^{n-k}$ 。

【随机变量】表示随机试验结果的量。它是一个至少可取两个值的变量, 但每次试验后它取什么值事先不能预知, 由试验结果而定, 因此, 可将它看作样本空间 Ω 上的一个函数。随机变量在某个范围内取值通常应表示某个事件。由此, 随机变量严格的数学定义应

如下: 设 X 是样本空间 Ω 上的实函数, 即 $X=X(\omega)$, 对一切 $\omega \in \Omega$, 如果对一切实数 X , 均有 $\{X \leq X\} = \{\omega: X(\omega) \leq X\} \in F$ (事件域), 即 $\{X \leq X\}$ 是事件, 则称 X 是 (Ω, F) 上的一个随机变量。又称 X 为可测空间 (Ω, F) 上的可测函数。而在概率论中, 常说 X 是概率空间 (Ω, F, P) 上的随机变量。

设 $X_1, X_2, \dots, X_n$ 是概率空间 (Ω, F, P) 上的 n 个随机变量, 称 $X=(X_1, X_2, \dots, X_n)$ 为 n 维随机向量, 称 $X_k (1 \leq k \leq n)$ 为它的第 k 个分量。特别, 随机变量亦可称为一维随机向量。

设 X, Y 是概率空间 (Ω, F, P) 上的随机变量, 称 $X+iY$ 为复随机变量, 其中 i 为虚数单位。

【随机变量的概率分布】描述随机变量取值的概率规律一般用分布函数, 但对特殊的随机变量还有更方便的方式。

设 X 为随机变量, 对任意实数 X , 令 $F(X)=P(X \leq X)$, 其中 $P(X \leq X)$ 表示事件 $\{X \leq X\}$ 的概率, 称如此定义在 R (全体实数) 上的函数 $F(X)$ 为随机变量 X 的分布函数。由概率性质可知, 任何分布函数 $F(X)$ 都满足下面三个条件: (1) 单调非降, 即当 $a < b$ 时, $F(a) \leq F(b)$; (2) $\lim_{X \rightarrow -\infty} F(X) = 0, \lim_{X \rightarrow +\infty} F(X) = 1$; (3) 右连续, 即

$\lim_{X \rightarrow a^+} F(X) = F(a)$, 其中 $X \rightarrow a^+$ 表示 $X > a$ 且趋于 a 。反之, 任一满足

上面三个条件的函数定是某个随机变量的分布函数。

分布函数刻划了随机变量取值的概率规律, 由它可得到随机变量在某个范围取值的概率。

取有限多个或可列多个值的随机变量称为离散型随机变量。凡表示离散型随机变量 X 可能取的值和取各个值的概率的各种形式都称之为 X 的分布列。设 X 可能取的值为 $X_1, X_2, \dots, X_n, \dots$

(包括取有限值的情况), 称 $p_k = P(X = X_k), k = 1, 2, \dots$ 为随机变量 X 的分布列, 简记为 $\{p_k\}$ 。由概率性质知, 任一分布列 $\{p_k\}$ 都满足条件: (1) $p_k \geq 0$; (2) $\sum_{k=1}^{\infty} p_k = 1$ 。反之, 任一满足 (1), (2) 条件的数列 $\{p_k\}$ 定是某个离散型随机变量的分布列。用分布列刻画离散型随机变量取值的概率规律比用它的分布函数来得简单、方便。

设 X 为一随机变量, 如果存在一个非负实函数 $p(t)$, X 在任一形如 $(-\infty, X)$ 的区间上取值的概率 $P(X \leq X)$ 可用 $p(t)$ 在这区间上的积分表示, 即

$$P(X \leq X) = F(X) = \int_{-\infty}^X p(t) dt,$$

其中 X 为任一实数, 则称 X 为连续型随机变量, 称 $p(t)$ 为 X 的概率分布密度函数, 简称密度函数。这种随机变量 X 在任何区间 $[a, b]$ 上取值的概率可用它的密度函数 $p(t)$ 在区间上的积分表示, 即

$$P(a \leq X \leq b) = \int_a^b p(t) dt,$$

它又可用图 1 中带阴影部分的面积表示。显然, X 取任何一个值的概率都等于 0, 即 $P(X = a) = 0, a$ 为任一实数。连续型随机变量 X 的密度函数 $p(X)$ 常由它的分布函数求得, 因为在一定条件下, $p(X) = F'(X)$ 。

由概率性质知, 密度函数 $p(X)$ 满足条件: (1) 对一切实数 $X, p(X) \geq 0$; (2) $\int_{-\infty}^{\infty} p(X) dX = 1$, 反之, 满足条件 (1), (2) 的函数定是某个连续型随机变量的密度函数。

【随机向量的联合分布与边缘分布】随机向量 $X = (X_1, X_2, \dots, X_n)$ 的概率分布称为 X 的联合分布或 $X_1, X_2, \dots$,

图 1 连续型随机变量 X 在 $[a, b]$ 上取值的概率

$\cdots, X_n$ 的联合分布。由 X 中的任意 $k (< n)$ 个分量作成它的子随机向量, 如 $(X_{i_1}, X_{i_2}, \cdots, X_{i_k})$, 它的概率分布称为 X 的联合分布的边缘分布。具体表示形式如下:

(1) 联合分布函数与边缘分布函数 设 $X = (X_1, X_2, \cdots, X_n)$ 为 n 维随机向量, 对任一 n 维实向量 $(X_1, X_2, \cdots, X_n)$, 令 $F(X_1, X_2, \cdots, X_n) = P(X_1 \leq X_1, X_2 \leq X_2, \cdots, X_n \leq X_n) = P(\bigcap_{m=1}^n \{X_m \leq X_m\})$, 则称如此定义的 n 元实函数 $F(X_1, X_2, \cdots, X_n)$ 为 X 的联合分布函数或 $X_1, X_2, \cdots, X_n$ 的联合分布函数。在不至于混淆的情况下, 可省略联合二字。联合分布函数具有和随机变量的分布函数类似的性质和结论。

X 的任一子随机向量, 如 $(X_{i_1}, X_{i_2}, \cdots, X_{i_k}), k < n$, 有如上方法定义的分布函数, 相对于 X 的联合分布函数称它为 $F(X_1, X_2, \cdots, X_n)$ 的边缘分布函数, 任何一个边缘分布函数都可由联合分布函数确定。

(2)联合分布列与边缘分布列 只取有限个或可列个向量值的随机向量称为离散型随机向量。设 $X=(X_1, X_2, \dots, X_n)$ 为离散型随机向量, X 取值 $(X_{i_1}, X_{i_2}, \dots, X_{i_n})$ 的概率记为 $P(X_1=X_{i_1}, X_2=X_{i_2}, \dots, X_n=X_{i_n})$ 。称 $p_{i_1, i_2, \dots, i_n} = P(X_1=X_{i_1}, X_2=X_{i_2}, \dots, X_n=X_{i_n}) = p(\bigcap_{m=1}^n \{X_m=X_{i_m}\})$, $i_m=1, 2, \dots, m=1, 2, \dots, n$, 为 X 的联合分布列, 简记 $\{p_{i_1, i_2, \dots, i_n}\}$ 。它表示了 X 可能取的值和取各个可能值的概率。有时也可省略地称它为分布列。

X 的任一子随机向量, 如 $(X_{k_1}, X_{k_2}, \dots, X_{k_l})$, $l < n$, 亦是离散型随机向量, 它的分布列称为联合分布列 $\{p_{i_1, i_2, \dots, i_n}\}$ 的边缘分布列。由联合分布列可求出任意一个边缘分布列。

(3)联合密度函数与边缘密度函数 设 n 维随机向量 $X=(X_1, X_2, \dots, X_n)$ 的分布函数为 $F(X_1, X_2, \dots, X_n)$, 如果存在一个非负的 n 元函数 $p(t_1, t_2, \dots, t_n)$, 对任一 n 维实向量 $(X_1, X_2, \dots, X_n)$ 都有

$$F(X_1, X_2, \dots, X_n) = \int_{-\infty}^{X_1} \int_{-\infty}^{X_2} \cdots \int_{-\infty}^{X_n} p(t_1, t_2, \dots, t_n) dt_1 dt_2 \cdots dt_n$$

成立, 则称 X 为连续型随机向量, 称 $p(t_1, t_2, \dots, t_n)$ 为 X 的联合密度函数, 有时, 就称密度函数。 X 在 n 维区域 G 上取值的概率可表为它的密度函数

$p(t_1, t_2, \dots, t_n)$ 在区域 G 上的积分, 即

$$P((X_1, X_2, \dots, X_n) \in G) = \int_G \cdots \int p(t_1, t_2, \dots, t_n) dt_1 dt_2 \cdots dt_n$$

X 的任一子随机向量, 如 $(X_{i_1}, X_{i_2}, \dots, X_{i_k})$, $k < n$ 的密度函数称为联合密度函数 $p(t_1, t_2, \dots, t_n)$ 的边缘密度函数。边缘密度函数均可由联合密度函数求得。

【条件分布】设 X, Y 为随机变量, 在 $Y=y$ 的条件下, X 的概率分布定义为 X 关于 $Y=y$ 的条件分布。条件分布具有随机变量概率分布相同的性质。例如, 设 (X, Y) 为二维离散型随机向量, X 可能取值 $X_1, X_2, \dots$, Y 可能取值 $y_1, y_2, \dots$ 。在 $Y=y_j$ 的条件下, X 的分布列称为关于 $Y=y_j$ 的条件分布列, 记为 $p_{i|j} = P(X=X_i | Y=y_j)$, $i=1, 2, \dots$ 。若 (X, Y) 的联合分布列为 $\{p_{ij}\}$, 则由条件概率知 $p_{k|j} = \frac{P(X_k=X_k, Y=y_j)}{P(Y=y_j)} = \frac{p_{kj}}{\sum_{i=1}^{\infty} p_{ij}}$, $k=1, 2, \dots$ 。 $\{p_{ij}\}$ 具有通常分布列的性质: (1) $p_{i|j} \geq 0$; (2) $\sum_{i=1}^{\infty} p_{i|j} = 1$ 。

【随机变量的独立性】称随机变量 $X_1, X_2, \dots, X_n$ 相互独立, 指对任何实数 $X_1, X_2, \dots, X_n$, 事件 $\bigcap_{m=1}^n \{X_m \leq X_m\}$ 的概率等于事件 $\{X_m \leq X_m\}$, $m=1, 2, \dots, n$ 的概率的乘积, 即 X 的联合分布函数 $F(X_1, X_2, \dots, X_n) = F_1(X_1)F_2(X_2) \cdots F_n(X_n)$, 其中 $F_m(X_m)$ 是 X_m 的分布函数。特别, (1) 当 $X=(X_1, X_2, \dots, X_n)$ 为离散型随机向量时, 如果所有事件 $\bigcap_{m=1}^n \{X_m \leq X_m\}$ 的概率都等于事件 $\{X_m = X_m\}$, $m=1, 2, \dots, n$, 的概率的乘积, 则称 $X_1, X_2, \dots, X_n$ 相互独立; (2) 当 $X=(X_1, X_2, \dots, X_n)$ 为连续型随机向量时, 如果它的所有分量的密度函数相乘为 X 的联合密度函数时, 则称 $X_1, X_2, \dots, X_n$ 相互独立。在解决实际问题时, 常常凭经验来判定一些随机变量取值相互之间无影响, 从而认为它们是相互独立的。

【随机变量的函数的概率分布】一个或多个随机变量的初等函数或连续函数仍是随机变量,它的概率分布有的可由随机变量的概率分布求出。例如,

(1)线性函数 设 X 的分布函数为 $F(X)$, $Y=aX+b$ ($a>0$, b 均为实数) 亦为随机变量, 它的分布函数

$F_Y(y)=F(\frac{y-b}{a})$ 。特别, 当 X 为连续型随机变量时, $p(X)$ 为它的密度函数, 这时 Y 的密度函数 $p_Y(y)=$

$\frac{1}{|a|}p(\frac{y-b}{a})$, 这里线性函数中的 a 只要求不等于 0。

(2)和 设 (X_1, X_2) 的联合密度函数为 $p(X_1, X_2)$, 和 $X=X_1+X_2$ 是一随机变量, 它的密度函数

$$p_X(X) = \int_{-\infty}^{\infty} p(X-z, z) dz \\ = \int_{-\infty}^{\infty} p(z, X-z) dz。$$

(3)商 设 (X_1, X_2) 的联合密度函数为 $p(X_1, X_2)$, 商 $Y=\frac{X_1}{X_2}$ ($X_2 \neq 0$) 是一随机变量, 它的密度函数

$$p_Y(X) = \int_{-\infty}^X p(Xz, z) |z| dz。$$

【数学期望】又称期望或均值, 数学定义如下:

(1) X 为离散型随机变量, 设 X 的分布列为 $p_k=P(X=X_k)$, $k=1, 2, \dots$, 若级数 $\sum_{k=1}^{\infty} |X_k| p_k < \infty$ 时, 则称 X 的期望存在且期望

$$EX = \sum_{k=1}^{\infty} X_k p_k。$$

(2) X 为连续型随机变量, 设 X 的密度函数为 $p(X)$, 若 $\int_{-\infty}^{\infty} |X| p(X) dX < \infty$ 时, 则称 X 的期望存在且期望 $EX = \int_{-\infty}^{\infty} X p(X) dX$ 。

(3) X 为一般的随机变量(包括离散型和连续型), $F(X)$ 为它的分布函数。若 $\int_{-\infty}^{\infty} |X| dF(X) < \infty$ 时, 则称 X 的期望存在且期望 $EX = \int_{-\infty}^{\infty} X dF(X)$, 其积分是黎曼-斯蒂尔杰斯积分(参见实变函数)。

随机变量的期望有下列性质(设 X, Y 期望存在): (1) 常数 a 看成随机变量时, $Ea = a$; (2) $E(X+Y) = EX + EY$; (3) 若 $a \leq X \leq b$, 则 $a \leq EX \leq b$; (4) 若 X, Y 独立, 则 $E(XY) = EX \cdot EY$ 。

随机变量的函数的期望有如下公式:

(1) 随机变量 X 的函数 $f(X)$ 的期望可用 X 的概率分布直接计算:

$$Ef(X)$$

$$= \begin{cases} \sum_{k=1}^{\infty} f(X_k) p_k & X \text{ 为离散型;} \\ \int_{-\infty}^{\infty} f(X) p(X) dX, & X \text{ 为连续型,} \end{cases}$$

其中 $\{p_k\}, p(X)$ 分别为 X 的分布列和密度函数。

(2) 随机变量 X, Y 的函数 $f(X, Y)$ 的期望可由 (X, Y) 的联合分布直接计算:

$$Ef(X, Y)$$

$$= \begin{cases} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} f(X_k, X_j) p_{k,j}, \\ (X, Y) \text{ 为离散型;} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(X, y) p(X, y) dX dy, \\ (X, Y) \text{ 为连续型,} \end{cases}$$

其中 $\{p_{k,j}\}$, $p(X, y)$ 分别为 (X, Y) 的联合分布列和联合密度函数。

若复随机变量 $Z = X + iY$, 则定义其数学期望 $EZ = EX + iEY$ 。

若随机向量 $X = (X_1, X_2, \dots, X_n)$, 则 X 的数学期望定义为 $EX = (EX_1, EX_2, \dots, EX_n)$, 又称它为 X 的均值向量。

【分位数】给定 $0 < \alpha < 1$, 如果数 X_α 满足条件 $P(X \leq X_\alpha) \geq 1 - \alpha$ 和 $P(X \geq X_\alpha) \geq \alpha$, 则称 X_α 为 X 的 α 分位数, 又称 $X_{1-\alpha}$ 为 X 的下 α 分位数。特别, 当 $\alpha = \frac{1}{2}$ 时, 称分位数 $X_{\frac{1}{2}}$ 为 X 的中位数, 它刻划了 X 的概率分布的中心位置, 对任何随机变量都有中位数, 但可能不唯一。当随机变量期望不存在时, 它可起到类似期望的作用。它与期望比, 主要优点是受极端值的影响较小, 因此, 在一些应用统计问题中常以它代替期望。

【方差】刻划随机变量取值与期望值平均偏离程度的一个量。若 EX 存在, 则定义 $E(X - EX)^2$ 为 X 的方差, 记作 DX 或 $VarX$ 。称 $\sqrt{DX}$ 为标准差或均方根差; 称 $\sqrt{DX}/EX$ 为变异系数。

方差的计算可归为计算随机变量 X 的函数 $(X - EX)^2$ 的期望, 由随机变量的函数期望公式知, 它完全决定于 X 的概率分布。

此外,还有 $DX = EX^2 - (EX)^2$ 。

方差的主要性质如下;(1)常数 a 看成随机变量时, $Da = 0$;
(2) $D(aX) = a^2DX$, a 为任意常数;(3)当 X, Y 独立时, $D(X+Y) = DX + DY$;(4)对任意实数 a , $E(X-a)^2$ 在 $a=EX$ 处达到最小值, 即 $DX = E(X-EX)^2 \leq E(X-a)^2$, 对一切实数 a ;(5)若 $DX=0$, 则 X 以概率 1 等于 EX , 即 $P(X=EX)=1$ 。

若复随机变量 $Z=X+iY$, 则定义 Z 的方差 $DZ=E(Z-EZ)\overline{(Z-EZ)}=E|Z-EZ|^2$, 其中 $\bar{Z}$ 是 Z 的共轭复值 $X-iY$ 。

【 k 阶原点矩和 k 阶中心矩】对任一整数 k , 若 EX^k 和 $E(X-EX)^k$ 存在, 则分别称为 X 的 k 阶原点矩和 k 阶中心矩。期望是一阶原点矩, 方差是二阶中心矩。 k 阶原点矩和 k 阶中心矩可用 X 的函数 X^k 和 $(X-EX)^k$ 的期望公式分别算出。矩在描述和确定概率分布时起重要作用。

【协方差与相关系数】刻划两个随机变量 X, Y 之间线性联系程度的量。定义 $E(X-EX)(Y-EY)$ 为 X, Y 的协方差, 记为 $cov(X, Y)$ 。它的计算可用 X, Y 的函数 $(X-EX)(Y-EY)$ 的期望公式。此外, 还有 $cov(X, Y) = E(XY) - EX \cdot EY$ 。

协方差具有以下性质:(1) $cov(X, Y) = cov(Y, X)$;(2) $cov(aX, bY) = abcov(X, Y)$, 其中 a, b 为任意实常数;(3) $cov(X_1 + X_2, Y) = cov(X_1, Y) + cov(X_2, Y)$ 。

定义 $\frac{cov(X, Y)}{\sqrt{DX \cdot DY}}$ 为 X, Y 的相关系数, 记为 ρ_{XY} 。显然 $|\rho_{XY}| \leq 1$ 。

$|\rho_{XY}|$ 的值越接近 1, 它表明 X, Y 之间线性联系越好。当 $|\rho_{XY}| = 1$ 时, X, Y 之间就以概率 1 有线性关系 $Y=aX+b$, 即 $P(Y=aX+b)=1$ 。

$b)=1$; 当 $\rho_{XY}=0$ 时, 称 X, Y 不相关或零相关, 即表示 X, Y 之间没有线性关系, 但不一定无其它关系。若 X, Y 独立, 则 $\rho_{XY}=0$; 反之不成立。

设 n 维随机向量 $X=(X_1, X_2, \dots, X_n)$, 若令 $\sigma_{ij}=\text{cov}(X_i, X_j)$, 则称 $n \times n$ 矩阵。

$$\begin{aligned} \Sigma &= E(X-EX)'(X-EX) \\ &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \end{aligned}$$

为 X 的协方差阵, 亦称方差阵, 其中 $(X-EX)'$ 是 $(X-EX)$ 的转置矩阵

若 Z_1, Z_2 是两个复随机变量, 则定义它们的协方差 $\text{cov}(Z_1, Z_2)=E(Z_1-EZ_1)\overline{(Z_2-EZ_2)}$

【条件期望】由条件分布确定的期望。例如, 当 (X, Y) 为离散型随机向量时, X 关于 $Y=y_j$ 的条件分布列为 $p_{k|j} =$

$P(X=X_k | Y=y_j), k=1, 2, \dots$, 则以它为分布列求得的期望称为 X 关于 $Y=y_j$ 的条件期望, 记作 $E(X | Y=y_j)$, 即

$$\begin{aligned} E(X | Y=y_j) &= \sum_{k=1}^{\infty} X_k p_{k|j} \\ &= \sum_{k=1}^{\infty} X_k P(X=X_k | Y=y_j). \end{aligned}$$

一般地说, 条件期望 $E(X | Y=y)$ 依赖于 Y 的取值, 因此, 它是 Y 的函数, 通常将此函数记为 $E(X | Y)$, 并称为 X 关于 Y 的条件期望。这时, 它是一个随机变量。在很多情况下, 无法用初等的方法来求 $E(X | Y)$, 但由测度论的知识已证明: 在 EX 存在的条件

下, $E(X|Y)$ 存在, 虽不唯一, 但彼此只在一个零概率事件上有差异。条件期望除有通常期望的性质外, 还有 $E(E(X|Y)) = EX$ 。

类似地, 还可定义 X 关于 $Y_1, Y_2, \dots, Y_n$ 的条件期望, 记为 $E(X|Y_1, Y_2, \dots, Y_n)$ 。更一般地, 还可定义 X 关于某个 σ 代数 B (若随机变量定义在 (Ω, F, P) 上, B 应是 F 的一个子 σ 代数, 即 $B \subset F$) 上的条件期望, 记为 $E(X|B)$ 。

【特征函数】 设 X 为随机变量, 称 Ee^{iXt} , $-\infty < t < \infty$, 为 X 的特征函数, 其中 i 为虚数单位, 记为 $f(t)$ 。当 X 为离散型随机变量时, $f(t) = \sum_{k=1}^{\infty} e^{iX_k t} p_k$, 其中 $p_k = P(X = X_k)$, $k = 1, 2, \dots$ 为 X 的分布列; 当 X 为连续型随机变量时,

$f(t) = \int_{-\infty}^{\infty} e^{iXt} p(X) dX$, 其中 $p(X)$ 为 X 的密度函数, 这时, $f(t)$ 正是 $p(X)$ 的傅里叶变换; 当 X 为任一随机变量时, $f(t)$

$= \int_{-\infty}^{\infty} e^{iXt} dF(X)$, 其中积分是黎曼-斯蒂尔杰斯积分, $F(X)$ 是 X 的分布函数。此外, 还有 $Ee^{iXt} = E(\cos Xt) + iE(\sin Xt)$ 。

特征函数 $f(t)$ 具有下面性质: (1) $f(0) = 1$; (2) $|f(t)| \leq 1$; (3) $f(t)$ 在 $(-\infty, \infty)$ 上一致连续且具有非负定性, 即对任意实数 $t_1, t_2,$

$\dots, t_n$ 和任意复数 $z_1, z_2, \dots, z_n$ 有 $\sum_{j=1}^n \sum_{k=1}^n f(t_j - t_k) z_j \bar{z}_k \geq 0$; (4) 若 X

的特征函数为 $f(t)$, 则 $aX + b$ (a, b 为常数) 的特征函数 $f_{aX+b}(t) = e^{ibt} f(at)$; (5) 若 X_1, X_2 独立, 它们的特征函数分别为 $f_1(t)$ 和 $f_2(t)$, 则 $X_1 + X_2$ 的特征函数

$f_{X_1+X_2}(t) = f_1(t) \cdot f_2(t)$; (6) 若 X 的 k 阶矩存在, 则对一切 $r \leq k$,

$EX^r = f^{(r)}(0)/i^r$, 其中

$f^{(r)}(0)$ 表示 X 的特征函数 $f(t)$ 的 r 阶导数在 0 点的值。这性质把求 X 的 r 阶矩变成求它的特征函数的 r 阶导数。(7)特征函数与概率分布相互唯一决定。利用性质(5)和(7)可求出许多独立随机变量和的概率分布。一个定义在 $(-\infty, \infty)$ 上的复值函数, 如果它连续、非负定且 $f(0)=1$, 则它定是某个随机变量的特征函数。

设 $X=(X_1, X_2, \dots, X_n)$ 是一 n 维随机向量, 定义 $f(t_1, t_2, \dots, t_n)=e^{i(X_1t_1+X_2t_2+\dots+X_nt_n)}$ 为 X 的特征函数。

【母函数】设 X 是取非负整数值的随机变量, 分布列为 $p_k=P(X=k)$, $k=0, 1, 2, \dots$, 称级 $P(s)=\sum_{k=0}^{\infty} p_k s^k$, $-1 \leq s \leq 1$, 为 X 的母函数。母函数在 19 世纪初被拉普拉斯所引进, 它是特征函数出现的先导。它有许多类似于特征函数的性质。对取非负整数值的随机变量, 在很多场合下用它比用特征函数更方便。

【常见概率分布】由实际问题导出或在数理统计中常用的概率分布。

(1)单点分布 若随机变量 X 以概率 1 取常数 a , 即 $P(X=a)=1$, 则称 X 服从单点分布, $EX=a$, $DX=0$, 特征函数 $f(t)=e^{iat}$ 。

(2)两点分布 若随机变量 X 的分布列为 $P(X=k)=p^k q^{1-k}$, $k=0, 1$, 其中 $0 < p < 1$, $q=1-p$, 则称 X 服从参数为 p 的两点分布。描述只有两个结果的随机试验用服从两点分布的随机变量。 $EX=p$, $DX=pq$, 特征函数 $f(t)=q+pe^{it}$ 。

(3)二项分布 若随机变量 X 的分布列为 $b(k; n, p)=P(X=k)=\binom{n}{k} p^k q^{n-k}$, $k=0, 1, \dots, n$, 其中 $0 < p < 1$, $q=1-p$, $\binom{n}{k}=n!/(k!(n-k)!)$ 是从 n 个对象中任意选取 k 个的组合数, 则称 X

服从参数为 n, p 的二项分布或 X 服从 $B(n, p)$ 分布。它的命名来源于 $b(k; n, p)$ 恰好是 $(p+q)^n$ 的二项展开式中的第 $k+1$ 项。

若 X 服从 $B(n, p)$ 分布, 则 X 取值由 0 依次增大到 n 时, 其概率先增大后减小。当 $(n+1)p$ 不是整数时, X 取 $[(n+1)p]$ 值的概率最大, 其中记号 $[a]$ 表示不超过 a 的最大整数; 当 $(n+1)p$ 是整数时, X 取 $(n+1)p$ 和 $(n+1)p-1$ 两值的概率相等且最大。 $EX=np$, $DX=npq$, 特征函数 $f(t)=(q+pe^{it})^n$, 母函数 $P(s)=(q+ps)^n$ 。若 $X_1, X_2, \dots, X_n$ 为相互独立且同服从参数为 p 的两点分布, 则 $X=X_1+X_2+\dots+X_n$ 服从 $B(n, p)$ 分布。

(4) 泊松分布 若随机变量 X 的分布列为 $p(k; \lambda)=P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, $k=0, 1, \dots$, 其中 $\lambda>0$, 则称 X 服从参数为 λ 的泊松分布或 X 服从 $P(\lambda)$ 分布。这个分布是泊松研究二项分布的渐近公式提出来的。

若 X 服从 $P(\lambda)$ 分布, 则 X 取值由 0 依次增大时, 其概率先增大后减小。当 λ 不是整数时, X 取 $[\lambda]$ 值的概率最大; 当 λ 是整数时, X 取 λ 和 $\lambda-1$ 两值的概率相等且最大。 $EX=DX=\lambda$, 特征函数 $f(t)=e^{\lambda(e^{it}-1)}$, 母函数 $P(s)=e^{\lambda(s-1)}$ 。

若随机变量 $X_1, X_2, \dots, X_n$ 相互独立, X_k 服从 $P(\lambda_k)$ 分布, 则 $X=X_1+X_2+\dots+X_n$ 服从参数为 $\sum_{k=1}^n \lambda_k$ 的泊松分布。

(5) 几何分布 若随机变量 X 的分布列为 $p_k=P(X=k)=pq^k$, $k=0, 1, \dots$, 其中 $0<p<1, q=1-p$, 则称 X 服从参数为 p 的几何分布。 $EX=qp^{-1}$, $DX=qp^{-2}$, 特征函数 $f(t)=p(1-qe^{it})^{-1}$, 母函数 $P(s)=p(1-qs)^{-1}$ 。

(6) 超几何分布 若随机变量的分布列为 $p_k=P(X=k)=$

$\binom{M}{k} \binom{N-M}{n-k} / \binom{N}{n}$, $\max(0, n+M-N) \leq k \leq \min(n, M)$, 其中 N, M, n 为正整数, $M \leq N, n \leq N$, 称则 X 服从参数为 N, M, n 的超几何分布。 $EX = \frac{nM}{N}$, $DX = \frac{nM}{N} \cdot \left[1 - \frac{M}{N}\right] \frac{\binom{N-n}{N-1}}{\binom{N-1}{N-1}}$ 。

(7) 负二项分布(帕斯卡分布) 若随机变量 X 的分布列为 $f(k; r, p) = P(X=k) = \binom{-r}{k} p^r (-q)^k = \binom{r+k-1}{k} p^r q^k, k=0, 1, \dots$, 其中 r 为正整数, $0 < p < 1, q = 1-p$, 则称 X 服从参数为 r, p 的负二项分布。

(8) 均匀分布 若随机变量的密度函数

$$p(X) = \begin{cases} 1/(b-a), & X \in [a, b]; \\ 0, & X \notin [a, b], \end{cases}$$

其中 $a < b$ 且均为常数, 则称 X 在 $[a, b]$ 上均匀分布或 X 服从 $U[a, b]$ 分布。 X 的密度函数曲线见图一。 $EX = \frac{a+b}{2}$, $DX = \frac{(b-a)^2}{12}$, 特征函数 $f(t) = (e^{ibt} - e^{iat}) / i(b-a)t$ 。

若二维随机向量 $X = (X_1, X_2)$ 在平面区域 G 上均匀分布, 则 X 的联合密度函数为

$$p(X) = \begin{cases} 1/G \text{ 的面积}, & (X, y) \in G; \\ 0, & (X, y) \notin G. \end{cases}$$

图 1 均匀分布的密度函数

(9) 指数分布 若随机变量 X 的密度函数 $p(X) = \begin{cases} \lambda e^{-\lambda x}, & X > 0; \\ 0, & X \leq 0, \end{cases}$ 其中 $\lambda > 0$ 为常数, 则称 X 服从参数为 p 的指数分布。

密度函数曲线见图二。在很多情况下, 表示寿命, 如灯管寿命的随机变量就服从这种分布。 $EX = \lambda^{-1}$, $DX = \lambda^{-2}$, 特征函数 $f(t) = \left[1 - \frac{it}{\lambda}\right]^{-1}$ 。

图 2 $\lambda = 0.5$ 的指数分布的
密度函数曲线

(10) 正态分布 又称高斯分布。若随机变量 X 的密度函数 $p(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$, $-\infty < X < \infty$

其中 μ 为实数, $\sigma > 0$, 则称 X 服从参数为 μ, σ^2 的(一维)正态分布或 X 服从 $N(\mu, \sigma^2)$ 分布, 又称 X 为一维正态变量。正态分布的概率密度函数曲线如钟形, 见图三。 $EX = \mu, DX = \sigma^2$, 特征函数 $f(t) = e^{(i\mu t - \frac{1}{2}\sigma^2 t^2)}$ 。

$$\sigma_1 < \sigma_2 < \sigma_3$$

特别, 当 $\mu = 0, \sigma^2 = 1$ 时, X 服从 $N(0, 1)$ 分布, 这时, 密度函数 $\varphi(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}}, -\infty < X < \infty$, 称 X 为标准正态变量。

图 3 正态分布的密度函数曲线

$$\text{令 } \Phi(X) = \int_{-\infty}^X \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

它表示标准正态变量 X 取值不超过 X 的概率。对任意 $X \geq 0$, 可查正态分布表得到 $\Phi(X)$ 的值, 对 $X < 0$, 可利用公式 $\Phi(X) = 1 - \Phi(-X)$, 其中 $\Phi(-X)$ 可查表得到, 从而可计算出 $\Phi(X)$ 的值。若 X 服从 $N(\mu, \sigma^2)$ 分布, 则可利用公式 $F(X) = P(X \leq X) = \Phi(\frac{X - \mu}{\sigma})$ 查表得到 $F(X)$ 的值。

若 X 服从 $N(\mu, \sigma^2)$ 分布, 则 X 的线性函数 $aX + b$ ($a \neq 0, b$ 均为任意实数) 服从 $N(a\mu + b, a^2\sigma^2)$ 分布, 特别, $\frac{X - \mu}{\sigma}$ 服从 $N(0, 1)$ 分

布。

若 $X_1, X_2, \dots, X_n$ 相互独立, $X_k (1 \leq k \leq n)$ 服从 $N(a_k, \sigma_k^2)$ 分布, 则它们的和 $X_1 + X_2 + \dots + X_n$ 服从 $N\left(\sum_{k=1}^n a_k, \sum_{k=1}^n \sigma_k^2\right)$ 分布。

若二维随机向量 $X = (X_1, X_2)$ 的联合密度函数

$$p(X, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ = eXp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(X-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(X-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

$-\infty < X < \infty, -\infty < y < \infty$, 其中 μ_1, μ_2 为实数, $\sigma_1 > 0, \sigma_2 > 0, |\rho| \leq 1$, 则称 X 服从参数为 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$ 的(二维)正态分布或 X 服从 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 分布。又称 X 为二维正态变量, 均值向量 $EX =$

$$(\mu_1, \mu_2), \text{协方差阵 } \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}。$$

(11)伽玛(Γ)分布 若随机变量 X 的密度函数

$$p(X) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} X^{\alpha-1} e^{-\beta X}, & X > 0; \\ 0, & X \leq 0, \end{cases}$$

其中 $\alpha > 0, \beta > 0$ 且均为常数, 则称 X 服从参数为 α, β 的伽玛分布或称 X 服从 $\Gamma(\alpha, \beta)$ 分布。密度函数曲线见图四。 $EX = \frac{\alpha}{\beta}, DX = \frac{\alpha}{\beta^2}$, 特征函数

$$f(t) = (1 - \frac{it}{\beta})^{-\alpha}。$$

图 4 伽玛分布的密度函数曲线

当 $\alpha=1$ 时,伽玛分布为指数分布;当 $\alpha=\frac{n}{2}, \beta=\frac{1}{2}$ 时,伽玛分布为 χ^2 分布。

(12) χ^2 分布 若随机变量 X 的密度函数

$$p(X) = \begin{cases} \frac{\chi^{\frac{n-2}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{X}{2}}, & X > 0; \\ 0, & X \leq 0, \end{cases}$$

其中 n 为正整数,则称 X 服从自由度为 n 的 χ^2 分布或 X 服从 $\chi^2(n)$ 分布。密度函数见图五, $EX=n$, $DX=2n$, 特征函数 $f(t)=(1-2it)^{-\frac{n}{2}}$ 。

若 $X_1, X_2, \dots, X_n$ 相互独立, X_k

$(1 \leq k \leq n)$ 服从 $\chi^2(n_k)$ 分布, 则和 $X_1 + X_2 + \dots + X_n$ 服从 $\chi^2(\sum_{k=1}^n n_k)$ 分布。

(13) t 分布 又称学生分布。若随机变量 X 的密度函数

$$p(X) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \cdot \left(1 + \frac{X^2}{n}\right)^{-\frac{n+1}{2}},$$

其中 n 为正整数,则称 X 服从自由度为 n 的 t 分布,或 X 服从 t

(n)分布。密度函数曲线见图六。 $EX=0, DX=\frac{n}{n-2} (n>2)$ 。

图 5 χ^2 分布的密度函数曲线

图 6 t 分布的密度函数曲线

(14)F 分布 若随机变量的密度函数

$$p(X) = \begin{cases} \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} n_1^{\frac{n_1}{2}-1} n_2^{\frac{n_2}{2}-1} X^{\frac{n_1}{2}-1} (n_1 X + n_2)^{-\frac{n_1+n_2}{2}}, & X > 0; \\ 0, & X \leq 0, \end{cases}$$

其中 n_1, n_2 为正整数, 则称 X 服从第一自由度为 n_1 , 第二自由度为 n_2 的 F 分布或 X 服从 $F(n_1, n_2)$ 分布, 密度函数曲线见图七。 $EX =$

$$\frac{n_2}{n_2 - 2} (n_2 > 2), DX = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_1 - 2)^2(n_2 - 4)} \quad (n_2 > 4)。$$

【概率论中的收敛】 设 $\{X_n, n \geq 1\}$ 是概率空间 (Ω, F, P) 上的随机变量序列, 它们都是 Ω 上的函数, 即 $X_n = X_n(\omega), \omega \in \Omega, n = 1, 2, \dots$ 。要求这一函数列对每个 ω 都收敛一般很难实现, 因此建立了下面一些收敛的概念:

(1) 以概率 1 收敛 若 $P(\omega: \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)) = 1$, 简记 $P(\lim_{n \rightarrow \infty} X_n = X) = 1$, 则称 $\{X_n, n \geq 1\}$ 以概率 1 收敛于 X 。以概率 1 收敛又称为几乎必然 (简记 $a. s.$) 收敛, 因此记为 $X_n \xrightarrow{a. s.}$

图 7 F 分布的密度函数曲线

X 。若把 $\{X_n, n \geq 1\}$ 看成可测函数列, 则上述收敛相当于测度论中的几乎处处 (简记 $a. e.$) 收敛。

(2) 依概率收敛 若对任意正数 ϵ , 都有 $\lim_{n \rightarrow \infty} P(\omega: |X_n(\omega) - X(\omega)| \geq \epsilon) = 0$ (简记 $\lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) = 0$) 或 $\lim_{n \rightarrow \infty} P(\omega: |X_n(\omega) - X(\omega)| < \epsilon) = 1$ (简记 $\lim_{n \rightarrow \infty} P(|X_n - X| < \epsilon) = 1$), 则称 $\{X_n, n \geq 1\}$ 依概

率收敛于 X , 记为 $X_n \xrightarrow{P} X$ 。若将概率看作测度, 则依概率收敛就是测度论中的依测度收敛。

(3) r 阶平均收敛 对 $r \geq 1$, 若 $\lim_{n \rightarrow \infty} E |X_n - X|^r = 0$, 则称 $\{X_n, n \geq 1\}$ r 阶平均收敛于 X , 记为 $X_n \xrightarrow{r} X$ 。特别, 当 $r=1$ 时, 称为平均收敛; 当 $r=2$ 时, 称为均方收敛。

(4) 依分布收敛 设随机变量序列 $\{X_n, n \geq 1\}$ 相应的分布函数列为 $\{F_n(X), n \geq 1\}$ 。设随机变量 X 的分布函数为 $F(X)$, 若对 $F(X)$ 的每一个连续点都有 $\lim_{n \rightarrow \infty} F_n(X) = F(X)$, 则称 $\{X_n, n \geq 1\}$ 的分布函数列 $\{F_n(X), n \geq 1\}$ 弱收敛于 X 的分布函数 $F(X)$, 记为 $F_n(X) \xrightarrow{w} F(X)$, 又称 $\{X_n, n \geq 1\}$ 依分布收敛于 X , 记为 $X_n \xrightarrow{c} X$ 。

$\{F_n(X), n \geq 1\}$ 弱收敛于 $F(X)$ 还有其它等价条件。例如, 对任一有界连续函数 $f(X)$, 都有

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(X) dF_n(X) = \int_{-\infty}^{\infty} f(X) dF(X)$$

成立就是一等价条件, 这里的积分是黎曼-斯蒂尔杰斯积分。

上述各种收敛之间有如下蕴含关系, 其中 $r' \geq r \geq 1, A \Rightarrow B$ 表示由 A 可得到 B :

$$\begin{array}{ccc} r' \text{ 阶平均收敛} & & \text{以概率 1 收敛} \\ \Downarrow & & \Downarrow \\ r \text{ 阶平均收敛} & \Rightarrow & \text{依概率收敛} \\ \Downarrow & & \Downarrow \\ \text{平均收敛} & & \text{依分布收敛} \end{array}$$

【大数定律】关于随机变量序列 $\{X_n, n \geq 1\}$ 的算术平均值 $\frac{1}{n} \sum_{i=1}^n X_i$ 向

期望平均值 $\frac{1}{n} \sum_{i=1}^n EX_i$ 收敛的定律。由于收敛有不同的方式,按其收敛为依概率收敛和以概率 1 收敛,分别称为弱大数定律和强大数定律,弱大数定律又常称为大数定律。这些定律是概率论和数理统计学中的基本定律之一。

若随机变量 X 的数学期望、方差分别为 EX 和 DX ,则对任何正数 a ,都有 $P(|X-EX| \geq a) \leq \frac{DX}{a^2}$ 。在 1867 年,这不等式由切比雪夫明确叙述和论证,故称之为切比雪夫不等式。它是证明弱大数定律的重要工具。

弱大数定律 关于随机变量序列 $\{X_n, n \geq 1\}$,对任意正数 ϵ , $\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n EX_i\right| < \epsilon\right) = 1$,即 $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \frac{1}{n} \sum_{i=1}^n EX_i$ 成立的定理。

(1) **伯努利大数定律** 在 n 次独立重复试验中,若 μ_n 表示事件 A 出现的次数, p 为事件 A 的概率,则对任意正数 $\epsilon > 0$, $\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \epsilon\right) = 1$,即 $\frac{\mu_n}{n} \xrightarrow{P} p$ 成立。由于 $\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \epsilon\right) = 1$ 等价于 $\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - p\right| \geq \epsilon\right) = 0$,因此,上述结论表明:只要 n 充分大,事件 A 的频率 $\frac{\mu_n}{n}$ 与概率 p 之间出现偏离大于 ϵ 的可能性可以任意小。

在 n 次独立重复试验中,考虑随机变量序列 $\{X_n, n \geq 1\}$,其中 X_i 表示第 i 次试验中 A 出现的次数,它是只取 0,1 两个值的随机变量且 $EX_i = p$,对一切 i 。因此, $\frac{\mu_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$,

$p = \frac{1}{n} \sum_{i=1}^n EX_i$,这样上述结论可表为

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n EX_i \right| < \epsilon \right\} = 1,$$

因此称为大数定律。

(2) 泊松大数定律 在一个独立试验序列中, 若 μ_n 表示事件 A 出现的次数, p_k 为在第 k 次试验中 A 出现的概率, 则 $\frac{\mu_n}{n} \xrightarrow{P} \frac{1}{n}$

$\sum_{k=1}^n p_k$ 。它之所以称为大数定律, 如(1)类似的解释。

(3) 切比雪夫大数定律 如果随机变量序列 $\{X_n, n \geq 1\}$ 两两独立, 其方差一致有界, 即存在常数 C, 对一切 k 有 $DX_k \leq C$, 则

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{P} \frac{1}{n} \sum_{k=1}^n EX_k。$$

(4) 辛钦大数定律 若 $\{X_n, n \geq 1\}$ 为独立同分布的随机变量序列, $EX_n = \mu$ 存在, 则

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu,$$

其中 $\mu = \frac{1}{n} \sum_{i=1}^n EX_i$ 。此定律为估计随机变量的数学期望提供了一条实际可行的途径。

强大数定律 关于随机变量序列 $\{X_n, n \geq 1\}$,

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} \frac{1}{n} \sum_{i=1}^n EX_i \text{ 成立的定理。最重要的是柯尔莫哥洛夫}$$

强大数定律: 若 $\{X_n, n \geq 1\}$ 为相互独立的随机变量序列, $EX_n = \mu$ 存在, 则

$$P \left(\lim_{m \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mu \right) = 1,$$

即

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} \mu$$

其中 $\mu = \frac{1}{n} \sum_{i=1}^n EX_i$ 。

【中心极限定理】关于随机变量序列部分和的分布渐近于正态分布的一类定理。1920年,波伊亚称它们为中心极限定理。这类定理有着广泛的实际背景。在自然界与科学试验中,如果某个量受到许多相互独立的随机因素的影响,而其中每个因素的影响又都很微小时,那么这个量受到叠加起来的总影响可近似地看作服从正态分布。例如测量误差。中心极限定理对此现象从数学上给予了严格的证明。最简单而又最重要的中心极限定理有:

(1)棣莫弗-拉普拉斯中心极限定理 在 n 重伯努利试验中,若事件 A 出现的次数为 μ_n ,事件 A 的概率为 p ($0 < p < 1$), $q = 1 - p$,则对任意实数 X ,有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n - np}{\sqrt{npq}} \right| < X \right\} \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^X e^{-\frac{t^2}{2}} dt。$$

如果用 X_i 表示在第 i 次试验中 A 出现的次数,它是只取 0,1 两个值的随机变量,从而得到一个独立的同为两点分布的随机变量序列 $\{X_n, n \geq 1\}$,则 $\mu_n = \sum_{i=1}^n X_i$ 。由这个定理可得到序列 $\{X_n, n \geq 1\}$ 的部分和 μ_n 近似地服从 $N(np, npq)$ 分布,即参数为 np 和 npq 的正态分布。

在 1716 年前后,棣莫弗对定理中 $p = \frac{1}{2}$ 的情形给予了证明。随后,拉普拉斯推广为上述的定理。

(2)林德伯格-莱维中心极限定理 若 $\{X_n, n \geq 1\}$ 是独立同分

布的随机变量序列且 $EX_k = \mu$ 和 $DX_k = \sigma^2$ 存在, 则部分和 $S_n = \sum_{i=1}^n X_i$ 的标准化 $S_n^* = \frac{S_n - n\mu}{\sigma \sqrt{n}}$ 的分布渐近于标准正态分布, 即对任意实数 X , 有

$$\lim_{n \rightarrow \infty} P(S_n^* < X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^X e^{-\frac{t^2}{2}} dt.$$

由此可知, 部分和 S_n 近似地服从 $N(n\mu, n\sigma^2)$ 分布。

这个定理在 20 世纪 20 年代由林德伯格和莱维证明。它在数理统计大样本理论中起重要的作用。

1935 年, 林德伯格和费勒提出了林德伯格条件, 从而使独立的(可不同分布)随机变量序列的中心极限定理得到根本解决。

数 理 统 计

【数理统计】研究如何进行有效地收集、整理和分析带有随机性的数据, 以对所研究的问题作出描述、推断或预测, 为采取一定的决策和行动提供依据和建议的数学分支学科。可用于物理、化学、医学、生物、地学、经济、社会等各种领域, 只涉及其中有关带随机性的数据处理问题。

数理统计是随着概率论的发展而发展起来的。当人们认识到必须把数据看成是来自具有一定概率分布的总体, 所研究的对象是这个总体而不能局限于数据本身之日, 也就是数理统计学诞生之时。在 19 世纪前, 已出现了高斯和勒让德关于观测数据误差分

析和最小二乘法的研究。高斯等的工作揭示了正态分布的重要性。德国的大地测量学者赫尔梅在 1875 年发现了统计上十分重要的 χ^2 分布。K. 皮尔森在 1900 年提出了检验拟合优度的 χ^2 统计量, 他的学生戈塞特于 1908 年导出了 t 分布, 即正态总体下 t 统计量的精确分布。费希尔在 20 年代系统地发展了正态总体下种种统计量的抽样分布; 建立了以最大似然估计为中心的点估计理论; 与耶茨合作创立了试验设计及方差分析; 在假设检验的发展中也起过重要作用。但假设检验理论的系统化和深入的研究, 应归功于奈曼和 E. S. 皮尔森(他是 K. 皮尔森之子)。奈曼对数理统计做出了另一项很重要的贡献, 他在 1934—1937 年间建立的置信区间理论。

多元统计分析是数理统计学中有重要应用价值的分支。1928 年以前, 费希尔已在狭义的多元分析(多元正态总体的统计分析) 方面做过一些工作。1928 年维夏特导出了著名的维夏特分布, 标志着多元分析成为一个独立的学科。以后, 狭义多元分析发展很快, 中国著名的数理统计学家许宝騄是做出重要贡献的学者之一。

在第二次世界大战后, 数理统计学在理论和应用两方面继续取得很大的进展。电子计算机的广泛应用, 也对数理统计学的发展起着很大的作用。有了计算机, 过去一些停留在理论上的方法得以实现, 而这又反过来促使人们提出和解决一些理论上的问题。通过计算机模拟, 可以在实际应用中避开一些难于解决的复杂的抽样分布问题。另外, 计算机在短时间内处理大量数据的能力, 使人们有可能从各个角度对数据进行透彻的分析, 从中提取更多的信息, 有的学者把这方面的工作称为“数据分析”, 并认为是数理统计中的一个生长点。总之, 计算机的广泛应用为数理统计学提供了巨大的机会, 也提出了一些很有价值的课题。

用数理统计方法去解决一个实际问题时, 一般有如下步骤: 建

立数学模型,收集整理数据,进行统计推断,预测或决策。其环节不一定按此顺序,有时是互相交错、不能截然分开的。在数理统计学中,模型是指所研究问题的总体的某种假定,一般是给总体分布规定一定的类型。有的假定不直接涉及总体分布形式。收集数据一般有全面观测、抽样观测和安排特定的试验。整理数据的目的是把包含在数据中的若干有用信息提取出来。一种形式是制定适当的图表,以反映隐含在数据中的一般趋势及粗略的规律。另一种形式是计算若干特征数字,如样本均值和方差分别反映了样本内各数值的平均水平及差异程度。这种特征数字在数理统计中称为统计量,其研究有重要的理论和实际意义,所得到的结果可用于指导数据的整理和统计推断方法的选择。统计推断是根据总体模型以及由总体中抽出的样本,做出有关总体分布的某种论断。统计预测的对象,是随机变量在未来某个时刻所取的值,或设想在某种条件下对该变量进行观测时将取的值。统计推断与预测相似之处在于:一是都要依据一定的统计模型和观测数据;二是都要超出已观察的事物的范围。不同之处在于:统计推断的对象是总体分布的一个方面,统计预测的对象不仅未知,且是随机的。但预测和推断又不能截然分开,在许多情况下,为了进行预测,必须先做统计推断。统计决策是依据所做的统计推断或预测,并考虑到行动的后果而制定的一种行动方案,目的是使损失尽可能小。

【统计量】样本的已知函数。其作用是把样本中有关总体的信息概括起来,是数理统计学中一个重要的基本概念。

一个统计问题所研究的对象的全体称为总体或母体,总体中的每一个基本单位称为个体。一般我们所关心的不是每一个个体的全部特征,而是它的某一个或几个数量特征。例如,小麦的植株

高度、粒重;鸡的产蛋量、每天增重量等。这些数量特征一般都是以随机变量的形式出现。因此对于一个总体的某个数量特征来说,可以用一个随机变量 X 来表示。

按一定方式从总体中抽取的若干个体称为一个样本,它用于提供总体的信息及由此对总体做统计推断。用样本对总体作推断时,常利用样本的某个或某些特定的函数,例如样本均值、样本方差、样本相关系数等。用样本推断总体,常通过统计量来实现。若样本由 n 个个体 $X_1, X_2, \dots, X_n$ 所组成,则可把样本记为 $X = (X_1, X_2, \dots, X_n)$ 或简记为 X 。 n 称为样本容量。样本 X 所能取的一切值的集合称为样本空间。

从总体中抽取样本的过程叫抽样。最常用的抽样方法是简单随机抽样,按这种方式抽样,总体中的每个个体都有同等的机会被抽入样本,且每次抽出的个体,在记下其指标后,还要放回总体中去,以保证在下次抽取时每个个体仍有与第一次抽取时同等的机会被抽入样本,这样得到的样本称为简单随机样本,简称随机样本或简单样本。简单随机抽样的具体结果 $X_1, X_2, \dots, X_n$ 是一组确定的值,但它是随每次抽样而改变的,因此每个 X_i 可看作是描述总体的随机变量 X 的一个观测值,一个随机样本又可记为 $X = (X_1, X_2, \dots, X_n)$,其中 X_i 是相互独立且均与总体同分布的随机变量,故 X 是一个 n 维的随机向量,它的分布称为样本分布。

统计量只依赖于样本 $X_1, X_2, \dots, X_n$,它不含总体分布的任何参数。从样本推断总体通常是通过统计量进行的。常用的统计量有样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 和样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$,前者反映总体中心位置的信息,后者反映总体离散情况。其它常用的统计量如样本标准差 $S = \sqrt{S^2}$,样本变异系数等。若 $(X_i, y_i) (i=1, 2, \dots,$

n)是从二维总体 (X, Y) 中抽取的随机样本,则样本协方差 $S_{XY} =$

$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X} - \bar{Y})$ 和样本相关系数

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2 \right]^{1/2}}$$

也是常用的统计量, r 可用于推断 X 和 Y 的相关性。

统计量的分布叫抽样分布。它与样本分布不同,样本分布是指样本 $X_1, X_2, \dots, X_n$ 的联合分布。统计量的性质以及使用某一统计量作推断的优良性,取决于其分布,所以抽样分布的研究是数理统计中的重要课题。对一维正态总体,有三个重要的抽样分布,即 t 分布、 F 分布和 χ^2 分布。

一个统计量若服从某分布,常以该分布的名字命名该统计量,如 t 统计量、 F 统计量、 χ^2 统计量等。

多元正态总体的重要抽样分布维夏特有分布和霍特林的 T^2 分布等。

【参数估计】统计推断的基本问题之一。从总体抽得的样本估计总体分布中包含的未知参数。有两种类型:要求构造一个统计量作为参数的一个估计量的点估计;在一定的置信度之下用区间给出参数的一个取值范围的区间估计。

【点估计】参数估计的一种方法。其目的是根据样本 $X = (X_1, X_2, \dots, X_n)$ 估计总体分布所含的未知参数 θ 或 θ 的函数 $f(\theta)$ 。一般 θ 或 $f(\theta)$ 是总体分布的某个特征值,如数学期望、方差、相关系数等。

点估计就是要构造一个只依赖于样本 X 的量 $\hat{f}(X)$, 作为 $f(\theta)$ 的估计值。 $\hat{f}(X)$ 称为 $f(\theta)$ 的估计量。

构造点估计的一种重要而普遍的方法是最大似然估计法, 由英国统计学家费希尔于 1912 年提出。设样本 $X=(X_1, X_2, \dots, X_n)$ 的分布密度为 $L(X, \theta)$, 若固定 X 而将 L 看作 θ 的函数, 称为似然函数, 当 X 是简单随机样本时, 密度为 $p(X_1, \theta)p(X_2, \theta) \cdots p(X_n, \theta)$, 其中 $p(X, \theta)$ 是总体分布的密度函数。若得到样本值 X , 就可确定 $\hat{\theta}(X)$, 使 $L(X, \hat{\theta}(X)) = \sup_{\theta} L(X, \theta)$, 然后用 $\hat{f}(X) = f(\hat{\theta}(X))$ 估计 $f(\theta)$, 此即为 $f(\theta)$ 的最大似然估计。不难证明, 估计正态分布 $N(\mu, \sigma^2)$ 中的参数 μ 和 σ^2 , 其估计量 $\bar{X}$ 和 $\hat{\sigma}^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / n$ 就是 μ 和 σ^2 的最大似然估计。

【区间估计】参数估计的一种方法。其目的是根据从总体中抽取的样本及一定的精确度和正确度的要求, 构造出适当的区间, 以作为总体的分布参数的真值所在范围的估计。例如, 估计某种皮棉的单行平均产量在 $4.0-4.3kg$ 之间; 估计某种雄鼠在 X 射线照射前后体重减少在 $0.6-1.7kg$ 之间等。

在数理统计学中, 待估计的未知量是总体分布的参数 θ 或 θ 的函数 $f(\theta)$ 。区间估计可一般地表示为: 要求构造一个仅依赖于样本 $X=(X_1, X_2, \dots, X_n)$ 的适当的区间 $[A(X), B(X)]$, 若得到了样本 X 的观测值 X , 就把区间 $[A(X), B(X)]$ 作为 θ 或 $f(\theta)$ 的估计。由于样本 X 是随机抽取的, 区间 $[A(X), B(X)]$ 是否真包含待估计的 θ , 取决于所抽的样本 X 。因此, 区间 $[A(X), B(X)]$ 只能以一定的概率 $\pi(\theta) = P_{\theta}(A(X) \leq \theta \leq B(X))$ 包括未知的 θ 。对于不同的 θ , $\pi(\theta)$ 之值可以不同, $\pi(\theta)$ 对不同的 θ 取的最小值 $1-\alpha (0 < \alpha < 1)$

称为区间 $[A(X), B(X)]$ 的置信系数, $[A(X), B(X)]$ 称为 θ 的一个置信区间。求置信区间的一种常用方法是利用已知的抽样分布。

【假设检验】又称统计假设检验。一种基本的统计推断,是数理统计学的一个重要分支。“假设”是指关于总体分布的一个命题。假设是否正确,要对从总体中抽取的样本进行检验,与此有关的理论和方法构成假设检验的内容。

设 A 是关于总体分布的一个命题,所有使命题 A 成立的总体分布构成一个集合 H_0 ,称为原假设(或称假设、零假设)。使命题不成立的所有总体分布构成另一个集合 H_1 ,称为备择假设。若 H_0 可用通过有限个实参数来描述,则称为参数假设,否则称为非参数假设。如果 H_0 (或 H_1)只包含一个分布,则称原假设(或备择假设)为简单假设,否则为复合假设。对一个假设 H_0 进行检验,即要制定一个规则,使得有了样本后,根据这规则可以决定是接受它(承认命题 A 正确),还是拒绝它(否认命题 A 正确)。拒绝 H_0 的集合称为检验的拒绝域。选定一个检验法,就是选定一个拒绝域,故常把检验法本身与拒绝域等同起来。

有时根据一定的理论或经验,认为某一假设 H_0 成立。如通常认为指定的一群人的身高服从正态分布。当得到样本后,可以考虑实际数据与理论假设 H_0 间的偏离,如果偏离达到了“显著”的程度就拒绝 H_0 ,这种检验法称为显著性检验。如何规定什么时候偏离达到显著程度,通常是指定一个很小的正数 α ($= 0.05$, 或 0.01),使当 H_0 正确时,它被拒绝的概率不超过 α , α 称为显著性水平。

在检验一个假设 H_0 时可能犯两类错误:第一类错误是真实情况为 H_0 成立,但判定 H_0 不成立,犯了“以真为假”的错误;第二

类错误是真实情况为 H_0 不成立,但判定 H_0 成立,犯了“以假为真”的错误。

常用的重要检验有:

(1) U 检验。若总体服从正态分布 $N(\mu, \sigma^2)$, 其中 σ 已知, $X = (X_1, X_2, \dots, X_n)$ 是从总体中抽取的随机样本, 记 $\bar{X} = \sum_{i=1}^n X_i/n$, 则 $U = \sqrt{n}(\bar{X} - \mu)/\sigma$ 服从标准正态分布 $N(0, 1)$, 故可考虑对 μ 的以下几种假设检验。其中 μ_0 是给定的常数, α 为检验的显著性水平, u_α 为标准正态分布的上 α 分位数。这些检验称为 U 检验。

原假设 H_0	拒绝域
$\mu \leq \mu_0$	$\bar{X} \geq \mu_0 + u_\alpha \sigma / \sqrt{n}$
$\mu \geq \mu_0$	$\bar{X} \leq \mu_0 - u_\alpha \sigma / \sqrt{n}$
$\mu = \mu_0$	$ \bar{X} - \mu_0 \geq u_{\alpha/2} \sigma / \sqrt{n}$

(2) t 检验。若总体服从正态分布 $N(\mu, \sigma^2)$, 但 σ 未知, 记 $\bar{X} = \sum_{i=1}^n X_i/n$, $S^2 = \left(\sum_{i=1}^n (X_i - \bar{X})^2 \right) / (n-1)$, 则 $t = \sqrt{n}(\bar{X} - \mu)/S$ 服从自由度为 $n-1$ 的 t 分布, 对 μ 有以下几种水平为 α 的假设检验。其中 t_α 为自由度为 $n-1$ 的 t 分布的上 α 分位数。这些检验称为 t 检验。

原假设 H_0	拒绝域
$\mu \leq \mu_0$	$\bar{X} \geq \mu_0 + t_\alpha S / \sqrt{n}$
$\mu \geq \mu_0$	$\bar{X} \leq \mu_0 - t_\alpha S / \sqrt{n}$
$\mu = \mu_0$	$ \bar{X} - \mu_0 \geq t_{\alpha/2} S / \sqrt{n}$

(3) F 检验。若 $X=(X_1, X_2, \dots, X_{n_1})$ 及 $Y=(Y_1, Y_2, \dots, Y_{n_2})$ 分别来自正态总体 $N(\mu_1, \sigma_1^2)$ 及 $N(\mu_2, \sigma_2^2)$ 的随机样本, 记 $\bar{X} = \sum_{i=1}^{n_1} X_i / n_1$, $\bar{Y} = \sum_{i=1}^{n_2} Y_i / n_2$, $S_1^2 = \left(\sum_{i=1}^{n_1} (X_i - \bar{X})^2 \right) / (n_1 - 1)$, $S_2^2 = \left(\sum_{i=1}^{n_2} (Y_i - \bar{Y})^2 \right) / (n_2 - 1)$, 则 $F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$ 服从自由度为 $(n_1 - 1, n_2 - 1)$ 的 F 分布。比较 σ_1^2 与 σ_2^2 的假设有以下的水平为 α 的假设检验。其中 F_α 为自由度为 $(n_1 - 1, n_2 - 1)$ 的 F 分布的上 α 分位数。这些检验称为 F 检验, 它在方差分析中有大量的应用。

原假设 H_0	拒绝域
$\sigma_1^2 \leq \sigma_2^2$	$S_1^2 / S_2^2 \geq F_\alpha$
$\sigma_1^2 \geq \sigma_2^2$	$S_2^2 / S_1^2 \leq F_{1-\alpha}$
$\sigma_1^2 = \sigma_2^2$	$S_1^2 / S_2^2 \geq F_{\alpha/2}$ 或 $S_2^2 / S_1^2 \leq F_{1-\alpha/2}$

(4) χ^2 检验。是考虑试验数据与理论之间拟合程度的一个重要的拟合优度检验。设原假设 H_0 是“总体分布等于某个已知的分布函数 $F(X)$ ”。把 $(-\infty, +\infty)$ 分为若干个两两无公共点的区间 $I_1, I_2, \dots, I_k$ 对任一区间 $I_j = (a_j, a_{j+1})$, 以 v_j 记大小为 n 的样本 $X_1, X_2, \dots, X_n$ 中落在 I_j 内的个数, 称为区间 I_j 的观测频数。对 $I_j (j=1, 2, \dots, k)$ 求出其理论频数 $\mu_j = n[F(a_{j+1}) - F(a_j)]$, 算出由 $\chi^2 = \sum_{j=1}^k (v_j - \mu_j)^2 / \mu_j$ 定义的 χ^2 的统计量, 则若 $F(a_{j+1}) - F(a_j) > 0$ 对 $j=1, 2, \dots, k$, 则当 $n \rightarrow \infty$ 时, χ^2 的极限分布是自由度为 $k-1$ 的 χ^2 分布。于是在样本 n 充分大时, 从 χ^2 分布表可查得 χ^2 分布的上 α 分位数 $\chi_\alpha^2(k-1)$ 。由此即得显著性检验水平为 α 的拒绝域 $\{\chi^2 > \chi_\alpha^2(k-1)\}$ 。如果原假设 H_0 为总体服从一族分布, 也可得到类似结果,

这只要在计算理论频数 v_j 时,将所包含的未知参数用适当的点估计代替,即可计算 χ^2 统计量。但此时极限分布的自由度为 $k-1-1$,其中 1 为独立参数的个数。

(5)列联表中的独立性检验。观测数据按两个或多个属性(定性变量)分类时所列出的频数表称为列联表。例如,有人统计了小家鼠到两种诱捕器(以同样的食物为诱饵)去的频数。一号诱捕器沾了以往小家鼠的尿,二号是清洁的,结果得到两行两列的列联表,或称 2×2 表或四格列联表。

结 果	一号	二号	行和 $n_{i.}$
进 入 诱捕器	170	40	210
不进入 诱捕器	90	240	330
列和 $n_{.j}$	260	280	540

一般,若总体中的个体可按两个属性 A 与 B 分类,A 有 r 个等级 $A_1, A_2, \dots, A_r$, B 有 c 个等级 $B_1, B_2, \dots, B_c$,从总体中抽取大小为 n 的样本,设其中有 n_{ij} 个属于等级 A_i 和 B_j , n_{ij} 称为频数,将 $r \times c$ 个 n_{ij} ($i=1, 2, \dots, r; j=1, 2, \dots, c$) 排列为一个 r 行 c 列的二维列联表,简称 $r \times c$ 列联表。若所考虑的属性多于两个,也可按类似的方式作出列联表,称为多维列联表。列联表分析在应用统计,特别是在医学、生物学、社会科学中有重要应用。

A	B				行和
	B_1	B_2	$\dots$	B_c	
A_1	n_{11}	n_{12}	$\dots$	n_{1c}	$n_{1.}$
A_2	n_{21}	n_{22}	$\dots$	n_{2c}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
A_r	n_{r1}	n_{r2}	$\dots$	n_{rc}	$n_{r.}$
列和	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.c}$	n

(6)非参数统计检验。在一个统计问题中,其总体分布不能用有限个实参数来刻画,只能对它做一些分布连续、有密度等一般性的假定,则称为非参数统计问题。秩检验是一类重要的非参数统计方法。设有随机样本 $X_1, X_2, \dots, X_n$, 把它们从小到大排列,若 X_i 在这个次序中占第 R_i 个位置(最小的占第一个位置),则称 X_i 的秩为 $R_i (i=1, 2, \dots, n)$ 。对两样本的秩检验,设 $X_1, X_2, \dots, X_m$ 和 $Y_1, Y_2, \dots, Y_n$ 分别是来自分布为 $F(X)$ 和 $F(X-\theta)$ 的总体中抽出的样本, F 是连续但未知, θ 也未知,检验假设 $H_0: \theta=0$ 。记 Y_i 在混合样本 $(X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n)$ 中的秩为 R_i , 且 $W = \sum_{i=1}^n R_i$ 为诸秩的和。当 $W > C$ 时,否定假设 H_0 , 这里 C 决定于检验的水平。

非参数统计问题中对总体分布的假定要求的条件很宽,因而针对这类问题而构造的非参数方法,不致因为对总体分布的假定不当而导致重大错误,它往往有较好的稳健性。但因为该方法需要照顾很广的分布,在某些情况下会导致其效率的降低。

(7)相关系数检验。若 $(X_i, Y_i) (i=1, 2, \dots, n)$ 是从二维总体 (X, Y) 中抽取的随机样本,记样本相关系数

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2 \right]^{1/2}},$$

其中 $\bar{X} = \sum_{i=1}^n X_i / n, \bar{Y} = \sum_{i=1}^n Y_i / n$ 。检验 X 与 Y 是否相关的假设 $H_0: \rho_{XY} = 0$, 用 r 估计 ρ_{XY} , 查 r 值表的上 α 分位数 $r_\alpha(n-2)$, 由此即得显著性水平为 α 的拒绝域 $\{r > r_\alpha(n-2)\}$ 。

【方差分析】在生产和试验中,常常需要分析影响产品质量和产量

的诸多因素中哪些起重要作用,并找出它们的最优状态或数值,以便获得最佳生产或试验条件。例如,欲比较四种水稻品种。比较的指标是单位产量,假定每种品种的水稻都种了几公顷,每公顷的产量分别为 $y_{i1}, \dots, y_{in}, i=1, 2, 3, 4$, 现在要通过分析这些数据来判断这四种水稻品种有无差异,若有,哪一种好?在数理统计学中,研究如有效地分析这类数据的统计理论和方法,称为方差分析。

记

$$\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}, \quad \bar{y} = \frac{1}{4n} \sum_{i=1}^4 \sum_{j=1}^n y_{ij}$$

它们分别为第 i 种水稻的平均单产量和四种水稻的平均单产量。而

$$SS_T = \sum_{i=1}^4 \sum_{j=1}^n (y_{ij} - \bar{y})^2$$

是水稻产量的总变差平方和,它度量了全部数据 $y_{i1}, \dots, y_{in}, i=1, 2, 3, 4$ 的散布程度。这散布来自两个原因:一个是四种水稻品种本身的差异;二是各种各样的随机误差的影响,如果四种水稻品种的差异愈大,由它们所引起的散布度就愈大。反过来,若四种水稻品种无差异,相应的散布度就等于零,这时 SS_T 全部由随机误差而来,因此,我们可以把总变差平方和 SS_T 按其来源分解成两部分

$$SS_T = SS_{\text{品种}} + SS_{\text{误差}},$$

其中

$$\begin{aligned} SS_{\text{品种}} &= \sum_{i=1}^4 \sum_{j=1}^n (\bar{y}_i - \bar{y})^2 \\ &= n \sum_{i=1}^4 (\bar{y}_i - \bar{y})^2, \\ SS_{\text{误差}} &= \sum_{i=1}^4 \sum_{j=1}^n (\bar{y}_{ij} - \bar{y}_i)^2, \end{aligned}$$

它们分别是由品种差异和随机误差所引起的变差平方和。比值 $SS_{\text{品种}}/SS_{\text{误差}}$ 愈大,我们就愈有理由判断四种水稻品种有差异。否则就认为这四种水稻无差异。至于这个比值多大,就认为有显著差异,这就需要进行显著性检验。在判定了它们之间有差异之后,我们还关心哪种品种好,好到什么程度,这就需要用多重比较方法,这些就构成了方差分析的内容。

在上面的讨论中,我们所关心的因素只有水稻品种一个,所以分析这种数据的统计方法叫做单因素方差分析,如果我们要研究的除了四种水稻品种之外,还有三种化学肥料,目的是找出肥料与稻种的最佳搭配,那么就包含了稻种和肥料两个因素,相应的分析方法叫做两因素方差分析。很自然,如果我们处理多个因素的问题就需要多因素方差分析,基本原理是一样的。在多因素方差分析中,我们把总变差平方和 SS_T 按其来源分解成

$$SS_T = SS_1 + \cdots + SS_k + SS_{\text{误差}},$$

这里 $SS_1, \cdots, SS_k$ 分别表示由因素 1, $\cdots$, 因素 k 所引起的变差平方和,我们根据比值 $SS_i/SS_{\text{误差}}$ 的大小做显著性检验,来判断第 i 个因素是否重要。然而,对一组具体数据能否做上述分解,如何分解,这取决于试验的设计。因此方差分析跟数理统计的另一个分支——试验设计有着密切的关系。

【试验设计】又称实验设计。数理统计学的一个分支,研究如何制定试验方案,以提高试验效率,缩小随机误差的影响,并使试验结果能有效地进行统计分析的理论与方法。其基本思想是减少偶然性因素的影响,使试验数据有一个合适的数学模型,以便使用方差分析方法对数据进行分析。试验设计应遵循三个原则:随机化、局部控制和重复。随机化的目的是使试验结果尽量避免受到主客观系

统性因素的影响而呈现偏倚性;局部控制是用划分区组的方法,使区组内部条件尽可能一致;重复是为了降低随机误差的影响,以保证结果的重现性。

析因设计是将全部因素的水平组合起来做试验,使得既能估计各个因素的主效应,又能估计因素之间的交互作用的一种试验设计法。主效应,指同一因素各水平之间的差异;交互作用指一个因素的效应因另一个因素的水平变化而起的变化。主效应和交互作用统称效应。析因设计在农业试验、工业试验、心理试验及其它科学技术试验中得到了大量的应用。

正交试验是一种试验次数少而代表性大、提供的信息丰富、反映情况全面的一种试验设计法。这里给出一个正交表 $L_9(3)^4$,表示有 9 行 4 列,每行都包含 3 个水平,它可以用来安排 3 水平 4 因素的试验。按正交表安排并进行分析的试验称为正交试验。正交表有性质:一是任一系列的每个水平出现的次数相同;二是任意两列的各种不同水平组合出现的次数相同。在应用中,当把因素对应于正交表各列时,各行则表示应做试验的水平组合。由上述二性质,任一因素的效应可不受其他因素干扰。

【回归分析】在现实世界中存在着两种变量依赖关系。一种是由自变量的值能够完全准确地确定因变量的值,这种关系称为确定型关系。例如,圆的面积 S 依赖于其半径 r ,由半径 r 可通过函数 $S = \pi r^2$ 完全确定面积 S 的值。另一种依赖关系情况则不同,虽然因变量依赖于自变量的值,但它们的关系尚没有密切到由自变量可以完全确定因变量的程度。例如,众所周知,人的身高与其体重有一定关系,表现在高个人的体重一般比较重。但知道身高并不能严格标出其体重。这种变量具有某种程度的不确定性,称为非确定型关

系或相关关系。这种关系大量地存在于客观世界的各类问题之中,在数理统计学中,研究如何有效地通过大量观测或试验数据建立描述这种相关关系的理论和方法,通称为回归分析,它是数理统计学中应用最广泛的分支之一。

用 Y 和 X 分别记我们所研究的因变量和自变量。如果它们之间存在一定的相关关系,那么在很多情况下,我们可以用下面的简单数学形式来近似地刻画

$$Y = \alpha + \beta X + e,$$

称之为理论回归方程。其中 e 是随机误差,它代表了 Y 依赖于 X 的不确定部分,而主要的确定部分是由 $\alpha + \beta X$ 来描述,这里 α 和 β 是未知参数,需要通过数据来估计。最常用的估计方法叫做最小二乘法。假定我们对自变量 X 和因变量 Y 有一些观测或试验结果: $(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)$, 它们自然满足关系式

$$y_i = \alpha + \beta X_i + e_i, \quad i = 1, 2, \dots, n.$$

所谓最小二乘法就是选择 α 和 β , 使得误差平方和。

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \alpha - \beta X_i)^2$$

达到最小,应用微积分知识,对上式关于 α 和 β 求偏导数,并令其等于零,得到关于 α 和 β 的一组线性代数方程。解这个方程得到 α 和 β 的最小二乘估计分别为

$$\begin{aligned} X &= \bar{y} - \hat{\beta} \bar{X}, \\ \hat{\beta} &= \frac{\sum_{i=1}^n X_i y_i - n \bar{X} \bar{y}}{\sum_{i=1}^n (X_i - \bar{X})^2}, \end{aligned}$$

其中

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

这样我们就得到经验回归方程

$$Y = \hat{\alpha} + \hat{\beta}X.$$

之所以称为经验回归方程,是因为它是通过一组观测数据而得到的。至于这个经验回归方程是否能够真正刻画 Y 与 X 的关系,一方面我们可以从数理统计学的角度做一些假设检验,另一方面,也是更主要的,应该通过实践去检验。如果检验的回答是肯定的,那么我们就认为上述经验回归方程确实描述了两变量之间的相关关系。据此可分析 Y 对 X 的依赖强度,例如 X 每增加一个单位,平均说来,相应的 Y 增加 $\hat{\beta}$ 个单位。还可以通过 X 值来预报 Y 的值。例如,当 $X = X_0$ 时,对应地 Y 平均说来取值应为 $X + \hat{\beta}X_0$ 等等。

在实际问题中,可能自变量不止一个。假定有 p 个,分别记为 $X_1, X_2, \dots, X_p$, 那么在很多情况下,他们之间的相关关系可近似地描述为

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + e.$$

和单个自变量的情形一样, e 是误差部分,有了实际数据应用最小二乘法,我们可以求到未知参数 $\beta_0, \beta_1, \dots, \beta_p$ 的最小二乘估计 $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ 和经验回归方程

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_p X_p.$$

在前面的讨论中我们都假定因变量 Y 和诸自变量之间关系的主要部分是线性关系,回归分析中相应的理论和方法称为线性回归分析。很明显,在一些情况下,变量之间的关系需要用非线性函数来描述,这就导致了非线性回归分析。

【多元统计分析】简称多元分析。当总体的分布是多元概率分布时,

处理该总体的数理统计理论和方法。数理统计学中一个有重要应用价值的一个分支学科。

多元分析发展初期,主要讨论把一元正态总体的统计理论和方法推广到多元正态总体。多元正态总体的分布由两组参数,即数学期望 μ 和协方差矩阵 Σ 所决定,记为 $N_p(\mu, \Sigma)$ 。设 $X_1, X_2, \dots, X_n$ 为来自 $N_p(\mu, \Sigma)$ 的样本,则 μ 和 Σ 的无偏估计分别是

$$\hat{\mu} = \bar{X} = \sum_{i=1}^n X_i / n \text{ 和 } \hat{\Sigma} = S =$$

$\left[\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \right] / (n-1)$, 分别称为样本均值向量和样本协方差阵,它们是在各种多元分析中常用的统计量。另一个常用的统计量是样本相关矩阵 R , 其元素 $r_{ij} = v_{ij} / \sqrt{v_{ii}v_{jj}}$, 其中 v_{ij} 为样本协方差阵 S 的元素。 S 的分布是维夏特分布,是一元统计中 χ^2 分布的推广。另一问题是假定二个多元正态分布协方差阵相同,检验其均值向量是否相等。设样本 $X_1, X_2, \dots, X_n$ 抽自 $N_p(\mu_1, \Sigma)$, $Y_1, Y_2, \dots, Y_m$ 抽自 $N_p(\mu_2, \Sigma)$, 检验假设 $H_0: \mu_1 = \mu_2$ 。其统计量为

$$T^2 = \frac{nm}{n+m} (\bar{X} - \bar{Y})' \hat{\Sigma}^{-1} (\bar{X} - \bar{Y}), \text{ 其中 } \bar{X} = \sum_{i=1}^n X_i / n, \bar{Y} = \sum_{j=1}^m Y_j /$$

$$m, \hat{\Sigma} = \left[\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' + \sum_{j=1}^m (Y_j - \bar{Y})(Y_j - \bar{Y})' \right] / (n+m-2), T^2 \text{ 的分布称为 } T^2 \text{ 分布,它是一元统计中 } t \text{ 分布的推广。}$$

多元分析有狭义与广义之分。当假定总体分布是多元正态分布时称为狭义的,否则称为广义的。

多元分析的主要研究内容有:化简数据结构,考虑是否能用简单的方法来表示较复杂的样本或指标,如降低指标的维数、将相关

变量组变换为独立变量组等;对具有多元数据的样本或指标进行分类;分析一个变量或几个变量对其它变量的依赖关系、二组变量间的相互依赖关系等;各种统计假设的建立和检验;各种多元统计方法的计算和研究。

按多元分析所处理的实际问题的性质分类,重要的有以下几种:

(1)聚类分析。是研究物以类聚的一种方法。其作用是建立一种分类方法,将一批样本或变量按照它们在性质上的紧密程度进行分类。一般有两种方法:一是把每个样本看成是 m 维(变量个数为 m 个)空间的一个点,在 m 维坐标系中,定义点与点之间的某种距离;另一种是用某种相似系数来描述变量或样本点间的紧密程度。当确定了样本或变量间的距离或相似系数后,就要对其分类。方法很多,有系统聚类法、动态聚类法、分解法、模糊聚类法等。系统聚类法是应用最广泛的一种方法。

(2)多元回归分析。见“回归分析”。

(3)判别分析。是利用若干个变量的一个或几个线性或非线性组合,将待判样本的多变量观测值划分为几个已知类之一的一种方法。按判别的组数来分,有两组判别分析和多组判别分析;按区分不同类所用的数学模型来分,有线性判别和非线性判别;按判别对所处理的变量不同有逐步判别、序贯判别等;有不同的判别准则,如费希尔准则和贝叶斯准则等。应用判别分析必须以事先存在几类为前提,而应用聚类分析则不必事先确定类的个数。

(4)主成分分析。又称主分量分析,是把原来多个变量化为少数几个相互独立的综合变量的一种常用的方法。设 $X = (X_1, X_2, \dots, X_p)'$ 是 p 维向量,它的第 i 个主成分 $Y_i = u_{i1}X_1 + u_{i2}X_2 + \dots + u_{ip}X_p$ ($i = 1, 2, \dots, p, u_{i1}^2 + u_{i2}^2 + \dots + u_{ip}^2 = 1$, 即 $U'U = 1$), 记 $Y_i = u'_i X$,

满足: Y_1 是一切形如 $Y=U'X$ 中方差达到最大者; Y_k 是一切形如 $Y=U'X$ 中与 $Y_1, Y_2, \dots, Y_{k-1} (k=2, 3, \dots, p)$ 都不相关且方差达到最大者。

(5) 因子分析。又称因素分析, 是研究相关矩阵的内部依赖结构, 用少数不可观测的公共因子的线性函数以及特定因子来描述总体, 以分析总体与若干公因子之间的关系。即在互为关联的许多因子中, 找出能反映它们内在联系和起主导作用的、数目较少的公因子。根据这些公因子进行研究, 既无损于多个因子所反映的信息, 又便于对原变量进行分类和解释。因子分析是主成分分析的推广, 也是降维方法之一。

(6) 典型相关分析。又称典范相关分析, 两组随机向量之间相关关系的分析方法, 而主成分分析和因子分析只涉及到一组变量的相互依赖关系。在两组变量中, 分别选取若干有代表性的变量组成有代表性的综合指标, 通过研究这两组综合指标之间的相关关系, 来代替这两组变量间的相关关系, 这些综合指标称为典型变量。

(7) 对应分析。又称相应分析。在因子分析中, 如果研究的对象是变量, 则采用 R 型因子分析; 如果研究的对象是样本, 则采用 Q 型因子分析。这两种分析通常是对立的概念, 分别进行处理。对应分析是对非负值变量的样本资料阵做适当的处理后, 将 R 型与 Q 型因子分析相结合的公共因子分析方法。其结果把变量和样本同时反映在一个因子平面上, 从而可以进行分类和解释, 且主要提供以下信息: 样本间的关系: 邻近的样本点具有相似性质, 属于同一类; 变量间的关系: 在以因子轴为坐标轴的图形上, 相邻近的一些变量点表示这些变量相关性密切; 变量与样本间的关系: 同一类型的样本点将为邻近的变量所表征。

随 机 过 程

【随机过程】依赖于某种参量的一族随机变量。最开始研究的是依赖于时间参量的一族随机变量。这时,随机过程描述的是随机变量或随机现象随着时间推移而演变的过程。例如,某国每年人口的增长数,它是一个随机变量,而且随着年限的变化而发生变化,因此,需用一族随机变量来描述。若用随机变量 X_n 表示第 n 年人口的增长数,于是可用依赖于时间的一族随机变量 $\{X_n, n \in T\}$, 其中 $T = \{1, 2, \dots\}$, 来描述人口增长随时间演变的过程。又如,某售票处在售票时间的每个时刻等待购票的人数,它也是随机变量,若用随机变量 $X(t)$ 表示在 t 时刻等待购票的人数,那么,依赖于时间 t 的一族随机变量 $\{X(t), t \in T\}$, 其中 T 表示售票时间的时间间隔,便描述了售票处等待购票人数随时间的演变过程。再如,一年内温度随时间的变化过程类似地可用一族随机变量 $\{X(t), t \in T\}$ 来描述,其中 $X(t)$ 表示 t 时刻的温度, T 表示一年的时间间隔。上述的每一族随机变量都是一个随机过程。

随机过程的起源是和物理中噪声和波动的研究密切联系着的,因为随机过程给这种研究提供了数学模型。现在,随机过程不仅在物理学中,而且在工程技术、医学、生物学、经济学等等许多领域中出现。随机过程也随着其它学科的发展而不断地丰富和完善。

设 (Ω, F, P) 为概率空间,如果对每个 $t \in T$, 有定义在 (Ω, F, P) 上的随机变量 $X(t) = X(t, \omega)$ ($X(t)$ 亦可记为 X_t) 与之对应,则

称随机变量族 $X = \{X(t), t \in T\}$ 为定义在 (Ω, F, P) 上的随机过程, 简称过程, 其中 T 称为参量集。通常 T 为实数集 $R = (-\infty, \infty)$ 的子集。特别, 当 T 为自然数集时, 称 $\{X_n, n \geq 1\}$ 为随机序列。 T 也可以是 d 维欧几里得空间 R^d (d 为大于 1 的正整数) 的子集, 这时称 $\{X(t), t \in T\}$ 为多指标随机过程。 T 还可以是更一般的集合, 如 T 可为 R 中的集合类或某种抽象空间, 等等。

过程 $X = \{X(t), t \in T\}$ 可看作两个变元 (t, ω) ($t \in T, \omega \in \Omega$) 的函数

$X(t, \omega)$ 。当 t 固定时, 它是一个随机变量, 即 Ω 上的函数, 记为 $X(t, \cdot)$ 或

$X(t)$; 当 ω 固定时, 它是 T 上的函数, 记为 $X(\cdot, \omega)$, 称此函数为随机过程关于 ω 的轨道或样本函数, 它表示随机过程的一个可能的观察结果。

如果在过程 $X = \{X(t), t \in T\}$ 中, 对每个 $t, X(t)$ 都是 d (大于 1 的正整数) 维随机向量, 则称 X 为 d 维向量过程。

过程中的随机变量 $X(t)$ 取值的空间称为状态空间或相空间, $X(t)$ 所取的值称为过程在 t 时刻所处的状态。通常状态空间为一维或 d 维欧氏空间, 也可为抽象的可测空间。

随机过程的基本概念有:

(1) 有穷维分布族。刻划过程的一族随机变量取值的概率分布规律。过程的一族随机变量有无穷多个, 对一切 $n \geq 1$, 从中任取 n 个, 这 n 个随机变量有(联合)分布函数, 这样的分布函数全体称为过程的有穷维分布族, 具体地说, 设 $X = \{X(t), t \in T\}$ 是一随机过程, 从 T 中任取 n ($n \geq 1$) 个 $t_1, t_2, \dots, t_n$, 对应的随机变量 $X(t_1), X(t_2), \dots, X(t_n)$ 的(联合)分布函数为

$$F_{t_1, t_2, \dots, t_n} = F_{t_1, t_2, \dots, t_n}(X_1, X_2, \dots, X_n) = P(X(t_1) \leq X_1, X(t_2) \leq$$

$X_2, X(t_n) \leq X_n$)。称 $\{F_{t_1, t_2, \dots, t_n}, n \geq 1, t_i \in T, 1 \leq i \leq n\}$ 为过程 X 的有穷维分布族。它满足下面相容性条件:

(1) 对 $(1, 2, \dots, n)$ 的任一排列 $(i_1, i_2, \dots, i_n)$,

$$F_{t_1, t_2, \dots, t_n}(X_1, X_2, \dots, X_n) = F_{t_1, t_2, \dots, t_n}(X_{i_1}, X_{i_2}, \dots, X_{i_n});$$

(2) 若 $m < n$, 则

$$\begin{aligned} & F_{t_1, t_2, \dots, t_m}(X_1, X_2, \dots, X_m) \\ &= \lim_{\substack{X_{i_{m+1}} \rightarrow \infty, \dots, X_{i_n} \rightarrow \infty}} F_{t_1, t_2, \dots, t_n}(X_1, X_2, \dots, X_n). \end{aligned}$$

反之, 由柯尔莫哥洛夫定理: 设已给 T 上的一族分布函数 $F = \{F_{t_1, t_2, \dots, t_n}, n \geq 1, t_i \in T, 1 \leq i \leq n\}$, 如果它满足上面相容性条件, 则必存在定义在某概率空间 (Ω, F, P) 上的过程 X , 使 X 的有穷维分布族重合于 F 。

(2) 期望函数和协方差函数。设过程 $X = \{X(t), t \in T\}$, 如果对一切 t , 随机变量 $X(t)$ 的期望存在, 则称函数 $m(t) = EX(t), t \in T$, 为 X 的期望函数或均值函数。如果对一切 $s, t \in T$, 随机变量 $X(S')$ 和 $X(t)$ 的协方差存在, 则称函数 $\Gamma(s, t) = E[X(s) - EX(s)][X(t) - EX(t)]$, 对一切 $s, t \in T$, 为 X 的协方差函数。

(3) 过程随机连续。设过程 $X = \{X(t), t \in T\}$, 如果对任一 $t_0 \in T$, 当 t 趋于 t_0 时, $X(t)$ 依概率收敛于 $X(t_0)$, 即 $X(t) \xrightarrow{P} X(t_0) (t \rightarrow t_0)$ 。

(4) 两过程等价。设 $X = \{X(t), t \in T\}$ 和 $Y = \{Y(t), t \in T\}$ 是定义在 (Ω, F, P) 上的两个过程, 如果对任何 $t \in T, P(X(t) = Y(t)) = 1$, 则称 X 与 Y 等价, 这时, 它们有相同的有穷维分布族。

(5) 停时。在一个随机过程中, 某种现象发生的时间。这个时间是随机的, 故为一个取值于时间参量集合 T 的随机变量。停时

有严格的数学定义。停时是过程发展进程中引进的一个非常重要的概念,它的出现大大促进了过程的纵深发展。

根据一族随机变量之间的不同关系,出现了许多各具特色的随机过程,最重要的有以下几种:

(1)马尔科夫过程。具有下述特性的随机过程:在已知过程现在所处状态的条件下,它将来处于什么状态不依赖于它过去曾经处于过什么状态。这种在已知“现在”的条件下,“将来”与“过去”独立的特性称为马尔科夫性或无后效性。

(2)离散时间马尔科夫链。参量集是离散的时间集,状态空间也是离散集的马尔科夫过程。通常表示为 $\{X_n, n \geq 0\}$, n 为非负整数,状态空间 $E = \{0, 1, \dots, M\}$ 或 $E = \{0, 1, 2, \dots\}$ 或 $E = \{\dots - 2, -1, 0, 1, 2, \dots\}$ 。“ $X_n = k$ ”表示过程在时刻 n 处于状态 k 。这时马尔科夫性表示为:对任意的 $0 \leq n_1 < n_2 < \dots < n_k < m, n > 0, i_1, i_2, \dots, i_k, i, j \in E$, 有 $P(X_{n+m} = j | X_{n_1} = i_1, X_{n_2} = i_2, \dots, X_{n_k} = i_k, X_m = i) = P(X_{n+m} = j | X_m = i)$ (*) 只要其中条件概率有意义。如果(*)式与 m 无关,则称它为齐次的马尔科夫链。称 $p_{ij} =$

$P(X_{m+1} = j | X_m = i)$ 为一步转移概率,即过程在时刻 m 时处于状态 i , 经过一步(即下一时刻)过程转移到状态 j 的概率,对一切 $i, j \in E$, 全体 p_{ij} 可表为矩阵

$$P^{(1)} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0n} & \cdots \\ p_{10} & p_{11} & \cdots & p_{1n} & \cdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} & \cdots \\ \dots\dots\dots\dots\dots\dots \end{bmatrix}$$

简记 $P^{(1)} = [p_{ij}]$, 称为链的一步转移概率矩阵。

对齐次马尔科夫链,从状态 i 出发,经 n 步转移到状态 j 的概率称为 n 步转移概率,记为 $p_{ij}^{(n)}$,它必然是从 i 出发,先经过 $r (< n$

的整数)步转移到到某个状态 k , 然后再从状态 k 出发(与过去无关)经 $s(r+s=n)$ 步再转移到状态 j , 用以式表述如下:

$$P_{ij}^{(n)} = \sum_{k \in E} p_{ik}^{(r)} p_{kj}^{(s)},$$

此式称为柯尔莫哥洛夫-查普曼方程。由它可得 n 步转移概率矩阵 $P^{(n)} = [p_{ij}^{(n)}] = [P^{(1)}]^n$ 。再有称 $P(X_0=i), i \in E$ 为链的初始分布, 由它和 $n \geq 1$ 步转移概率可完全确定过程的有穷维分布族, 因此一步转移概率矩阵在研究齐次马尔科夫链中起重要的作用。

对于马尔科夫链, 需对其状态的特性进行研究。如果由状态 i 出发, 经有限步转移一定能返回状态 i , 则称 i 为常返态。否则称为瞬时态, 如果链的一个状态集合 C , 一旦过程到达 C 中某一状态, 它将永远在 C 中转移, 则称 C 为链的闭集, 对闭集 C , 如果从 C 中任一状态出发, 经有限步转移, 均可到达 C 中其他任一状态, 则称 C 为不可约闭集。可以证明, 链的状态空间可分解为由一切瞬时态组成的集合 E_0 和若干个由常返态组成的不可约闭集 E_α 的并。这样, 在链的转移中, 它或者总在 E_0 中转移或者转移到某个常返类 E_α 中, 只要它一旦转移到某 E_α 中, 它将永远在 E_α 中转移。

(3) 连续时间马尔科夫链。时间参量为连续的, 状态空间为离散的马尔科夫过程, 记为 $\{X_t, t \geq 0\}$ 。这时马尔科夫性表示为: 对一切 $0 \leq t_1 < t_2 < \cdots < t_k < s, t > 0, i_1, i_2, \cdots, i_k, i, j \in E$, 有 $P(X_{s+t}=j | X_{t_1}=i_1, 1 \leq l \leq k, X_s=i) = P(X_{s+t}=j | X_s=i)$ 。

若这式与 s 无关, 则称链为齐次的。

对齐次连续参数马尔科夫链, 称 $p_{ij}(t) = P(X_{s+t}=j | X_s=i)$ 为转移概率。称矩阵 $P(t) = [p_{ij}(t)]$ 为链的转移矩阵。 $P(t)$ 满足条件:

(1) $p_{ij}(t) \geq 0$, (2) $\sum_{j \in E} p_{ij}(t) = 1$; (3) 柯尔莫哥洛夫-查普曼方程 $p_{ij}(s$

$+t) = \sum_{k \in E} p_{ik}(s) \cdot p_{kj}(t), s, t > 0$ 。通常还假定: (4) 标准性 $\lim_{t \downarrow 0} p_{ij}(t) = p_{ij}(0) = \delta_{ij}$, 其中 $\delta_{ii} = 1, \delta_{ij} = 0 (i \neq j)$ 。

研究连续时间马尔科夫链, 常常用到矩阵 $Q = [q_{ij}] = [p'_{ij}(0)]$, 其中 $p'_{ij}(0)$ 表 $p_{ij}(t)$ 在 $t=0$ 处的右导数。 $p'_{ij}(0)$ 存在, 当 $i \neq j$ 时, $p'_{ij}(0)$ 非负有穷, 但 $p'_{ii}(0)$ 可能为无穷。称 Q 为链的密度矩阵, 又称 Q 矩阵。

(4) 一般马尔科夫过程。时间参量集为连续的, 状态空间为一可测空间 (E, B) 的马尔科夫过程, 记为 $\{X_t, t \geq 0\}$ 。这时, 马尔科夫性需用条件概率来表述。马尔科夫过程的许多性质可通过转移函数 $p(s, X, t, A)$ 的研究来得到, 转移函数 $p(s, X, t, A)$ 是一个四元函数 ($0 \leq s < t, X \in E, A \in B$), 它表示过程在时刻 s 处于状态 X , 到时刻 t 时转移到状态集 A 中的概率。如果它只依赖于 $t-s, X$ 和 A , 则称过程为齐次的, 这时, 转移函数写成 $p(t, X, A)$, 它表过程从状态 X 出发, 经过时间 t 转移到状态集 A 中的概率。

(5) 强马尔科夫过程。具有强马尔科夫性的随机过程。通常马尔科夫性中的“现在”是指某个固定时刻, 如果把“现在”换成停时 (这是一随机时间), 在已知“现在”的条件下, “将来”与“过去”仍独立, 这种特性称为强马尔科夫性。若过程具有强马尔科夫性则一定具有通常的马尔科夫性。

【平稳过程】某种统计特性在推移任意一段时间后仍不发生变化的随机过程。这种过程随时间变化的情况不仅与当时所处的状态有关, 还与过去的情况有着不可忽视的联系, 越相近, 联系越紧密。例如, 一台稳定工作的纺纱机纺出的纱的截面积大小, 它是一个随时

间变化的随机过程,由于纺纱机稳定工作,可以认为随时间推移,其统计特性不会发生变化,故是一平稳过程。由于纱是一束纤维组成的,因此,前面纺出的纱的截面积大小必然会影响后面的状况,而且当时间相距越近,影响就越显著。

设随机过程 $X = \{X(t), t \in T\}$, 其中 T 为整数或实数全体。如果对任意自然数 n 及任意 $t_1 < t_2 < \cdots < t_n, s \in T$, 随机向量 $(X(t_1), X(t_2), \cdots, X(t_n))$ 推移任意一段时间 S 后, 其统计特性——概率分布不变, 即 $(X(t_1 + s), X(t_2 + s), \cdots, X(t_n + s))$ 与 $(X(t_1), X(t_2), \cdots, X(t_n))$ 的联合分布函数相同, 则称 X 为严平稳过程或狭义平稳过程。如果对一 $t, E|X(t)|^2 < \infty$, 而且 X 的期望函数为常数, 即 $EX(t) = m$; X 的协方差函数推移任意一段 τ 时间不变, 即 $E(X(t + \tau) - m)(X(\tau) - m) = \Gamma(t)$ (与 τ 无关), 则称 X 为宽平稳过程或广义平稳过程。它是 X 的统计特性——均值函数和协方差函数推移任意一段时间不变。对于一严平稳过程 $X = \{X(t), t \in T\}$, 只要对一切 t , 二阶矩 $E[X(t)]^2$ 有穷, 它一定是宽平稳过程。

将傅里叶分析方法用于宽平稳过程, 可以把过程表成有不相关随机振幅的简谐波的叠加, 这样便能将过程按频率进行谱分解。因此, 这种过程的理论是处理随机信号等问题的重要工具, 故在无线电技术和自动控制等领域中有着广泛的应用。

【泊松过程】随着时间增长, 累计随机事件发生次数的一种最基本的随机过程。例如, 随着时间增长, 累计某电话交换台收到呼唤的次数; 累计放射物质放射的粒子数, 等等, 均可用泊松过程来描述。一般地, 从直观上看, 只要随机事件在不相交时间区间是独立发生的, 而且在充分小的区间上最多只发生一次, 它们的累计次数就可用泊松过程来描述。从数学上讲, 设随机过程 $X = \{X(t), t \geq 0\}$ (对

每个 $t \geq 0$, $X(t)$ 表示在时间间隔 $[0, t]$ 间随机事件累计发生的次数, 状态空间 E 为非负整数值的集合, 如果满足以下条件: (1) $P(X(0)) = 1$; (2) 过程在不相交区间上的增量相互独立, 即对一切 $0 \leq t_1 < t_2 < \cdots < t_n$, 增量 $X(t_1), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$ 是 n 个相互独立的随机变量; (3) 增量 $X(t) - X(s) (t > s)$ 的概率分布为只依赖于 $t - s$ 的泊松分布, 即分布列为

$$P(X(t) - X(s) = k) = \frac{[\lambda(t-s)]^k}{k!} e^{-\lambda(t-s)}, k = 0, 1, 2, \cdots,$$

其中 $\lambda > 0$, 则称 X 为齐次泊松过程。简称泊松过程, 其中 λ 称为泊松过程的强度。因为 $EX(t) = \lambda t$, 所以 λ 表示单位时间内随机事件平均发生的次数。泊松过程不仅是具有独立增量的过程, 而且具有马尔科夫性, 故又为连续时间马尔科夫链。此外, 泊松过程还有如下重要特性: (1) 泊松过程 X 的样本函数几乎都是跃度为 1 的左 (或右) 连续的阶梯函数, 反之, 若一过程的样本函数具有如上性质, 加上它又是随机连续的、具有独立增量的过程, 则它必是泊松过程; (2) 设序列 $\{T_n, n \geq 1\}$ 表示随机事件依次发生的时间, 令 $\tau_n = T_n - T_{n-1} (n \geq 1)$, 它表示随机事件第 $n-1$ 次发生与第 n 次发生的时间间距, 于是 $\{\tau_n, n \geq 1\}$ 为随机事件依次相连两次发生的时间间距序列, X 是强度为 $\lambda > 0$ 的泊松过程的充要条件是: 随机序列 $\{\tau_n, n \geq 1\}$ 相互独立同分布, 而且分布是参数为 $\lambda > 0$ 的指数分布, 即 $P(\tau_n > t) = e^{-\lambda t}, t \geq 0$; (3) 对任何 $n \geq 1, T_n$ 服从参数为 n, λ 的伽玛分布, 即 T_n 的密度函数

$$p(t) = \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t}, t \geq 0;$$

(4) 固定 $t, X(t)$ 服从参数为 λt 的泊松分布。在已知 $X(t) = k$ 的条

件下, k 个时刻 $T_1 \leq T_2 \leq \dots \leq T_k$ 的联合分布, 即 $(T_1, T_2, \dots, T_n)$ 的概率分布与 k 个在 $[0, t]$ 上均匀分布且相互独立的随机变量的次序统计量的分布相同, 即设 $U_1, U_2, \dots, U_k$ 都在 $[0, t]$ 上均匀分布且相互独立, 它的次序统计量为 $U_{(1)} \leq U_{(2)} \leq \dots \leq U_{(k)}$, $(T_1, T_2, \dots, T_k)$ 与 $(U_{(1)}, U_{(2)}, \dots, U_{(k)})$ 同分布。

齐次泊松过程是最简单的泊松过程, 此外还有如下几种:

(1) 非齐泊松过程。如果齐次泊松过程的强度 λ 换成与 t 有关的函数 $\lambda(t)$, 其它条件不变, 只是增量的分布列变为:

$$P(X(t) - X(s) = k) = \frac{[\Lambda(t) - \Lambda(s)]^k}{k!} e^{-[\Lambda(t) - \Lambda(s)]}, \quad k = 0, 1, 2, \dots, \text{其中 } \Lambda(t) = \int_0^t \lambda(s) ds,$$

则称 X 为非齐次泊松过程, $\lambda(t)$ 为其强度。非齐次泊松过程可通过时间尺度变换变成齐次泊松过程。

(2) 复合泊松过程。若 $\{X_n, n \geq 1\}$ 是一独立同分布的随机变量序列, $X = \{X(t), t \geq 0\}$ 是一齐次泊松过程, 令 $Y(t) = \sum_{n=1}^{X(t)} X_n$, 则称过程 $\{Y(t), t \geq 0\}$ 为复合泊松过程。

(3) 滤过泊松过程。称过程 $\{Y(t), t \geq 0\}$ 为滤过泊松过程, 如果

$$Y(t) = \sum_{n=1}^{X(t)} W(t, T_n, X_n), \text{ 其中}$$

$X = \{X(t), t \geq 0\}$ 是一齐次泊松过程; $\{T_n, n \geq 1\}$ 是过程 X 中随机事件依次发生的时间序列; $\{X_n, n \geq 1\}$ 是过程 X 中随机事件依次发生时的量值序列, 它们独立同分布; $W(t, T, X)$ 是一个三元函数, 通常称为响应函数。

由齐次泊松过程可以构造许多实际中有用的随机过程, 因而在随机过程中占有特殊的地位, 故有人称它为随机过程的基石。

【布朗运动】又称维纳过程。它描述了轻质的悬浮微粒由于周围弥散的介质分子(如水、气体分子等)碰撞而产生的随机涨落的无规则运动。其数学定义如下:设随机过程 $X = \{X(t), t \in R\}$

($X(t)$ 表示微粒在 t 时刻所处位置在某个轴上的坐标), 状态空间为 R 。如果 X 满足条件:(1) $X(0)=0$;(2)具有独立增量, 即对任意的 $0 \leq t_0 < t_1 < \dots < t_n$, 增量 $X(t_0), X(t_1) - X(t_0), \dots, X(t_n) - X(t_{n-1})$ 是相互独立的随机变量;(3)对任意 $t \geq 0, s > 0$, 增量 $X(t+s) - X(t)$ 为(一维)正态变量, 即它的密度函数为

$$p(s, X) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{X^2}{2s}} \text{ (与 } t \text{ 无关);}$$

(4) X 的一切样本函数为连续函数, 则称 X 为(一维)布朗运动。它既是马尔科夫过程, 又是鞅。

一维布朗运动的所有样本函数尽管处处连续, 但是几乎所有(即以概率 1)的样本函数处处不可微和在任一区间上不是有界变差函数。

如果上述定义中状态空间 R 换成 R^d (d 为大于 1 的整数), 而且满足上面类似的条件, 则称 X 为(d 维)布朗运动。

【鞅】起源于对“公平”赌博过程作数学描述。若 X_t 表示赌徒在 t 时刻拥有的赌金, 则随机过程 $X = \{X_t, t \in T\}$, 其中 $T = \{0, 1, \dots\}$ 或 $(0, \infty)$, 描述了这个赌徒在赌博中赌金的变化过程。设赌徒在 S 时刻拥有赌金 X_s 和到 s 时刻为止所获得的赌博经验, 在此基础上再往下赌。如果在未来时刻 t , 赌徒拥有的赌金平均(或期望)仍是 X_s , 即赌徒不会因有前面获得的赌博经验而使后来拥有的赌金增加或减少, 则称它是一种“公平”的赌博。这种“公平”特性用数学

形式表示为: $E(X_t | X_u, u \leq s) = X_s$, 等式前面是 $X(t)$ 关于 $\{X_u, u \leq s\}$ 的条件期望, $\{X_u, u \leq s\}$ 理解为赌徒到 S 时刻为止所获得的赌博经验。

如果随机过程 $X = \{X_t, t \in T\}$ 具有“公平”特性, 则称 X 为鞅。如果 X 不具备“公平”特性, 而满足: $E(X_t | X_u, u \leq s) \leq X_s$ 或 $E(X_t | X_u, u \leq s) \geq X_s$, 则称满足前者的 X 为上鞅; 满足后者的 X 为下鞅。它们描述的是“不公平”的赌博。上鞅表示赌徒的赌金平均地看是减少的, 故利于这个赌徒。下鞅则相反。

鞅的基本理论是以鞅的收敛定理和分解定理为核心内容。鞅的基本定理和方法现已成为研究其它随机过程的有力工具。人们对局部鞅和半鞅等的研究, 推动了随机积分和随机微分方程的发展。

新兴现代数学

运 筹 学

【运筹学】近代应用数学的一个分支。研究对象极为广泛, 主要是把生产、管理等过程中一些具有普遍性的运筹问题抽象为数学模型, 然后用数学方法加以解决。它包括众多的分支: 数学规划; 图论、网络流; 排队论; 对策论; 决策分析; 可靠性数学理论; 库存论; 搜索

论;模拟理论等。

数学规划包括线性规划、整数规划、0-1 规划、非线性规划、动态规划、多目标规划、参数规划、组合规划、几何规划和随机规划等。一般来说,求目标函数 $f(X)=f(X_1, X_2, \dots, X_n)$ 在约束条件 $X=(X_1, X_2, \dots, X_n) \in A \subseteq R^n$ 下的极值问题称为数学规划问题,简记为 (P)。但数学规划与古典的极值问题有本质的不同,其目标函数和约束条件都很复杂,很难用一个简单的公式求解,有效的途径是设计一种算法,通过迭代求出达到一定精度的数字解。在问题 (P) 中,若 $f(X)=\sum c_i X_i; A=\{X | \sum_{i=1}^n a_{ij} X_i = b_j, j=1, 2, \dots, m; X \geq 0\}$, 则称为线性规划,记为 (LP)。在 (LP) 中,若 $X_1, X_2, \dots, X_n$ 全部或部分只允许取整数,则称整数规划;若它们只能取值 0 或 1,则称为 0-1 规划。在问题 (P) 中,若目标函数或约束中出现非线性关系,则称非线性规划;若目标函数多于 1 个,则称多目标规划;

线性规划是运筹学中较早发展起来的分支,其理论与方法都比较成熟,应用也最广泛。下面用简例说明它的某些基本概念。

某工厂生产每吨甲种产品需原料 A、B 分别为 3 吨与 2 吨,生产每吨乙种产品需原料 A、B 各为 2 吨与 5 吨。生产甲、乙两种产品每吨利润分别为 2 万元与 3 万元。又受销售限制,甲种产品生产不能超过 3 吨,A、B 两种原料供应各不超过 12 吨与 20 吨,问生产甲、乙产品各多少吨时,其利润最大?

设生产甲、乙产品分别为 X_1 和 X_2 吨,由题意, X_1, X_2 可受的约束为 $3X_1+2X_2 \leq 12, 2X_1+5X_2 \leq 20, X_1 \leq 3, X_1, X_2 \geq 0$ 。目标函数为 $f=2X_1+3X_2$ 。要回答以上问题,只需求目标函在约束条件下的最大值就行了。

满足约束条件的点称为可行解,可行解的全体所成的集合称

为可行域。如图所示,本例中可行域为五边形 ABCDO 围成的区域。直线 $f(X)=d$ 称为目标函数的等值线(图中用虚线表示),等值线越往右移,目标函数的值就越大,显然,在顶点 $C(\frac{20}{11}, \frac{36}{11})$ 处,目标函数取到在可行域上的最大值。因此,最优解为 $X_1=\frac{20}{11}, X_2=\frac{36}{11}$,最大利润为 $f=\frac{20}{11} \times 2 + \frac{36}{11} \times 3 = \frac{148}{11}$ (万元)。

上例说明,线性规划的可行域为凸集,最优解对应于可行域的某一顶点。这一性质,对一般情况也是成立的。即若问题(LP)有最优解,则必定可在它的可行域的某一顶点上达到。求解(LP)最基本最有效的方法是单纯形法(G. B. 丹齐克,1947),它的基本步骤是适当选取可行域的一个顶点做为初始点,然后按一定的规则转到另一顶点上(迭代),使目标函数得到改进,再通过一定准则判断它是否为最优解,直到达到最优解或判定不存在最优解而停止迭代。

【图论】图论是数学的一个分支,属离散数学范畴。图论开始于1736年 Euler 解决哥尼斯堡七桥问题,它的真正起飞是20世纪50年代,70年代方获得它的大量追随者。

图论是描述处理离散系统的一个有力工具,在计算机科学、信息科学、生物学等高科技领域有着广泛的应用。

一个图 G 由一个非空的被称为顶点的元素的集合 $V(G)$ 和一个被称为边的元素的集合连同它们之间的关系组成。一个严格的数学定义为:设 V 是一个非空集合, $E \subseteq 2^V$, 序对 (V, E) 被称为一个图,如果对任意 $e \in E$ 有 $|e| \leq 2$ 。如果 $e = (u, v)$ 则称 u 和 v 是边 e 的端点,也称 u, v 同边 e 关联。如果对图的边或顶点赋权,则

称它是一个网络。如对边 e 的元素数目没有上述限制,则称 (V, E) 为一个超图。

近四十年来,图论获得了蓬勃发展,已形成了许多分支,主要有:基础图论、代数图论、拓扑图论、随机图论、极值图论、网络理论、超图理论、拟阵理论等。图论中已取得了若干杰出成就:Appel 和 Haken 证明了四色定理,Robertson 和 Seymour 发展了丰富而漂亮的图子式(Graph minor)理论。

基础图论主要研究图的路结构,圈结构、连通性质等。图的 Hamilton 圈问题同货郎问题密切联系,后者是一个重要的组合优化问题。人们对 H —图做了许多研究,给出了若干 H —圈存在的充分条件,早期的主要结果是 Dirac 定理和 Ore 定理,80 年代初 Fan 推广了 Ore 定理,得到了一个更一般性的充分条件,人们称之为 Fan 条件。该领域有许多猜想和问题有待解决。

匹配理论在组合优化中十分有用,取得了充分发展,关于偶图匹配,主要有 König 定理和 Hall 定理。一般图的完美匹配的存在性由 Tutte 解决了。Tutte 进一步得到 f —因子的存在性定理, Lovasz 给出了 (g, f) —因子定理。Petersen 早年得到了 2—因子分解定理。因子理论成为一个十分活跃的研究课题。许多组合学、代数学及其它数学结构都同图因子和图分解有密切联系,显示了它在理论上的重要性。另外,图因子的结果被应用到编码、组合设计、计算机和通讯网络、雷达天线设计等领域。一类重要图分解是把一个图分解为同构子图。Basak 的专著“Decompositions of Graphs”总结了这一领域的主要成果。图的同构因子分解同图的对称性密切相关。Alspach 猜想:所有顶点—传递图都是有理的。作为图类有下述关系: $\{\text{完全图}\} \subset \{\text{完全多部图}\} \subset \{\text{循环图}\} \subset \{\text{Cayley 图}\} \subset \{\text{顶点—传递图}\}$ 。到目前为止解决到 $\{\text{完全多部图}\}$ 的有

理性问题。

代数图论主要研究图的谱性质和对称性质, Biggs 的书“Algebraic Graph Theory”总结了这一领域的成果。Ringel 和 Youngs 解决了经典的 Heawood 地图着色问题, 这是拓扑图论的转折点。随后, 大量的推广和分析 Ringel 和 Yoangs 定理的方法研究涌现出来, 并且所产生的技巧应用到了许多其它图嵌入问题。White 的书“Graphs, Groups and Surfaces”, 把拓扑图论系统化了。Gross 和 Tucker 的书“Topological Graph Theory”做了进一步的完善。图论近三十年来最引人注目的是随机图论的发展。这一理论由 Erdős 和 Renyi 于 50 年代末 60 年代初奠定了基础。三十年来, 发现了源于概率论的许多工具, 例如随机游动、鞅、分支过程、马尔可夫链等等可用于图论, 于是随机图论开始繁荣, Bollobás 的专著“Random Graphs”, 把随机图系统化了。

超图(Hypergraphs)理论是研究有限集合族的组合性质的。超图是被当做图的自然推广的集合系统。面对离散数学快速增长着的用途, 1960 年组合数学家积极主张发展集合的系统组合工具。1966 年, Berge 提出了术语“超图”, Erdős 和 Hajnal 的先驱性论文“On the Chromatic Number of Graphs and Set Systems”, 建立了超图理论的前导。一般超图的结构十分复杂, 人们通常把某些超图问题化为图的问题进行研究。另外, 对某些特殊类型的超图, 例如 balance 超图, 具有 Helly 性质, König 性质的超图, 进行了较深入的研究。最近 Wang 和 Lee, 从关系数据库的拓扑结构出发, 提出了一些超图问题, 建立了超图与偏序集间的关系, 引入了一些组合不变量, 并建立了它们同 Möbius 函数, Betti 数, Euler 示性数等之间的关系, 研究了无圈超图的性质和一般超图的圈结构等一些很荃员的问题。

目前图论的一个发展趋势是同其它数学分支互相渗透,在图论中越来越多地使用其它数学分支的成果,例如,有限域上曲线的 Rieman 假设已被多次使用,上同调理论成为 Chung 和 Graham 关于伪随机超图的工作动力。

图论的最大优势或许在于它具有众多自然的漂亮的问题有待解决。Hilbert 强调问题对一门数学分支的重要性:“只要一门科学分支能提出大量问题,它就充满着生命力;而问题的缺乏则预示着独立发展的衰亡或中止”。什么是图论中的大问题呢?人们认为明显称得上这样的问题有两个,1. NP 问题是否等于 P 问题。2. Hadwiger 猜想。前者为整个数界尽人皆知。后者组合界所有人都熟悉。该猜想为:每个 K -色图可“收缩”为一个包含 K 阶完全图的图。

Erdős 提出了许多有名的组合图论问题。

另一个趋势是图论同计算机科学越来越密切,人们十分重视计算机科学中提出来的问题、重视寻求图论在计算机科学中的生长点。Steiner 树问题、搜索树问题,图的厚度和交叉数、图的带宽等都是来源于计算机科学。

【网络流】1955 年哈里斯提出在一个给定的网络上寻求两点间最大运输量问题,1956 年福特和富尔克森给出了算法,从而建立了网络流理论。

最短路问题是网络流理论的重要问题之一,它可以直接应用于解决诸如管道铺设、线路安排、厂区布局等实际问题。设 $G=(V,E)$ 为有向图,选定起点 v_1 和终点 v_p ,对每条边 $e=(e_i,e_j)\in E$,给定一个数 w_{ij} (称为该边的权),就得到一个赋权有向图,最短路问题就是找一条从 v_1 到 v_p 的路,它的权(各边权之和)最小。1959

年 E. W. Dijkstra 提了一种较好的算法。

网络是由有向图 $G=(V, E)$ 和 E 上的一个非负函数 C 构成, 设 v_1 和 v_p 分别为发点与收点, c_{ij} 为函数 c 在弧 $e=(v_i, v_j)$ 上的值, 网络上的流就是定义在 E 的函数 f , 它满足 (1) $0 \leq f_{ij} \leq c_{ij}$, 这里 f_{ij} 是 f 在弧 $e=(v_i, v_j)$ 上的值; (2) 对于任何一个中间点 $v_i (i \neq 1, p)$, 从 v_i 流出的总量等于流入 v_i 的总量, 即 $\sum_j f_{ij} = \sum_k f_{ki}$ 。在网络上求一个使流量 $v(f)$ (即从 v_1 流出的总量与流入 v_1 的总量之差) 最大的流 f , 称为网络的最大流问题, 是网络流的主要课题。福特的富尔克森提出的标号法是求最大流的一种有效算法。

网络流的另一问题是最小费用流问题, 即在给定的网络上求流 f , 使流量 $v(f)$ 等于给定的数值 v^* , 且使费用 $\sum a_{ij} f_{ij}$ 最小, 这里 $\sum$ 对所有的弧求和, a_{ij} 是单位流量在弧 (v_i, v_j) 上的费用, 求在已知运输网络上将一批货物从一地点运到另一地点的最小费用就属于此类问题。现在已有几种较为有效的求最小费用流的方法。

【排队论】研究服务系统中排队现象随机规律的一门学科。任何服务系统都由三个基本部分组成: 输入过程、排队规则和服务机构。

输入过程即顾客到来的规律。主要类型有 (1) 定长输入, 即每隔一定时间到达一位顾客; (2) 最简单流, 又称泊松流, 即顾客相继到达的间隔时间是相互独立的, 且具有相同的负指数分布; (3) 埃尔朗输入, 即顾客到达的间隔时间相互独立, 且有相同的埃尔朗分布; (4) 一般独立输入。

排队规则主要有: (1) 损失制, 顾客到达后, 若不能得到服务就立即离去; (2) 等待制; (3) 混合制, 限制队长 N , 顾客到达时, 若等

待队长小于 N , 则等待服务, 否则就离去。

服务机构也有多种类型: 服务台的个数有一个与多个之分; 组织形式有串联与并联的区别; 服务方式可以是单个服务或成批服务; 服务时间则有定长分布、负指数分布、埃尔朗分布及一般独立分布等不同类型。

不同的输入、排队规则和服务机构可以组合成许多不同的服务系统的模型, 它们都有特定的符号, 例如 $M/M/n$ 表示泊松输入、负指数服务分布、 n 个服务台的系统; $M/G/1$ 表示泊松输入、一般服务分布、1 个服务台的服务系统等。

如果在系统运行不太长的一段时间内去考察系统的数量指标, 一般它们要依赖于初始条件与运行时间, 系统此时的状况称为瞬时性态。在系统运行了相当长的时间之后, 它的各项数量指标的变化规律不再依赖于初始条件和运行时间, 此时系统的状况称为平稳性态。

20 世纪 50 年代以前主要是对系统的平稳性态的研究。最著名的结果是损失制的 $M/M/n$ 系统的埃尔朗公式, 即顾客到达后, 由于 n 个服务台都被占用失去顾客造成损失的概率为

$$P_n = \frac{\frac{(\lambda/\mu)^n}{n!}}{\sum_{j=1}^n \frac{(\lambda/\mu)^j}{j!}}$$

式中 λ 为单位时间内到达顾客的平均数, μ 为单位时间内每一服务台对顾客完成服务的平均人数。

这一公式对系统的性能分析与设计都很重要。例如在电话系统中, 统计检验已经证实, 呼唤流是最简单流, 会话时间服从负指数分布, 因此只要统计出单位时间收到呼唤次数的平均数 λ 及会

话的平均时间 $1/\mu$, 根据交换台的线路数目 n , 就能算出呼损率的大小, 反之, 如果先规定呼损率的大小, 例如要求小于 5% , 即平均 1000 次通话中最多有 5 次打不通, 则由此公式可求出所需要的最少的线路数目 n 。

另一著名公式是辛钦-波拉采克公式, 它主要包括两个公式: 其一是 $M/G/1$ 系统中顾客的平均数 $Q = \rho + (\rho^2 + \lambda^2 \sigma^2) / (2(1 - \rho))$, 式中 λ 为单位时间内到达顾客的平均数, $\rho = \lambda / \mu$ 称为服务强度, $1/\mu$ 为每个顾客的平均服务时间, σ^2 服务时间的方差。可以看出, 在 μ 不变的情况下, 减小 σ 也能使系统中顾客的平均数下降, 其二是 $\mu/G/1$ 系统中等待时间的拉普拉斯变换 $W^*(s) = (1 - \rho)s / (s - \lambda) \times B^*(s)$, 式中 λ, ρ 的意义同上, $B^*(s)$ 是服务时间分布的拉氏变换, 由此公式能求出等待时间的分布。

从 20 世纪 50 年代中期开始, 瞬时性态的研究逐渐成为排队论研究的焦点, 对此, 中国的数学工作者做出了重要贡献。

模糊数学

【模糊数学】研究和处理模糊性现象的数学。1965 年, 美国控制论专家查德首次提出了模糊集合的概念, 这标志着模糊数学的诞生。模糊性主要是指客观事物的差异在中介过渡时所呈现的“亦此亦彼”性。在普通集合论中, 一个对象对于一个集合, 要么属于, 要么不属于, 二者必居其一, 且仅居其一, 绝不模棱两可。这就限定了普通集合论只能表现“非此即彼”的现象。一个概念有其内涵和外延, 所谓内涵是指符合此概念的对象所具有的共同属性, 而外延是指

此概念的对象组成的集合,集合可以表现概念,集合的运算和变换可以表现判断和推理。因此,建立在集合论上的现代数学成为可以描述和表现各门学科的形式语言和系统。然而一个模糊的概念却没有明确的外延,例如“年轻人”就是一个模糊概念,对于一个人来说,他是否属于“年轻人”无法明确回答。打破普通集合论中元素对集合的绝对隶属关系,在“ $u \in A$ ”与“ $u \notin A$ ”之间考虑其中间状况,提出“隶属度”的思想是解决问题的关键之处。

给定一个论域 U (讨论范围),所谓指定了 U 上的一个模糊集合 A ,是指对任意 $u \in U$,都有一个数 $A(u) \in [0, 1]$ 与之对应。 $A(u)$ 叫做 u 对 A 的隶属度。这样,模糊集合 A 实际上是一个映射 $A: U \rightarrow [0, 1]$,称之为隶属函数。

以年龄作为论域,取 $U = [0, 100]$,“年轻人”这个模糊概念可用一个模糊集 Y 来表示,隶属函数规定为

$$Y(u) = \begin{cases} 0, & 0 \leq u < 50 \\ \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1}, & 50 \leq u \leq 100 \end{cases}$$

设 A, B 是两个模糊集,规定它们的并,交,余运算如下:

$$(A \cup B)(u) = A(u) \vee B(u)$$

$$(A \cap B)(u) = A(u) \wedge B(u)$$

$$A^c(u) = 1 - A(u)$$

其中 $\vee = \max$, $\wedge = \min$ 。这些运算满足下列几条性质:

1) 幂等律: $A \cup A = A, A \cap A = A$

2) 交换律: $A \cup B = B \cup A,$

$$A \cap B = B \cap A$$

3) 结合律:

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

$$4) \text{ 吸收律: } A \cap (B \cup A) = A, \\ A \cup (B \cap A) = A$$

5) 分配律:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

$$6) \text{ 两极律: } U \cup A = U, U \cap A = A \\ \emptyset \cup A = A, \emptyset \cap A = \emptyset$$

$$7) \text{ 复原律: } (A^c)^c = A$$

8) 对偶律:

$$(A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c$$

模糊数学中的模糊逻辑与模糊控制已经在人工智能等领域中取得很多成果。在日本,以模糊推理为基本原理的模糊产品已经进入世界范围内的市场,如模糊洗衣机,模糊电冰箱,模糊空调器,模糊摄像机等家用电器深受用户喜爱;在美国,模糊控制原理已经用于高科技领域,如宇航系统,高精度过程控制,机器人等。

【最大隶属原则与模式识别的直接方法】设 $A_1, A_2, \dots, A_n$ 论域 U 上的 n 个模糊集, u_0 是 U 中一个元素,如果存在指标 $i \in \{1, 2, \dots, n\}$, 使得

$$A_i(u_0) = \max\{A_1(u_0), A_2(u_0), \dots, A_n(u_0)\}$$

则认为 u_0 相对隶属于 A_i 。

另一种形式为: 设 $u_1, u_2, \dots, u_n$ 是 U 中 n 个元素, A 是 U 上的一个模糊集。如果存在指标 $i \in \{1, 2, \dots, n\}$, 使得

$$A(u_i) = \max\{A(u_1), A(u_2), \dots, A(u_n)\}$$

则认为 u_i 相对隶属于 A 。

上面是最大隶属原则的两种形式,在实际问题中,若把模糊集作为标准模型,则最大隶属原则便构成了模式识别的直接方法。

【贴适度与择近原则】设 A, B 是 U 上的两个模糊集,记

$$A \circ B = \bigvee_{u \in U} (A(u) \wedge B(u))$$

$$A \odot B = \bigwedge_{u \in U} (A(u) \vee B(u))$$

分别叫做 A 与 B 的内积和外积。再记

$$(A, B) = \frac{1}{2} [A \circ B + (1 - A \odot B)]$$

称之为 A 与 B 的贴适度。

给定 U 上的 n 个模糊集 $A_1, A_2, \dots, A_n$, 及另一模糊集 B 。若存在指标 $i \in \{1, 2, \dots, n\}$, 使得

$$(B, A_i) = \max_{1 \leq j \leq n} (B, A_j)$$

则认为 B 与 A_i 最贴近。这就是择近原则。

在模式识别中,假定 $A_1, A_2, \dots, A_n$ 为标准模式, B 为待识别模式,可用择近原则判定 B 应归属哪个模式。

【模糊关系与聚类分析】对事物按一定要求进行分类的数学方法叫做聚类分析。现实的分类问题多伴随着模糊性。一个模糊的分类,要由一个模糊等价关系来确定。

所谓从论域 U 到论域 V 之间的一个模糊关系 R 是指, R 是 $U \times V$ 上的一个模糊集。隶属程度 $R(u, v), (u, v) \in U$, 叫做 u 与 v 具有关系 R 的程度, 当 $U = V$ 时, R 称为 U 上的模糊关系。

当 $U = \{u_1, \dots, u_n\}, V = \{v_1, \dots, v_m\}$ 均为有限集时, 模糊关系

R 可表示为一个矩阵: $R = (r_{ij})_{n \times m}$, 其中 $r_{ij} = R(u_i, v_j)$ 。这样的矩阵叫做模糊矩阵。

设 $R = (r_{ij})_{n \times m}$, $S = (s_{jk})_{m \times r}$ 为两个模糊矩阵, 令

$$q_{ik} = \bigvee_{j=1}^m (r_{ij} \wedge s_{jk}) \quad i=1, 2, \dots, n, \\ k=1, 2, \dots, r$$

称 $Q = (q_{ik})_{n \times r}$ 为 R 与 S 的复合矩阵, 记作

$$Q = R \circ S$$

设 R 是 U 上的一个模糊关系(矩阵), 如果它满足

- 1) 反身性: $r_{ii} = 1$
- 2) 对称性: $r_{ij} = r_{ji}$
- 3) 传递性: $R \circ R \subset R$

则称 R 为 U 上的一个模糊等价关系。

有一个结论: R 是 U 上的一个模糊等价关系当且仅当对任意 $\lambda \in [0, 1]$, R_λ 是 U 的普通等价关系。依据这些等价关系, U 中的元素便可以分类。且当 λ 由 1 下降到 0 时, 所分的类由“细”变“粗”, 逐渐归并, 形成一个动态的聚类图。

模糊聚类分析以下三个步骤:

步骤 1: 设 U 是被分类对象集, 建立 U 上的模糊相似关系 R 。

步骤 2: 将 R 改造成模糊等价关系 R^* 。

步骤 3: 由 R^* 作动态聚类图。

【模糊变换与综合评判】设论域 $U = \{u_1, u_2, \dots, u_n\}$, $V = \{v_1, v_2, \dots, v_m\}$ 均为有限集, 若 A, B 分别是 U, V 上的模糊集, 则它们均可表示为向量形式(叫做模糊向量):

$$A = (a_1, a_2, \dots, a_n), B = (b_1, b_2, \dots, b_m)$$

其中 $a_i = A(u_i)$, $b_j = B(v_j)$ 。设 R 是 U 到 V 的一个模糊关系, 则根据矩阵的复合运算, 由 R 确定了一个变换: 任给 U 上的一个模糊集 A 便可决定 V 上的一个模糊集 $B = A \circ R$ 。

把 U 作为因素集, V 作为评语集, R 是 U 到 V 的模糊关系, r_{ij} 表示从第 i 个因素着眼, 对被评对象作出第 j 种评语的可能程度。固定 i , $(r_{i1}, r_{i2}, \dots, r_{im})$ 是 V 上的一个模糊集, 表示从第 i 个因素着眼, 对于被评对象所作的单因素评判。因此 R 也就是单因素评判所组成的矩阵。若把 A 作为诸因素的权重分配, 则 B 就是综合了各因素后所作的综合评判。

信 息 论

【信息论】信息论是建立在这样一个概念的基础上, 即: 信息是可测的。所以信息论也称为“通信的数学理论”, 是研究信息的传输、存储、处理的科学。各式各样的信息传输系统, 就其本质可概括成统一的理论模型, 通常称它为通信系统模型(见右上框图):

信息论研究的主要问题是在通信系统设计中如何实现有效性和可靠性。而这一问题是由美国科学家仙农(C. E. Shannon)提出来的, 他在 1948 年发表了著名的论文《通信的数学理论》, 为信息论奠定了理论基础。当前, 对信息论的研究分划成三个类型:

(1) 狭义信息论。研究信息的测度、信道容量以及信源和信道编码理论。

(2) 一般信息论。除了(1)之外, 还研究噪声理论、信号滤波和预测, 统计检测与估计理论、调制理论以及信息处理理论等。

(3)广义信息论。除了(1)、(2)之外,还包括所有与信息有关的领域,如模式识别、计算机翻译、心理学、遗传学、神经生理学、语言学、语义学等有关信息的问题。

【信息量】信息是一个可测的抽象的量;它的值不取决于信息具体讲些什么。信息是与一个事件发生的各种可能结果的集合有关,它的值取决于与这些结果相联系的概率。当两个事件具有相同数目的结果,并且和这两个事件结果集合联系着的两组概率均一一对应相等,那么称这两个事件有相等的信息量。若和这两个事件结果集合联系着的两组概率不相等,我们必须推导出一个只能以一个数目来表示对信息量的测度。设随机事件,用随机变量 X 表示。

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ P(a_1)P(a_2)\cdots P(a_n) \end{bmatrix}$$

其中 $\sum_{i=1}^n P(a_i) = 1$; $a_1, a_2, \cdots, a_n$ 是 X 可能发生的结果值; $P_1, P_2, \cdots, P_n$ 是 X 取这些值的概率。仙农(Shannon)根据这些结果把 Z 的信息量定义为:

$$H(X) = - \sum_{i=1}^n P(a_i) \ln P(a_i)$$

这个函数的单位是表示给定事件 = 进位数的位数,简称 *bit* (比特)。我们通常也称 $H(X)$ 为信息熵。从这个最基本的信息熵出发还可进一步引出联合信息熵以及其它一系列的有关的信息量概念。信息量是研究信息论的基础。

【通讯系统】我们谈的通讯系统,通常指数字通讯系统(其它形式可类比)。一个数字通讯系统是由如下几部分组成:一是信源;二是信息编码器,一般说来,将要输出的信息编成由 0,1 组成的二元信息

序列;三是信道编码器,将二元信息序列按一定的规则编成较长的二元序列,叫作码字序列,再经过调制器把码字变成波形传送到收方,经过收方解调器将收到的波形转换成二元数字序列,叫作接收序列,由于讯号在信道传送过程中受到噪声、天电、串音等干扰,使信号失真,接收序列与发送的码字有差异,因此要求编出的码字有一定的纠错能力;四是信道译码器,根据收到的序列、信道编码的法则和信道的特性、纠正接收序列在传送地程中出现的错误,然后经过信源译码器将纠过错的接收序列转换成信息序列。

【代数编码】若假定信息是用 q 个 3、, 符号的序列进行表示的, 则可对这个信息序列分成等长的信息段, 每段由 k 个信息符号组成 $m = (m_0, m_1, \dots, m_{(k-1)})$, 叫作信息段, k 称为信息次数。若按代数的方法将每个信息段变换成较长的几位符号组成的序列 $c = (c_0, c_1, \dots, c_{(n-1)})$ 叫作一个码字。通常 q 取作 2, 符号用 0, 1 两数字表示。在这个条件下, 所给出的代数编码理论主要建立在有限域 F_2 上的几维线性空间的研究上; 进一步也可转换在二元域 F_2 上的一元多项式环 $F_2[X]$ 上, 来探讨同余类环 $R = F_2[X]/X^n - 1$ 。所得到的重要码有循环码。研究循环码的问题, 完全归结到研究同余类环 R 的理想问题。事实上, 每个循环码 V 是由 $X^n - 1$ 的某个因式 $g(X)$ 在 R 中生成的, $V = (g(X))$; 反之 $X^n - 1$ 的每个因式 $g(X)$ 都生成一个循环码。有代表性的循环码有 BCH 码。代数编码的目的就是纠错, 具体说来就是对接收的码系列进行译码时利用编码的校验部分进行校正从而正确还原出信息序列。

控 制 论

【控制论】研究系统的调节与控制的理论。20 世纪 50 年代末开始形成和发展起来的现代控制论为系统的描述、分析综合和设计、预测等问题提供了系统的理论与方法。现在控制论主要有以下几个研究方向:线性控制系统理论、非线性控制理论、分布参数控制系统、最优控制理论等。

【分布参数控制系统】用偏微分方程或用偏微分-积分方程描述的系统。例如描述温度场、弹性振动的系统就是这种系统。它的状态空间是某一函数空间,它的每一瞬时状态是空间的一个函数。分布参数系统的研究的主要内容是系统的辨识、滤波和控制。

【最优控制理论】要求找到一个控制函数。使在它的控制下,系统从一个状态转移到所希望的状态。并且还要求控制的方式是最好的。

设有非线性系统

$$\dot{X} = f(t, X(t), u(t)) \quad (6)$$

选择最佳控制就是在给定的控制区 $U \subset R^r$ 内,选取控制向量 $u(t)$,使(6)的对应的解 $X(t)$ 从给定的初态 $X(0) = X_0$ 在某个时刻 T 转移到给定的终态 $X(T) = X_f$,且

$$\int_0^T \Phi_0(t, X(t), u(t)) dt \quad (7)$$

取到最小(或最大)值。其中 Φ_0 是给定的已知函数。这样的控制 u

(t)称为最佳控制,对应的解 $X(t)$ 称为最佳轨道,称积分(7)为系统的性能指标。

数学史及数学家

【数学史】科学史的一个分支。研究数学发展规律的学科。通过数学史的研究,既可总结数学的过去发展规律,又可预见其发展的未来,还可为现代数学研究提供借鉴的作用。数学史,可以按不同的地域、时代、民族、性质、特点、专题进行分类,因此可有多种数学史,例如可有国别数学史、地区数学史、断代数学史、某一民族数学史等。由于分类的标准不一,其不同种类的数学史,从宏观上看其特性固然无大差别,但从微观上看,则各具特色。

【中国数学史】研究中国数学的发展规律的科学;中国数学史的研究对象是中国历代的数学成果、数学学术活动、数学思想、数学的历史背景以及一切记录等。根据《易·系辞》记载,“上古结绳而治,后世圣人易之以书契”。这说明上古时候,先有结绳记事或记数,然后易之以契刻,三国时代虞翻《易九家义》也说:“事大,大结其绳;事小,小结其绳;结之多少,随物众寡。”不但有很多书籍上有“结绳”、“契刻”的记载,而近代也发现不少原始社会遗留下来的实物,因此可以说《易·系辞》的记载是可以信的。这也说明中国数学史从原始社会就开始了记数的工作。许多出土的原始社会的陶器上,可以发现刻画着很多不同的几何图形和数字符号,有菱形、圆形、

鱼形、矩形、三角形等、还有一、五、七、十、二十、三十等数目字。通过这些实物,说明在原始社会就形成了初步几何图形及数字的概念。因此可以说中国数学起源于原始社会,而中国数学史则也是起源于原始社会。随着时间的推移,到殷商、西周时代,由甲骨、青铜器皿上,可以发现许多数学资料,不但有完整的整数及部分分数记录,还有简单的数字运算。这说明到殷商、西周时代,已积累了很多数学知识,所可惜的是,尚没有发现有关的书籍。

从春秋到西汉末期,在一些典籍中记载着丰富的数学内容。例如,《周易》、《管子》、《墨子》、《考工记》等,都有一些记载。在《周易》里,记载了由阴、阳爻组成的八卦、六十四卦,实际上相当于是重复排列;在《管子》里记载着一些乘、除法例证;在《墨子》里则记载一些逻辑学、力学、光学以及一些几何方面的内容,企图用逻辑的方法论述几何的概念;在《考工记》里,除有一些工件的大小尺寸外,还涉及分数的运算、角度的概念、容量的进位等。这些书籍虽不是数学专著,但却记载了某些数学的片断知识,足证春秋时代的数学水平;1983年,在湖北出土一部《算数书》,根据初步考证,是成书于西汉初期的数学专著。其中有整数的各种运算、分数的运算、比例算法、简单的方程算法以及利息算法等。其内容虽比较零散,但自古以来这是第一部流传至今的数学书籍,由此说明,由于社会的需要,数学工作有了明显的提高。

两汉时期,数学有了长足的进展,达到了较高的学术水平;如《周髀》、《九章算术》都是传世的名著;在《周髀》里,记载有繁难的分数运算、等差级数算法、勾股原理的论述、重差算法以及用矩之道。在《九章算术》里,论述了完整的分数算法、平面图形面积算法、立体图形体积算法、比例算法、开方算法、配分比例算法、均输算法、盈不足算法、“方程”算法、正负数算法、勾股算法以及简单的测

量算法等。此外,在一批汉简中,也记录有相应的数学知识。

到了魏晋南北朝时代,中国数学有了飞跃的发展,三国吴人赵爽为《周髀》作注,魏人刘徽为《九章算术》作注,不仅有许多创造、发明,还为中国传统数学奠定了理论基础,形成具有东方特色的数学理论体系。使得中国传统数学进入第一次高峰;刘宋祖冲之及其子祖暅在圆周率方面取得惊人的成就,并巧妙地求得球的体积。《孙子算经》、《张邱建算经》分别记载了著名的“中国剩余定理”、“百鸡问题”;即使《元嘉历》、《大明历》也有一些数学成就。

在隋唐时代,太史丞王孝通撰写《缉古算经》一部,从而建立一套列三次方程、解三次方程的方法;为了给国子监选定教科学,李淳风奉命校勘十部算经,而隋唐的数学教育逐渐走向制度化。一方面吸收了大批留学生,一方面与朝鲜、日本、印度开拓了学术交流

活动。

宋元时期,中国传统数学形成了第二次高峰;北宋沈括首创垛积术,刘益、贾宪提出解方程的新方法,并创立了二项式系数表,南宋秦九韶给予一次同余式组的一般解法,还提出高次方程数值解的一般方法,杨辉对纵横图提出一些造术方法,元代李冶对天元术即一元代数学给以系统的研究,在此基础上,朱世杰创立了四元术,即是四元代数学,并发挥了垛积术,即高阶等差级数的求和问题。所有这些成果,都遥遥领先于世界。

明代政治腐败,重技巧而轻理论,商业数学逐渐代替了中国传统数学的发展,在计算方面,珠算代替了筹算,并得到了普及。到明末时期,西方数学便乘隙而入,先后传入欧几里得《原本》、克拉维斯《实用算术》、三角学以及对数等知识。到清代,除西方数学继续输入外,清代学人,如梅文鼎家族、汪莱、明安图、李善兰等人,也有许多建树;由于整理、收集四库全书,有一些人如戴震、李潢、李锐、

罗士琳等人,挖掘出一批古算书,从而为复兴古算作出了贡献。

中国传统数学,由上古时代开始,至清末为止,共经历了数千年之久,独自形成了以算为主、形数结合、程序化的特色。到清末之后,就走向中、西数学汇合的道路,从而进入世界数学的行列。

【中国古代数学】中国古代数学是指在秦汉之前的数学。在殷商时代的甲骨上,记载了完整的十进制系统,并有了“十”、“百”、“千”、“万”等数名,而且有最大的数字“三万”还有“五十犬,五十羊,五十豚;三十犬,三十羊,三十豚;二十犬,二十羊,二十豚;十五犬,十五羊,十五豚”。的记载。从这些甲骨文记载来看,殷商时代已有完整的十进制,而且具备了简单的运算法则。

春秋战国时代在书籍里,包含有一些数学知识。如在《周易》里,以阴爻“— —”、阳爻“——”为元素,进行重复排列,从而得到“八卦”、“六十四卦”。实即每取三个元素、每取六个元素的排列数。即 $U_2^3 = 2^3$, $U_2^6 = 2^6$ 。

在《墨子》里,有“经上”、“经下”、“经说上”、“经说下”四篇,通称为《墨经》,包括十数条相当于几何命题的记述,“经上”：“圜，一中同长也”。“经说上”：“圜，规写交也”。其中，“圜”与圆相通，是说“圜”是从一个中心到周围等距离的图形。这相当于圆的定义。后面是说圆是用圆规画出来的。又如，“经上”“穷，或有前不容尺也”。“经说上”：“穷，或不容尺，有穷；莫不容尺，无穷也”。其中，“或”即是域，表示区域或路程，“尺”是度量的单位，这里相当于给出“有穷”、“无穷”的概念。其余尚有平：同高也（高度相等）；方：柱隅四灌也（四边四觔止）等记述。

在《庄子·天下篇》征引了名家惠施的论述：“至大无外，谓之大一。”“至小无内，谓之小一。”其中“大一”、“小一”是惠施提出的

两个名词,并用“至大无外”、“至小无内”分别予以解释,这两句话的意思,相当于现今所说无穷大、无穷小的概念。此外《庄子·天下篇》还引用名家之说:“一尺之棰,日取其半,万世不竭。”这是说,有一尺长的木棰,每日取其前一天剩余的一半,虽取一万年,也取不足一尺。

在《考工记》里,除记载了工件的规格、尺寸外,还涉及一些数学的知识。如用“句”表示锐,用“倨”表示钝,用“倨句”表示角;在《考工记》里,称直角为“倨句中矩”或称作“矩”,并称“半矩”(45°)为“宣”;称“一宣有半”(67°30′)为“櫪”;称“一櫪有半”(101°15′)为“柯”;称“一柯有半”(151°52′5)为“髣折”;也称“一矩有半”(135°)为“髣折”;还用圆心角的大小描述弓形的曲率,如“天子之弓,合九而成规”;“诸侯之弓,合七而成规”;“大夫之弓,合五而成规”;“士之弓,合三而成规”。也记载了一些分数的简单运算。此外,在《管子》、《荀子》、《法经》里也有一些数学知识的零星记载。

在秦始皇统一中国之前,虽未发现有数学专著问世,但从这些记载来看,在某些方面,中国古代数学已达到成熟的境界。

【秦汉数学】指秦汉时代中国传统数学的成就。在《汉书·艺文志》里称:“《许商算术》二十六卷”。“《杜忠算术》一十六卷”可惜这两部专著早已散失,现传本只有《周髀》、《九章算术》以及新出土的《算数书》等著作。

由《周髀》表面看来,记载有很多西周周公、商高、陈子、荣方等人的对话,也记载了吕不韦的论说,又因盖天学说是西汉初期天文学家创建的一种宇宙学说,所以初步考证,《周髀》可能成书于西汉时期。它虽是一部盖天学说的天文书籍,但其中却记载了不少数学知识;例如,由于天文历法的需要,书中有相当繁杂的分数运算;书

中还记载了“七衡六间”、二十四节气晷影长度的计算问题,是涉及到等差级数的问题。在计算二十四节气晷影长度时,也是按等差级数算法计算的。在《周髀》里,还记载了六条“用矩之道”,即“平矩以正绳,偃矩以望高,覆矩以测深,卧矩以知远,环矩以为圆,合矩以为方”。其前四句是说利用相似勾股形性质,进行的一次测量问题,后两句则是说用矩绘制圆、方的问题。《周髀》还记载说,“若求斜至日者,以日下为勾,日高为股。勾、股各自乘,并而开方除之,得斜至日”。即:

$$\text{斜至日(弦)} = \sqrt{\text{勾}^2 + \text{股}^2}$$

这就是所谓“勾股原理”,可是在《周髀》里却缺乏严格的证明。

《九章算术》采用问题集的形式,全书共 246 问,202 术,厘为九类,计方田、粟米、衰分、少广、商功、均输、盈不足、方程、勾股九章;方田章,38 问 21 术,主要论述各种平面图形如方田、直田、圭田、邪田、箕田、圆田、宛田、弧田、环田的面积算法,除宛田、弧田是近似算法外,其余都是正确算法。方田章还论及分数的通分、约分、扩分、加、减、乘、除、比较大小、求平均数的各种运算;粟米章,46 问 33 术,主要论述了 20 种粮食、饭食的兑换比率表以及四项比例算法。四项比例算法,当时称之为“今有术”,即“以所有数乘所求率为实,以所有率为法,实如法而一”。即:

所求数 = (所有数 × 所求率) ÷ 所有率。在章末并论述了四项比例算法的特殊情形,即“经率术”、“经术”、“其率术”、“反其率术”;衰分章,20 问 22 术,主要论述配分比例算法,有的问题是按等差级数计算,也有的问题是按等比级数计算的;少广章,24 问 16 术,主要论述已知长方形宽边及其面积,推求其长边的算法,其特殊情形即是已知正方形面积求其一边之长,实即开平方算法,推广此法即是开立方算法,还列有根据圆面积、球体积求其直径的算法;商

功章,28问24术,主要论述立体图形的体积算法;均输章,28问28术,主要论述比较复杂的配分比例算法;盈不足章,20问17术,主要论述中国特有的盈亏算法,按现今说法,即是“双设法”或“弦位法”;方程章,18问19术,主要论述“方程”的解法,所谓“方程”,即今之线性方程组,其解法即相当于加减消元法,由于需要,还论及正、负数的运算法则;勾股章,24问22术,主要论述有关勾股问题的解法,并论及简单的勾股测量问题。

1983年,在湖北出土一部数学著作,一般称之为《算数书》,其中包括有60多个小标题,标题之下列有算题,题下有答,答下有术,《算数书》也是一部问题集。在《算数书·少广》有一问与《九章算术·少广》之第一问极为相似,而《算数书·钱息》有一问与《九章算术·衰分》第二十问完全一致。由这两对问题可看出,《九章算术》是在《算数书》的基础上修改而成的,而《算数书》乃是《九章算术》的母本之一。

在《九章算术》成书之前,已有《许商算术》、《杜忠算术》。而《九章算术》必然参考过这些著作,虽不便认为“许商、杜忠《算术》即是《九章》”,但认为《九章算术》是由《算数书》、《许商算术》、《杜忠算术》演变而成的,并非毫无根据。

《九章算术》既记录了秦汉时代的数学成就,也反映了秦汉时期的各种章典制度和社会情况。可见,秦汉数学的学术是有一定水平的。

【魏晋南北朝数学】从三国时代到隋文帝统一中国之前的数学各项成就辉煌,是中国数学的第一个高峰期。

三国时代吴人赵爽,其字君卿,身世不详。曾为《周髀》作注,在“注”中,补绘了“日高图”及“七衡图”;赵爽撰写了“勾股圆方图”

注、“日高图”注、“七衡图”注；在“勾股圆方图”注中，他用 500 余字，论证了勾股原理，还论证了有关勾、股、弦的 20 多条命题，在“日高图”注中，他对所谓“日高术”即重差术，也给予几何证明，在“七衡图”注中，他对盖天学说的理论作了说明，所有这些，对后世研讨《周髀》者都有很大裨益。例如“勾股圆方图”注说：“勾、股各自乘，并之为弦实。开方除之，即弦。案弦图又可以勾、股相乘为朱实二，倍之为朱实四。以勾、股之差自相乘为中黄实。加差实一，亦成弦实”。其前一句是重复论述勾股原理，而后一句则是以面积概念证明了勾股原理。设勾、股、弦分别为 a 、 b 、 c

勾股原理即 $\sqrt{a^2+b^2}=c$ 其证明过程为：

勾、股相乘为朱实二， ab ；

倍之为朱实四， $2ab$ ；

以勾、股之差自相乘为中黄实， $(b-a)^2$ 。

加差实一，亦成弦实。 $2ab+(b-a)^2=c^2$ 。这是中国对勾股原理第一次严密的证明。

略晚于赵爽的是刘徽，刘徽是魏、晋期间杰出布衣数学家，其身世不详，初步考证，当是现今山东人，曾为《九章算术》作过注解；刘徽《九章算术注》中，在数学理论上、方法上、技巧上、程序上多所建树和发明，为

中国传统数学奠定了理论基础,为中国传统数学形成了独特的理论体系。

刘徽利用出入相补原理,即利用图形的分、合、移、补的方法证明平面图形如圭田、邪田、箕田、圆田、宛田、弧田、环田的面积算法,还用以证明勾股原理以及整数勾股弦的一般公式。

设勾、股、弦分别为 a 、 b 、 c , 整数勾弦一般公式为: $a : b : c = [m^2 - 1/2(m^2 + n^2)] : mn : 1/2(m^2 + n^2)$ 至于直线型柱体如城、垣、堤、沟、堑、渠的体积或容积算法,也以出入相补原理予以证明;刘徽把出入相补原理作为一条普遍的原理,应用于几乎全部的几何理论之中。刘徽又利用三种基本几何体,即堑堵、阳马、鳖臑,他先证明三种基本几何体的体积算法,再把直线型非柱体的立体分割为三种基本几何体,然后证明其体积算法。可见,刘徽把三种基本几何体作为多面体体积算法的理论关键。对于圆型平面图形如圆田、弧田的面积算法,刘徽则是利用极限观念和出入相补原理证明其面积算法,并求得圆周率较精密之值分别为

$$\pi = 157/50 = 3.14,$$

$$\pi = 3927/1250 = 3.1416。$$

对于环田,乃是将环田之周拉直,使变为等积的梯形,再用出入相补原理证明其面积算法。刘徽在推证圆型立体时,采用了截割原理,即是在圆柱、圆锥、圆台上,各作一外切方柱、方锥、方台,以圆型立体与其外切方型立体体积之比为 $\pi : 4$,从方型立体可证得圆型立体的体积算法。刘徽在正方体上,作相互垂直的两圆柱,称两圆柱的公共部分为“牟合方盖”,在“牟合方盖”里作内切球,刘徽深知“牟合方盖”与其内切球体的体积之比为 $4 : \pi$,他虽未求得球体体积算法,但他提出其体积之比,在算法理论上是十分珍贵的。刘徽在推证勾股原理的基础上,还用面积理论推证有关勾、股的各种线

段的求法,还用相似勾股形性质推证“勾股容方”、“勾股容圆”的求法,也证明了简单的测量算法。刘徽对数学名词改变了“约定俗成”的惯例,对一些名词给予明确的定义,如把“率”定义为“凡数相与者,谓之率”。“等除法实,相与率也”。即是说,凡数与数之比,称之为“率”,约简两数之比,则称为“相与率”。并把“率”概念几乎应用到所有算法之中,作为各种算法的主线。给正负数所下定义为:“今两算得失相反,要令正负以名之”。还给“方程”下了正确的定义,既给出“方程”有确切解的条件,又给出“方程”的同解理论,并创造性的给出“方程”的新解法。此外,刘徽还对幂、齐同通、列衰、开平方、开立方、鳖臑、阳马、堑堵、勾、股、弦等数学名词都给出正确的定义。刘徽的这些定义,不但没有含混不清之词,也没有循环定义之举,都合于逻辑,因而成为演绎论证的理论依据。刘徽在推理演绎与证明方法上,既有归纳,也在演绎;既有综合法,也有分析法,还有反证法;在逻辑方面是十分丰富的。于是在中国传统数学理论的发展上,形成了第一次高峰。

西汉时期,主张盖天学说的天文学家创造了测量日高、远的方法,称之为重差术,到刘徽时代,几乎失传。刘徽乃潜心研究测量原理,使重差术加以发展,并由两次测望推广至三次、四次测望,编撰九道测量问题,缀于《九章算术》之终。唐代,李淳风为国子监审定数学教材,使另行单本,因其第一问为“今有望海岛”,故称之为《海岛算经》。可见刘徽在测量理论上的成就是卓著的。

生活在南北朝宋、齐时代的著名数学家祖冲之及其子祖暅,对数学有很出色的贡献,例如他们父子曾作《缀术》一书,可惜早已失传。保留至今的,只有祖冲之所创造的圆周率,即介于盈肭二限:

$$3.1415926 < \pi < 3.1415927,$$

密率: $\pi = 355/113$, 约率: $\pi = 22/7$ 。还有其子祖暅解决了刘徽所提

出的“牟合方盖”体积问题,从而得到球体积的巧妙证法。

在南北朝成书的《孙子算经》三卷,是一部数学普及著作,其卷下除有“鸡兔同笼”问题外,还有著名的“物不知数”题,即“今有物,不知其数。三三数之,剩二;五五数之,剩三;七七数之,剩二。问物几何”。也即现今之一次同余式组问题;设所求之物数为 N ,则有 $N \equiv 2 \pmod{3}$, $N \equiv 3 \pmod{5}$, $N \equiv 2 \pmod{7}$,《孙子算经》给出 N 的最小解为 $N=23$,西方学者一直称这问题为“中国剩余定理”,我们则称之为“孙子定理”。

《张邱建算经》也是一部数学普及著作,书中涉及最大公约数和最小公倍数的正确求法,还有七道问题涉及等差级数及其解法,有的固是继承《九章算术》的成果,但等差级数更多的内容则是创新;《张邱建算经》卷下,有一“百鸡问题”,是著名的不定方程问题。即“今有鸡翁一,值钱五;鸡母一,值钱三;鸡雏三,值钱一。凡百钱买鸡百只,问鸡翁、母、雏各几何”。书中共给出四组答案。虽然《九章算术》已有不定方程的记载,但一题数答却始自《张邱建算经》。

【隋唐数学】唐代初期,太史丞、历算家王孝通编撰《缉古算经》一书,全书由二十问集成。其第一问是算术问题,已知日月合朔时刻及夜半时日所在赤道经度,推求夜半时月所在赤道经度。第二至十四问,是土木建筑问题;需用三次方程求解。第十五至二十问,是勾股问题;需用三次方程或双二次方程求解。《缉古算经》虽未给出解三次方程的具体计算步骤,但王孝通必然是用带从开立方法来解的,这显然是一项辉煌成就;他还用自注形式给出方程各项系数的组成方法,这一作法,也是难能可贵的。唐代沿袭隋代的办学制度,建立国子监,设立数学专业,天算家李淳风奉命选定并校注数学专业教材,因此选定《周髀算经》、《九章算术》、《海岛算经》、《孙子算

经》、《五曹算经》、《张邱建算经》、《五经算术》、《缀术》、《缉古算经》、《夏侯阳算经》等 10 部算经,作为教科书。在唐代,还规定了学制,也规定了考试制度和考试成绩即是及格、不及格的标准。自隋代建立国子寺之后,到唐代才逐渐形成一定的教育制度,不但影响隋唐以后的各代,这种教育制度也直接传入朝鲜、日本等国。随着教育制度的传播,一些中国古代数学著作也传入朝鲜、日本;对朝鲜、日本的数学发展,产生了一定的影响。随着佛教及其经典的传入,印度的一些数学知识也传入中国,例如印度数码随着《开元占经·九执历》传入,由于没有把印度数码写法刊刻出来,以致印度数码当时没有在中国流传下来;又如,在《九执历》里记载有从 0 度到 90 度每隔 3.75 度的正弦线值表(其半径取为 3438),由于此表没有引起当时中国历算家的重视,此表在中国也就没有起到应有的作用和影响。但是,中国数学是否传播到印度,则是一个需要深入研究的课题。

【宋元数学】中国传统数学的发展中,宋元数学达到了较高水平,并遥遥领先于世界,成为中国传统数学的第二次高峰。北宋数学家刘益,著有《议古根源》,现已失传。只能从南宋杨辉的著作中看到一些端倪。《议古根源》打破了古代对方程系数的一些限制,提出新解法,即正负开方术,经后世数学家的发展,即形成著称的增乘开方法。北宋另一数学家贾宪,著有《算法斡古集》、《黄帝九章算法细草》二书,此二书均已失传;只能从杨辉的著作中,了解其中的梗概。对于方程解法,贾宪提出“立成释锁平方法”、“增乘开平方法”、“立成释锁立方法”、“增乘(开立)方法”四法,其中释锁平方、释锁立方与《九章算术》相似,而增乘开方法却比旧法简明得多。另外贾宪还提出“开方作法本源”图,现称为“贾宪三角”,实即二项式展开

系数表,这是中国传统数学的优秀成果之一,也是方程理论的重要内容。沈括,字存中,号梦溪,钱塘(今杭州)人,著有《梦溪笔谈》、《梦溪补笔谈》,其中给出隙积术、会圆术、棋局都数,这三项都是中国传统数学上的杰出成就。隙积术,是推求累棋、层坛、积罍的求和公式,也即垛积术的开端;会圆术即是由弦、矢求弧的算法;棋局都数是计算棋局的总数。如清代《畴人传》称:“隙积、会圆二术,补《九章》所未及”。南宋数学家秦九韶,著有《数书九章》,其中集有 81 问,分为 18 卷,厘为九类,每类九题。《数书九章》对数学的主要贡献计有:“大衍总数术”,即一次同余式组的一般解法;书中给出了三个乃至八个模数的问题;“正负开方术”,即高次方程的数值解法;书中也给出了 8 次乃至 10 次方程的问题;而这两项代表了中国传统数学的最高成就,也标志着当时世界数学的最高水平。另外,秦九韶还改进了《九章算术》的“方程”术,即改进线性方程组解法为“互乘对减”消元法。他也独立地由三角形三边求得三角形的面积,即所谓“三斜求积”术,这与古希腊海伦公式是等价的。南宋数学家杨辉,著有《详解九章算法》12 卷、《日用算法》2 卷、《乘除通变本末》3 卷、《田亩比类乘除捷法》2 卷、《续古摘奇算法》2 卷,共计 5 种算书,其后三种一般称为《杨辉算法》。在《详解九章算法》中,选《九章算术》80 问作为典范,进行详解,并增添“图草”、“乘除算法”、“纂类”三卷;在“纂类”里,引述了刘益的“正负开方术”贾宪的“增乘开方法”与“开方作法本源”,都是十分珍贵的数学史料。在《日用算法》里,把数学内容编成韵文,便于记诵,成为中国民间数学的特色。《乘除通变本末》是论述乘除算法、加减算法、求一算法诸术。《田亩比类乘除捷法》可看作是《详解九章算法》的延续。而《续古摘奇算法》则是搜集“诸家算法奇题及旧刊遗忘之文”,其上卷主要论述“纵横图”,下卷主要论说《海岛算经》。杨辉对一些纵横

图给出了造术方法,为古代数学著作之所缺;杨辉用面积概念论证《海岛算经》的“日高”、“日远”公式,成为后人研究《海岛算经》造术的依据之一。中国古代,很早就有以算筹表示方程的方法,但是,关于建立方程的一般方法则是当务之急;由于方程理论的高度发展,北方一些数学家开创了“天元术”,“天元术”就是用文字表示未知数以建立一元方程的方法。李冶虽不是天元术的首创学者,但是,他的著作《测圆海镜》12卷、《益古演段》3卷,都是天元术的专门著作,对于天元术的发展和推广却起到不可忽视的作用。在《测圆海镜》里,李冶以“太”字表示常数项,或以“元”字表示一次项,用“太”不用“元”,用“元”不用“太”,这就是他改进原来天元术建立方程的方法;他还把建立方程方法的步骤制度化,首先是“立天元一为某某”,相当于“设 x 为某某”;然后依题意列出一多项式,称为“天元筹式”;再依据题意列出另一多项式,称为“同数”;最后“天元筹式”与“同数”两式相消,即是相减,即得一元高次方程。用这种方法建立方程,不但改变古代用几何法建立方程的繁难,而且所建方程之次数也不受限制。对于高次方程,则采用“增乘开方法”以求其解。李冶在《测圆海镜》里,先提出 692 条几何命题,作为全书的理论依据,又给出得自“洞渊九容之说”,连同《九章算术》的“勾股容圆”,共计十种容圆公式,作为此书 170 问的基础。在《益古演段》中,一方面用天元术来解所列 64 问几何题,一方面对比蒋周《益古集》的旧解法,从而看出李冶在天元术方面的贡献。随着天元术的发展,北方数学家又创立了二元术、三元术,即是二元代数学、三元代数学,基于此,元代中期朱世杰创立了四元术,即四元代数学。朱世杰著有《四元玉鉴》、《算学启蒙》二书,书中用“天”、“地”、“人”、“物”四字分别表示四种未知数,根据题意,可以列出四个未知数的高次方程组,而四元术的精华是“消元”,通过“消元”,使四元方程化为

三元方程,再使三元方程化为二元方程,由二元方程化为一元高次方程,然后将用“增乘开方法”以求其解。朱世杰的“消元”法,是中国传统数学的一项杰出成就,不但领先于西方四百余年,用“消元”法解多元高次方程组的精神是值得称道的。在《四元玉鉴》及《算学启蒙》里,还在另项杰出的成就,就是在沈括隙积术的基础上,系统地发展了垛积术及其相关的内插法;也就是开创性地对高阶等差级数作了系统的研究,取得了辉煌的成绩,并领先于西方三四百年之久。在宋元期间,中国传统数学取得非常丰富的辉煌成果,可以说中国传统数学达到了全盛时期。

【明清数学】在明代,由于政治腐败、不注重理论、废除数学考试等原因,使得中国传统数学逐渐走向缓慢发展的道路。在当时,社会上流传着数十种数学著作,这些著作大都是反映日常应用数学和商业贸易数学,其内容比较浅显,文字通俗易懂,还编撰了各种歌诀,便于记忆,其算法也容易掌握,因此促使这种数学广为流传,而古代的《九章》、天元术、四元术等则几乎成了绝学。当时的代表作计有吴敬《九章算法比类大全》、程大位《直指算法统宗》、王文素《新集通证古今算学宝鉴》三书,这些书籍虽然采用《九章算术》各章的名目,但其内容却多是日用或商业的数学知识,还有一大部分内容属于珠算的技巧和算法。于是珠算之盛行与普及代替了古代的筹算法,以筹算为基础的传统数学也渐次匿迹、消失。

在中国传统数学走向低谷的时候,西方教士利玛窦首先来华,一方面传播教义,一方面传入西方的科学以及数学知识。万历三十五年(1607),他便与徐光启合译了欧几里得的名著,称为《几何原本》前6卷。继利玛窦之后,有大批西方教士接踵来华,有的在政府部门里工作,有的在教堂传播教义,他们编撰、翻译了不少西方科

学著作,仅就数学而论,所编、译的著作多属于初等数学范畴,因而对中国传统数学产生了一定的影响。

在中算濒于失传,西算刚刚输入之际,中国学者的态度极不一致,有人以为“薄古法为不足观”,有人以为“斥西人为异学”,而清初数学家梅文鼎则主张:“法有可采何论东西,理所当明何分新旧”。于是他融会贯通中西之学,自成一家之言。其著作颇丰,粗略统计,约 88 种之多,达 200 余卷。主要著作有:《筹算》、《笔算》、《度算释例》、《平三角举要》、《弧三角举要》、《方程论》、《几何补编》、《环中黍尺》、《堑堵测量》、《勾股举隅》以及《方圆幂积》等,几乎遍及初等数学的各个分支。

清朝第二代皇帝康熙。自幼喜爱科学,尤其酷爱天文与数学,并召法教士张诚、白晋及比利时教士南怀仁入宫讲解西方科学与数学,还接受陈厚耀的建议,开设“蒙养斋”,任梅文鼎之孙梅穀成为蒙养斋编修官,编修《律历渊源》、《数理精蕴》等书。《数理精蕴》共五 53 卷,其内容除有当时已传入的一些西方数学知识外,还有中国数学家如梅穀成的某些研究成果。《数理精蕴》不但是按算术、几何、三角、代数学学科分类的著作,也相当于是一部中国初等数学的百科全书,又因此书有“康熙御制”之名,故流传较广,其影响亦大。

受西方教士的影响,中国学者多所著述。如年希尧经常与意大利教士郎世宁共同切磋工程图学的技艺,领悟西方的透视学的精神,经多年钻研写成《视学》一书;《视学》不仅是中国图学的最早著作,也是世界上较早的图学专著,在国际上也产生了一定的影响。又如法教士杜德美曾将西方三个幂级数展开式传入中国,称为“杜氏三术”,但他只传入三个公式,却未明确其理;蒙族数学家明安图以多年精力,用初等数学方法证明了“杜氏三术”,并给出六个展开

式及其证明,共计“九术”。这“九术”对后世有一定影响,引起了如汪莱、董祐诚、项明达等人的关注,从而得到许多新的研究成果。

乾隆年间,开设四库全书馆,为了编撰《四库全书》,一方面辑录《永乐大典》所保存之书,一方面征集私家所藏之书;因而发现并挖掘出多部濒于失传的中国古代传统数学著作,为了探讨、校订古代的经籍,形成了所谓“乾嘉学派”,他们用分析、归纳的方法,对这些古籍进行研究、校勘、注解,这对于恢复古籍起到了积极作用。为了复兴古算,对于所挖掘出来的古算书进行了整理、校勘、注释、研究,其中做出贡献的,计有戴震、李潢、焦循、汪莱、李锐、罗士琳等人。

到清代后期,西方数学第二次输入中国,中国数学研究曾出现一时的活跃,畴人辈出,著述如林,如在幂级数、尖锥术、组合数学、不定分析等方面,也获得不少可喜成果,虽不及同期西方数学的水平,但有些成果却是中国学者独立创造出来的。李善兰在接触西方数学之前,受到垛积术及传统极限观念的启发,创造出一种相当于微积分的方法,即是尖锥术。他在《方圆阐幽》一书中,以十个“当知”也即十个命题,给尖锥术建立了理论的基础,并列举了一些例证加以核算。李善兰还对垛积术进行了深入研究,他借助并拓展了贾宪三角,别树一帜,自成系统,在《垛积比类》中创造出数十套垛积公式,成为组合数学的光辉成果。

鸦片战争之后不久,李善兰、华蘅芳分别翻译了不少西方古典高等数学著作,从此,西方变量数学开始传入中国,由于文字关系,中国学者接触、研究西方数学还有很大的局限性,于是,清政府为形势所迫,不得不设立同文馆,培养专门人才,并选派留学生出国深造。等到大批留学生回国之后,便兴办学堂,讲解洋课程,培养出许多数学人才,从而使得中国数学的研究与数学的教育,逐渐走向

世界数学的发展道路。

【外国数学史】中国以外的国家或地区的数学史。由于中国以外的国家或地区,其数学的发展并不一致,尤其在古代,各国之间或各地区之间,也未必有其必然的联系,很难统一在一起,直到近代(17—19 世纪)各个国家或地区数学的发展才汇合在一起,成为世界数学。外国数学史,在古代实际上是指各个地区的数学史,例如古巴比伦数学、古埃及数学、古希腊数学、古印度数学、阿拉伯数学等;在中世纪,是指欧洲数学史;在近代,才是世界数学史。由于中国数学有悠久的发展史,经历了数千年之久,而且具有很突出的特色,与任何一个国家或地区的发展,极不相似,所以把中国数学史单独列出很有必要,也有充分理论根据。相应地也把外国数学史单列一项。在古代,亚洲底格里斯河与幼发拉底河之间的地带,是人类文明发源地之一,公元前 19 世纪,苏美尔和阿卡德民族在这里建立了巴比伦王国。19 世纪,在美索不达米亚出土约 50 万块刻有楔形文字的泥板,经考证,这些泥板有的是公元前 20 世纪的遗物,有的是公元前 6 世纪的遗物。这些楔形文字中也包括巴比伦人在数学上的一些成就。由于古巴比伦对奴隶的剥削日趋严酷,农奴生活濒于绝境,于公元前 6 世纪,巴比伦王国覆灭,合并于波斯帝国,而巴比伦数学也告结束。

大约公元前 3000 年左右,在尼罗河一带,形成了古埃及王国。由于埃及人长期与大自然作斗争,逐渐掌握了一些科学、技术知识;又因需要以物易物、丈量土地、建筑房屋及坟墓,也积累了一些数学知识;为了传递信息,古埃及人也创造了一种像形文字,一般称为僧侣文。根据考证,尼罗河每年定期泛滥,泛滥之后,需要重新丈量被淹没的土地,因而长期以来,便由丈量土地的知识逐渐发展

成为所谓几何学。要了解古埃及的某些情况,只能通过“莫斯科纸草书”、“阿默斯纸草书”这两卷纸草书进行探讨。由于宗教的改革,古代埃及统治集团的内部斗争愈加剧烈,外部则经常受到欺凌,于公元前6世纪前后,被波斯吞并,成为一个省,而古埃及的文化也随之逐渐消失。

古代希腊人,为人类创造了历史上的文明,尤其对西方的文化有巨大的影响。古希腊文明可以追溯到公元前29世纪,一直延续到公元6世纪。古希腊的数学发展是由学派组成的,例如,最早是以泰勒斯为代表的爱奥尼亚学派。在爱奥尼亚学派之后,相继而起的是毕达哥拉斯学派,在数学方面,研究了一些初等数论的问题,并以发现勾股定理(西方称为毕达哥拉斯定理)驰名于世。与毕达哥拉斯齐名的学派,是芝诺为代表的悖论学派,悖论学派创立了一些悖论,给学术界造成了极大的震动。原子论学派,主张宇宙间的物质都由不可分割的元素组成。与悖论学派差不多同时,雅典出现了诡辩学派,在数学方面。他们提出了三大几何问题,即“化圆为方”、“倍立方”、“三等分角”。在雅典相继而起的是柏拉图学派,柏拉图是古希腊的著名哲学家,他注重数学,并十分推崇几何,认为几何可以培养思维能力。该学派培养出不少优秀学生,亚里士多德就是他的学生之一。在雅典以亚里士多德为首创办了吕园学派。吕园学派的贡献在于创立了逻辑学,因而为欧几里得的《原本》铺平了发展道路。公元前4世纪,亚历山大帝国瓜分为三个国家,最大的是托勒密王朝。托勒密王在亚历山大城建立了最大的图书馆,从而使得亚历山大城变为希腊文化的中心;但是,到公元5—6世纪,由于东罗马的入侵,希腊文化的发展即告终结,而保留下来的希腊文化遗产,为欧洲的文化提供了丰富的营养。

古印度也是古代文明国家之一,印度数学大约产生于公元前

4 世纪,当时是一种十进非位值制系统,经过千年的变迁,到公元 6 世纪,才形成印度数码,8 世纪传入阿拉伯,13 世纪输入欧洲,逐渐演变成现今所谓印度-阿拉伯数码。19 世纪出土了“巴克沙利手稿”,经考证,记载了印度 4、5 世纪的数学知识,其中论述了“反演法”及其例证。古印度人还以“库塔卡”来解某些不定方程;还改变了希腊人的“全弦”为“半弦”,即今之“正弦”线。

阿拉伯数学是指 8—15 世纪伊斯兰教国家所建立的数学。其代表人物之一,是阿尔·花拉子米,他首先提出所谓“代数学”一词,他对二次方程作了系统的研究。另一代表人物是 13 世纪的纳速拉丁,他首先从天文学里把三角分割出来,使成为一门独立的学科。阿拉伯人曾把一大批希腊、印度的著作翻译成阿拉伯文,使得这些濒于灭亡的著作获得新生,从而传入欧洲,使欧洲数学一跃而起。

中世纪数学是指以罗马为中心的数学发展概况。由于十字军东征,欧洲人从阿拉伯获得大批希腊著作以及阿拉伯文译本,或阿拉伯人的著作,12 世纪便进入大翻译时期;从而了解到希腊及阿拉伯在数学上的贡献。13 世纪初期,由一些教会学校转变为一些大学,这些大学成为后世数学发展的基地。欧洲经过了文艺复兴的洗礼,数学首先发展了起来,并为后世的数学发展做了准备。

【古巴比伦数学】底格里斯河与幼发拉底河流域是世界文明的发祥地之一。大约公元 5000 年前,苏美尔人、阿卡德人就先后活跃在这里。约公元前 1894 年,由阿摩利人建立了古巴比伦第一王朝(约公元前 1595 年灭亡)起,此后一些王朝也多冠以巴比伦之名,因而传统上称这一地区的古代各民族为巴比伦人。19 世纪初,这里发掘出大约 50 万块刻着楔形文字的粘土泥板,其中约有 300 块为载有

各种数表和一批数学问题的数学泥板,关于巴比伦数学的知识,即来源于此。这些数学泥板可分为三组:第一组约在公元前 2100 年苏美尔文化末期;第二组属于古巴比伦第一王朝时期,数量很大;第三组属于公元前 600—前 300 年,内容也较丰富。

巴比伦人对数学的第一个重大贡献是发明了 60 进位记制数法,并一直影响到今天的角度与时间计量。由于表示空位的记号(零号)直到大约公元前 3—前 2 世纪才出现,早期泥板中的数字表示常常发生混乱。他们有完整的加减乘除算法,并使用了 60 进制的小数。为了便于计算,编制了大量数表,如乘法表,平方表、立方表、平方根表、立方根表、倒数表,以及计算复利的指数表等。他们还发明了表示减与乘的记号。巴比伦数学的最高水平体现在代数方面。尽管没有一般的代数符号,解法也仅限于给出每一步的具体计算,但从其所解决的大量问题可以看出,巴比伦人已经有所了解方程的明确概念和方法,并能解决一些方程问题,如,

(1) $aX=b$; (2) $X^2=a$; (3) $X^3=a$; (4) $X^2 \pm aX=b$; (5) $X^2(X \pm 1)=a$; (6) $X \pm y=a, Xy=b$; (7) $X \pm y=a, X^2+y^2=b$; 以及多达 10 个未知数、10 个方程的线性方程组。当然,所有的解都只取正值。他们分别对 $n=9$ 和 $n=10$ 求出了下列级数的和:

$$\sum_{k=1}^n 2^k = 2^n + 2^n - 1$$

$$\sum_{k=1}^n k^2 = \left(\frac{1}{3} + \frac{2}{3}n\right) \cdot \sum_{k=1}^n k$$

巴比伦人的几何水平已很高,涉及的问题有:由平行线引出的比例,勾股定理,三角形和梯形的面积,圆的面积与周长(取 $\pi=3$),棱柱和圆柱的体积。对这些内容他们都得到了正确结果,有一块泥板上还出现了取 $\pi=3\frac{1}{8}$ 较好近似值。然而他们给出的正四棱台体

积公式却是错误的,或者说是有很大误差的。

巴比伦数学的基本特点是它的“算术-代数”性质,几何学是从属于算术与代数。数学还主要是一门实用技术,没有成为科学,但他们的绝大部分数学成果都是在公元前 1600 年以前取得的,并且和埃及数学一起成为希腊数学的先导。

【埃及古代数学】以金字塔闻名于世的埃及,很早就在数学上取得了引人注目的成就。我们了解埃及古代数学的主要依据,是大约公元前 1850—前 1650 年间的两份纸草书:莫斯科纸草书与阿默斯纸草书。前者因其收藏于莫斯科美术博物馆而得名,后者则得名于原件的书写者,人们还认为,阿默斯纸草书是一部更为古老的数学著作的抄写本。

埃及人建立了以 10 为基数的数学符号,由于完全没有位值的概念,他们需要为 10 的每一幂次设计一个新的符号。在埃及古老的像形文字中,1、10、100、1000 分别被记作 1、 $\cap$ 、 $\odot$ (或 $\ominus$)、 ⌘ (或 ⊖),一般数字是将 10 的各个幂次按加法原则累积而成,例如 2345 记作

$\text{⌘} \text{⌘} \odot \odot \odot \cap \cap \cap \cap ! ! !$

这就使得数的表示与各种运算都十分烦琐。古代埃及人的加减乘除运算是完整而正确的。他们发明了形如 $\frac{1}{n}$ (n 是自然数) 的单位分数,一般分数则用单位分数的和表示,为此编制了专门的分解表。除法是根据乘法的逆运算原理或直接用分数分解表实现的。

埃及人解决了一些在今天看来属于代数学范围的问题,有的相当于一元一次或二次方程,还有的问题涉及等差或等比数列中的几项。但是,所有这些问题都是用加、减、乘、除、比例等典型的算

术方法求解的,至多涉及开平方运算,而所谓代数问题的解法是以比例算法为核心的单假设法,从中难以看出代数学的思想与方法。

金字塔的建造说明埃及人有着丰富的几何知识。实际上,由于古埃及的主要经济部门是农业,尼罗河水泛滥后土地的重新测量,兴修水利、营建仓廩等都引起了对几何知识的大量需要。埃及人正确地计算了正方形、长方形、直角三角形,直角梯形以及能分割成这些形状的图形的面积,可能也已经知道了勾股定理的某些特殊情形。如果记面积为 S , 直径 d , 则他们的圆面积公式相当于 $S = (\frac{8}{9}d)^2$, 由此可以得到圆周率 π 的一个很好的近似值 3.16049。他们也正确地计算了立方体、长方体、棱柱、圆柱的体积,特别是以对称的形式给出了正四棱台的体积公式

$$V = \frac{1}{3}n(a^2 + ab + b^2)$$

被后人一语双关地称作最伟大的埃及金字塔,他们的圆台、半球体积公式则是相当粗糙的。

埃及数学的特征是它的“算术-几何”性质。他们创造了十进制(但非位值制)记数系统,发展了完整的整数及分数运算,并且运用比例原理解决了某些代数问题。他们的几何知识达到了很高水平。所有这些都对希腊数学产生了直接而深远的影响。

【古希腊数学】主要由学派组成的,较早的学派是以泰勒斯为代表的爱奥尼亚学派,其最主要的贡献是把演绎证明引入到数学中,例如论证了(1)直径平分其圆周;(2)半圆周角是直角;(3)等腰三角形之底角相等;(4)两直线相交其对顶角相等;(5)若三角形的一边和两角为已知,则此三角形为确定。从而便进入了数学发展中所谓初等数学时期。

在爱奥尼亚学派之后,相继兴起的是毕达哥拉斯学派,这一学派带有一定的神秘主义色彩,他们信奉“万物皆数”的信条,认为宇宙间一切事物、现象都依附于某些数值之间的相互关系;他们的研究范围属于比例论、初等数论、几何代数等范围,例如他们研究了“三角形数”、“正方形数”、“多角形数”、“亲和数”、“完全数”以及“整勾股数”等;并以发现勾股定理而著称于世(西方称勾股定理为毕达哥拉斯定理);他们还发现了正方形的边与其对角线不能公度的问题。

与毕达哥拉斯学派齐名的学派,是以芝诺为代表的悖论学派,悖论学派创立了一些悖论,给学术界造成极大震动;他们所提出的悖论分别是:(1)二分说,即一物体从 A 到 B,永远不能到达;(2)追龟说,即阿基里斯永远追不上乌龟;(3)飞箭静止说,即飞箭在任一时刻都是在静止的位置上,所谓运动即是静止的总和;(4)运动场说,即一段时间和其一半相等。他们还研究了物质的连续性、无限性等,从而为原子学派奠定了思想基础。

原子论学派的代表人物是德漠克利特,他是古希腊的数学家、哲学家,他认为任何线段、面积、体积都是由有限个不可再分割的原子所构成,并成功地把这种观点运用于计算之中,所谓计算长度、面积、体积就等于将这些原子集合在一起。

与悖论学派差不多同时,在雅典出现了诡辩学派,这学派主要的研究是用数学来解释宇宙的现象,在数学方面,他们提出三大几何问题,即(1)作一正方形,使与已知圆面积相等,简称为“化圆为方”;(2)作正方体之一边,使其体积等于已知正方体体积之二倍,简称为“倍立方”;(3)三等分任意一角,简称为“三等分角”。由于这三大几何问题的提出,对后世数学增加了许多新的研究课题。

在雅典相继而起的是柏拉图学派,柏拉图是希腊的著名哲学

家、数学家,他十分注重数学,并非常推崇几何学,以为几何可以培养思维能力。柏拉图及其学派,使用准确的定义、清晰的题设、严密的证明和系统的推理,对数学的发展以及对后世的影响,建立了莫大的功绩。他们研究了棱柱、棱锥、圆柱、圆锥以及五种正多面体,还研究了无理数理论、比例论理论、圆锥曲线性质等。柏拉图学派在柏拉图的指导下,培养出不少优秀学生,欧多克斯、亚里士多德等人都是他的学生。

在雅典,以亚里士多德为首创办了吕园学派。亚里士多德不但在机械学、物理学、气象学、心理学、伦理学上有所贡献,他最大的贡献是创立了逻辑学。其逻辑思想影响整个数学界达 2000 年之久,而且为欧几里得取得的成功铺平了道路。

公元前 4 世纪,亚力山大帝国瓜分为三个国家,最大的为托勒密王朝。托勒密在亚力山大建立了最大的图书馆,并集中大批优秀人才进行科学研究,从而使亚力山大变成希腊的文化中心。不久,亚力山大城就出现了一批著名数学家,其中最杰出的是欧几里得、阿基米德、阿波罗尼等人;欧几里得总结了希腊的数学,写出十三卷的《原本》,对数学的公理化体系产生了巨大影响;阿基米德在他的著作中,记载了关于圆周率的计算,阿基米德螺线的意义,抛物线弓形面积求法,球和球冠面积求法等;阿波罗尼在他的《圆锥曲线论》里,论证了抛物线、椭圆、双曲线可由一个圆锥截得,还探讨了这三种曲线的异同等。

公元 5—6 世纪,由于东罗马的入侵,希腊的文化即告终结,所保留下来的希腊文化遗产,为欧洲的数学发展提供了丰富的营养。

【古印度数学】起源可以追溯到公元前 2000 年左右,从建筑的遗迹、出土的器皿上的古代刻划、铭文,可发现某些数学知识。在公元

前 4 世纪,印度出现了数字,经过了千余年的演变,才变成所谓印度数码,印度人所创造的这套数码和位值记数法,于 8 世纪传入阿拉伯,又经过阿拉伯人的改变,于 13 世纪初期传入欧洲,才形成现今的形式。在印度创立数码的同时,也建立了一些简单的运算法则。

1881 年出土一份所谓“巴克沙利手稿”,经考证,认为是印度 4—5 世纪的遗物,其中记载了用“反演法”解的算术问题,并记载有最初零的记号“0”。7 世纪印度数学家婆罗摩笈多,对整数、分数、数列、比例、面积、体积以及晷影都有所贡献。例如他求得二次方程的一根公式,并引入负数及其加、减算法,还给出一次不定方程的通解的特殊二次不定方程的解。在求解一次不定方程时,他们使用了所谓“库塔卡”的解法。12 世纪在印度数学家婆什迦罗的《丽罗娃祇》以及《算法本原》中,也记录了印度在代数方面的某些成就,他引进了负数的乘、除法运算,认为正数有两个平方根,因而求得二次方程的两个根。

印度人把希腊人的“全弦”,改变为“半弦”,即今之正弦线。这一改变,对三角学的发展起到了积极的推动作用。同一时期阿耶波多第一次用几何方法把圆周作 21600 等份,又取半径为 3438 等份,制造了正弦线表,还论证了一些简单的三角恒等式;到 15 世纪,他们建立了关于 π 的无穷级数展开式以及反正切级数展开式。

在几何方面,印度人的一些公式不十分严密,虽然给出了一些几何体的体积算法,如圆锥、圆台的体积近似算法,但却缺少推导过程。又如给出四边形的面积算法为

$$S = \sqrt{(s-a)(s-b)(s-c)(s-d)}$$

其中 a, b, c, d 分别为四边形之各边, s 为其周长之半。但他们却疏忽了这四边形必需内接于圆的条件。

【阿拉伯数学】公元7世纪初伊斯兰教创立后,政教合一的伊斯兰国家迅速统一了阿拉伯半岛并向外扩张,到8世纪中叶,形成了一个包括中亚、北非、西班牙在内的庞大帝国。所以“阿拉伯数学”实际是阿拉伯帝国统治下的各民族学者共同创造的。

阿拉伯很早就受到印度文化的影响,而且通过各种途径也受到中国文化的影响,与罗马人的战争又使阿拉伯人获得了大量希腊文化遗产,因而阿拉伯数学从一开始就表现出兼收并蓄、交融发展的特征。8—9世纪是阿拉伯数学的翻译时期,希腊和印度的大批天文、数学著作被译为阿拉伯文并被迅速吸收融合,随即进入9—13世纪的鼎盛时期。代数学和三角学是阿拉伯人取得了辉煌成就的领域,数论、几何学方面也有一些创造。重要的数学家有:花拉子米(约780—850)、巴塔尼(约858—929)、阿布尔·维发(940—998)、卡拉吉(10世纪末—11世纪初)、奥马尔·海亚姆(约1048—1131)、纳西尔丁(约1201—1274)、卡西(?—1429)。

8世纪,印度数码传入阿拉伯,经改造后于12世纪传入欧洲,成为近代计算的三大动力之一,这是对数学发展的一个重要贡献。阿拉伯数学家系统地论述了整数、分数、无理数的各种算法,熟知比例算法,却排斥了负数(据说个别学者偶尔使用过)。他们的格子乘法曾广泛流行于欧洲并传入中国。

在代数学方面,花拉子米系统地研究了一次、二次方程的分类和解法,他已注意到二次方程有两个根、但总是排斥负根,而且从他开始,他们的代数著作都用文字叙述,没有接受印度人的缩写记法。卡拉吉重复了花拉子米关于一次及二次方程的工作。并且研究了可根据二次方程的原理求解的双 n 次方程,系统地研究了多

项式的运算,还证明了 $\sum_{r=1}^n r \sum_{r=1}^n r^2$ 、 $\sum_{r=1}^n r^3$ 的求和公式。奥玛尔·海亚姆的三次方程研究在阿拉伯代数学中有重要地位。他对三次方程作了细致的分类,并且研究了利用圆锥曲线的交点求出这些方程解的几何解法。他还讨论了二项式的展开。15 世纪,卡西受到中国宋元数学的影响,给出了一般的二项式系数表,研究了高次方程近似解法,并明确认识了指数运算法则 $a^m \cdot a^n = a^{m+n}$, $a^m \div a^n = a^{m-n}$ 。在 11—13 世纪阿拉伯人的数学著作中还多次出现了中国的盈不足术,称之为天秤术或契丹算法。

阿拉伯人研究了一些二次、三次不定方程,例如 $X^3 + y^3 = z^2$, $X^2 - y^2 = z^3$, $X^2 y^3 = z^2$, $X^3 + 10X^2 = y^2$ 的整数解, $X^2 - y^3 = z^2$, $X^3 + y^2 = z^3$ 的分数解等。还曾有人试图证明 $X^3 + y^2 = z^3$ 没有整数解,即费尔马大定理的一个特例。阿拉伯人的数论成就还有泰比特·伊本·柯拉(约 826—901)推求亲和数的一个法则。

阿拉伯的几何学中既有简单、实用的测量与计算,也继承了希腊几何学的某些理论传统。从 9 世纪开始,有不少阿拉伯数学家研究过欧几里得的平行线,作过试证平行公设的努力,其中有泰比特·伊本·柯拉、奥玛尔·海牙姆、纳西尔丁等,这些工作一直影响着 18 世纪以前欧洲平行线理论的发展。此外,阿布尔·维发首创直尺与定脚圆规作图的研究。15 世纪,卡西运用阿基米德的方法研究圆的度量问题,得出圆周率的 17 位准确数字。打破了中国祖冲之保持 900 多年的记录,并在中国以外最先使用了十进小数。

阿拉伯人以希腊人和印度人的工作为基础,对三角学的研究达到了一个全新的阶段。大约与花拉子米同时的摩瓦兹(以哈巴士之名著称)因测量日晷影长考虑了“直阴影”和“反阴影”(分别相当于余切和正切线值),也有记载说这是巴塔尼的工作。巴塔尼从代

数角度出发研究了多种三角关系式和各种斜三角形的解法,并发现了球面三角的余弦定理。阿布尔·维发明确地引入了正切、余切、正割、余割、并把包括正弦、余弦在内的所有三角线都定义在一个圆上。他编制了间隔为 $15'$ 的正弦表、正切表,证明了正弦的半角及倍角公式。13 世纪,纳西尔丁在两部三角学专著中完成了它的系统化,使之成为一个独立于天文学的数学分支。比欧洲的同类工作早了 200 年。

还应该指出,除了具体的数学工作外,阿拉伯数学以其保存希腊古代著作,吸收和发展希腊、印度及中国古代数学成果,及其对欧洲中世纪与文艺复兴时期数学的巨大影响,在世界数学史上起着沟通东西,继往开来的重要作用。

【中世纪数学】从 5 世纪西罗马帝国灭亡到 12 世纪,一般称之为黑暗的“中世纪”,在欧洲罗马对数学仅有一些粗浅的记载。文艺复兴时代的初期,由于十字军东征,使欧洲人获得大批希腊人和阿拉伯人的著作,欧洲人对这些学术著作倍感兴趣,因而在欧洲掀起了学习、翻译希腊著作的高潮。他们翻译了希腊的欧几里得、阿基米德、海伦、德奥多西阿、托勒密、亚里士多德等人的著作,也把阿拉伯的阿尔·花拉子米等人的著作译成拉丁文,使得即将失传的著作能得以复活。既给欧洲人增加了新鲜血液,又为文艺复兴的科学发展作了准备。13 世纪初期,意大利的斐波纳契《算盘书》传入欧洲,引起欧洲人极大兴趣。在《算盘书》里介绍了很多希腊、阿拉伯的数学知识,还引进了印度-阿拉伯数码。例如介绍了整数、分数的计算方法,还介绍了比例计算、利息计算、算术级数算法、几何级数算法以及斐波纳契级数的算法等。在此之后,欧洲人不但学习、研究了希腊、阿拉伯的著作,也开始建立自己的数学,并分别在法国、英国、

意大利、德国等地的大学开设了数学学科,这些大学以后都成为数学发展的基地。

文艺复兴时期,由于艺术家所创建的透视法,逐步形成了射影几何学;在斐波纳契《算盘书》之后,欧洲也出版了一些数学著作,从而促进了十进分数的理论及其运算的发展;16 世纪初期,最出色的数学成就,是意大利数学家发现了三次、四次方程的代数解法,有的使用了虚数,还改进了当时的数学符号;在三角学发展方面,欧洲人也把三角学从天文学独立出来,使之成为一门独立的学科,并重新定义了各种三角函数的概念,还编制了非常精密的三角函数表。

中世纪,欧洲数学是在吸收并消化希腊、阿拉伯的数学知识之后才逐渐得到了发展的。

【近代数学史】指 17—19 世纪的数学发展概况。具体来说,就是自笛卡儿、费马创立了解析几何之后,把变量引入到数学中,使数学拓展了新的领域;而牛顿、莱布尼茨创立了微积分学;纳白尔、比尔吉发明了对数;巴斯卡、费马、惠更斯兴起了概率论;使得 17 世纪欧洲数学由定量数学发展成为变量数学,并达到了一定的高峰,称为古典高等数学时期。到 18 世纪,在数学里,逐渐形成几何学、代数学、分析学的三大分支;尤其是欧拉把以曲线为主要研究对象的微积分学拓广成以函数为主要对象,使微积分学提到极高的层次,又由于实际的需要,出现了微分方程,不久使得微分方程成为一支重要的学科。到 19 世纪,由于非欧几何的诞生,射影几何的复兴,分析学的严格化,数学的公理化,成为当时的主要研究对象;并为 20 世纪的数学发展,作了必要而充分的准备。

【17 世纪数学】数学史上的辉煌时期。17 世纪初期继续着上一世纪的研究。30 年代,费尔马与笛卡儿分别以古希腊的圆锥曲线理论为基础,通过引入坐标和变量的概念建立了几何中的曲线与代数中的方程之间的内在联系,创立了解析几何学。费尔马的著作完成于 1630 年左右,虽然到 1679 年才得以出版,但其思想与方法已在同时代人中产生了影响,笛卡儿的《几何学》作为其巨著《方法论》的附录,于 1637 年正式出版,标志着解析几何的诞生,并为微积分的创立做了准备。微积分是 17 世纪最辉煌的数学创造,也是自希腊时代以来数学中一系列重要创造的继续和发展,尤其是自文艺复兴以来,由于科学技术中各种实际问题的推动,对变速运动规律的研究,对曲线切线、函数极值、物体重心和引力的研究,以及对曲线、曲面各种度量问题的研究,到 17 世纪中期已经积累了大量具体成果和方法。1666 年 10 月,牛顿完成了第一篇系统的微积分论文,此后在将近 40 年的时间里不断改进和发展了这一理论。莱布尼茨于 1673 年左右独立于牛顿接触到微积分的实质性问题,大约在 1675 年完成了创建微积分的工作。与牛顿的工作相比,他更注重于发展微积分的形式化算法和建立一套简洁、明确而有效的符号,他于 1684 年先于牛顿发表了第一篇微积分论文。牛顿和莱布尼茨的历史功绩在于从众多零散成果中确立了微积分的基本概念,普遍方法和一般形式,使之最终成为一门完整而统一的数学分支。17 世纪,在几何领域发生的另一场重大变革就是射影几何的建立。1639 年,笛沙格在一篇论文中把无穷远元素引入几何学,得到射影几何中的一些基本命题,特别是“笛沙格定理”,是全部射影几何的基本定理。通过研究笛沙格的著作,巴斯卡得到射影几何中另一些重要定理,尤其是著名的巴斯卡定理,并于 1640 年发表了《圆锥曲线论》是自阿波罗尼奥斯以来圆锥曲线论的最大进步。

17 世纪,由于使用字母系数而使证明有了一种尺度,代数学已上升为一门科学,方法和理论都得以大大扩展,1637 年,笛卡儿在《几何学》中给出了关于高次方程正根与负根个数的笛卡儿符号法则。1653 年,巴斯卡在《论算术三角形》一书(1665 年出版)中深入地讨论了二项式系数和基本的组合关系,并给出了数学归纳法的最早陈述。1665 年,牛顿给出了有理指数的二项式定理,1671 年他又给出了求方程实根近似值的牛顿法。1693 年,莱布尼茨创立了行列式理论。

17 世纪的数论主要是在费尔马的推动下进步的,他给出了关于素数、完全数、亲和数、不定方程等方面的许多重要结果,但通常只是给出命题却很少证明。其证明大多由欧拉和拉格朗日在 18 世纪给出,而最著名的费尔马大定理至今仍未获得证明。此外,默森尼研究了形如 2^p-1 (p 为素数)的素数,笛卡儿给出了一条探索亲和数的规则。莱布尼茨得到了后人所说的用于素数检验的威尔逊定理。

1654 年,巴斯卡与费尔马在通信中讨论了“赌博中断问题”,从而共同创立了概率论。在此基础上,1657 年惠更斯发表了概率论的第一篇正式论文——《论赌博中的推理》,其中首次引入了“数学期望”这一重要概念。

这一时期计算技术的一个十分引人注目的进步是原始计算机的发明,1623 年,德国科学家席卡德制造了第一台机械计算机的模型。1642 年,巴斯卡制成了第一台可供实用的计算加减法的机械,1671 年,莱布尼茨制成了可进行乘除运算的计算机。这些工作标志着计算开始由手工时代进入机械时代,并成为后世电子计算机的源头。

17 世纪的数学不仅由于解析几何与微积分的创立而成为近

代数学的开端,它在数学成果、方法与思想各方面的丰富创造也对后世数学的发展产生了极为深远的影响。

【18 世纪数学】18 世纪最引人注目的是微积分的迅速发展并发挥出巨大威力,一些重要概念被不断明确和深化,一些强有力的方法被建立,18 世纪中叶,多元微积分的概念与方法也已初步建立。1748 年,欧拉出版了《无穷分析引论》,标志着微积分发展的一个新阶段,与此同时,微积分向更加广阔的领域扩展,产生了无穷级数、常微分方程、偏微分方程、变分法等重要分支,从而使现代数学中最广阔的领域——数学分析初具规模。这些进展主要是伯努利家族、泰勒、欧拉、克莱罗、达朗贝尔、拉格朗日、拉普拉斯等人的推动下实现的。

与分析领域的巨大成就相比,其他领域略为逊色,但仍有十分明显的进步。

1713 年,雅各布·伯努利的遗著《猜度术》出版,这是概率论的第一部专著,对早期概率论作了全面总结。随后,棣莫弗在《机会论》(1718)、《分析杂论》(1730)中给出了大量新内容,1777 年,法国人蒲产提出了著名的“投针问题”,开始了几何概率的研究。

解析几何的进步与微分几何的创建交织在一起。1704 年,牛顿发表《三次曲线》、研究了 78 种可能情形中的 72 种,1715 年,约翰·伯努利引入了三个坐标面,1731 年,克莱罗开创了空间曲线理论,是三维微分几何的第一个重大发展。此后,欧拉对解析几何、空间曲线与曲面理论都有重要贡献,18 世纪末至 19 世纪初,经过蒙日及其学生们的努力,终于把微分几何作为一门独立的学科建立起来,同时他还是画法几何的奠基人。

在代数学方面,对五次以上代数方程式解的探讨仍在继续,

其中拉格朗日的工作对 19 世纪群论的诞生有着重要意义。线性方程组及行列式已经建立了独立而系统化的理论。1799 年,高斯继承欧拉、达朗贝尔、拉格朗日的工作,终于给出了代数基本定理的第一个证明。

18 世纪的数论主要是由欧拉推动的,他证明了费尔马提出的许多数论命题,系统地探讨了亲和数,并提出了二次互反律,这是 18 世纪数论中最富创造性的成果,于 1801 年被高斯证明,并称之为“算术中的宝石”。哥德巴赫(1742)与华林(1770)分别提出了他们的著名猜想,成为后世数论中的重大课题。18 世纪后期,拉格朗日、勒让德也有重要的工作。与 17 世纪相比。18 世纪的数学虽然没有提供那样众多新颖而基本的概念与方法,却施展了高度的技巧,并根据科学技术特别是力学与天文学的需要,提出和解决了大量新的问题。在数学方法上则完成了从几何方法向解析方法的转变。18 世纪的数学在迅速发展的同时也暴露出其弱点:忽略了数学方法与基础的严密性,从而导致了一系列矛盾,而 18 世纪的数学家对此似乎并不十分担心。这一时期的数学家大多又是物理学家,人们坚信概念与方法在物理学上的正确性保证了它们在数学上的正确性。

【19 世纪数学】1801 年高斯发表了《算术探究》一书,表面看来似是“算术”著作,但实际是一部非常严格的数论专著;书中不但建立了严格论证的典范,还表达了追求严密论证的思想。1837 年,狄利克雷在高斯的基础上,进一步开辟了解析数学论的新领域;之后,黎曼的工作也促进了解析数论的发展。

高斯、罗巴切夫斯基、鲍耶各自独立地发现了非欧几何学,从而打破了 2000 年来欧几里得几何学一统天下的局面。19 世纪 20

年代,高斯开拓了微分几何学,经过黎曼的工作发扬了微分几何学的研究,并建立了任意维空间内蕴几何学。

在 19 世纪,微积分学的严格化成了当务之急,而柯西对微积分学严格化起到了巨大作用;此外,魏尔斯特拉斯发展了柯西的论说,提出用 ϵ - δ 定义极限概念,并给予函数连续性的确切定义以及一致收敛的定义,使得微积分学以至分析学趋于完善的境界;尤其魏尔斯特拉斯、康托尔、戴德金等人建立了无理数的精密定义之后,更使得分析学达到完善的地步。

在级数方面,1822 年,傅立叶不仅扩展了偏微分方程的研究,而且还提出三角级数的研究,从而获得许多丰硕的成果。到 30 年代,伽罗瓦创造性地提出代数的最基本的结构——群论,也是抽象代数的重要组成部分。

1899 年,希尔伯特发表《几何基础》,书中阐述了欧几里得几何学的公理系统;因而数学家们为数学的各个分支纷纷建立各自的公理体系,于是形成数学走向公理化的高潮。1900 年,希尔伯特在第二届国际数学家大会上作了题为“数学问题”的报告,提出当时数学中尚未解决的 23 个问题,这既是 19 世纪的未决问题,又是 20 世纪数学的发展方向。

【赵爽】中国古代数学家。字君卿。约生活于东汉末至三国时代。研究过张衡的天文数学著作《灵宪》和刘洪的《乾象历》。他为《周髀算经》作序并逐段对经文作详细注解。他撰成“勾股圆方图”注,附录于《周髀算经》首章的注文中,附图六张,简明总结了后汉时期勾股算术的辉煌成就。他重新叙述了勾股原理并严加证明,然后用短短 500 余字写出了 20 多条与勾股原理有关的命题,并给予严谨证明。他将勾股定理表述为:“勾、股各自乘,并之为弦实。开方除之,

即弦”。证明方法为：“按弦图，又可以勾股以相乘为朱实二，倍之为朱实四。以勾股之差自相乘为中黄实，加差实一，亦成弦实”。证法的基本思想是：图形经过割补后，面积不变。这种方法后来发展为演段术，成为中国古代数学方法的特色之一。虽然赵爽证明勾股原理比古希腊数学家毕达哥拉斯及欧几里得为晚，但他用面积理论证明的具体步骤是十分精巧的。印度数学家婆什迦罗二世曾用过类似的证明方法，但晚了约 900 年。赵爽在《勾股圆方图》中还研究了二次方程问题，得出了世界上最早的求根公式之一，并得出与“韦达定理”类似的结果。他对分数也有研究，是“齐同术”的创始人，在计算天文问题时将“齐同术”应用于乘除法。在《日高图》下的注文中，他用面积证明了“日高术”，这后世中国“重差术”的发展，奠定了理论基础。

【刘徽】中国古代数学家。约生活于魏、晋期间。据《隋书·律历志》称：“魏陈留王景元四年(263)刘徽注《九章》”。他在长期精心研究《九章算术》的基础上，采用高理论，精计算，潜心为《九章》撰写注解文字。他的注解内容详细、丰富，并纠正了原书流传下来的一些错误，更有大量新颖见解，创造了许多数学原理并严加证明，然后应用于各种算法之中，成为中国传统数学理论体系的奠基者之一。如他说：“徽幼习《九章》，长再详览。观阴阳之割裂，总算术之根源，探赜之暇，遂悟其意。是以敢竭顽鲁，采其所见，为之作注”。又说：“析理以辞，解体用图。庶亦约而能周，通而不黷，览之者思过半矣。”他除为《九章》作注外，还撰写过《重差》一卷，唐代改称为《海岛算经》。他的主要贡献在于创造了割圆术，运用极限观念计算圆面积和圆周率；创造十进分数、小单位数及求微数思想；定义许多重要数学概念，强调“率”的作用；运用直角三角形性质建立并推

广重差术,形成特有的准确测量方法;提出“刘徽原理”,形成直线型立体体积算法的理论体系,在例证方面,他采用模型、图形、例题来论证或推广有关算法,加强说服力和应用性,形成中国传统数学风格;他采用严肃、认真、客观的精神,差别粗糙、错误的论述,创造精细、有逻辑的观点,以理服人,为后世学人树立良好的学风;在等差、等比级数方面也有一些涉及和创意。经他注释的《九章算术》影响、支配中国古代数学的发展 1000 余年,是东方数学的典范之一,与希腊欧几里得(约前 330—275)的《原本》所代表的古代西方数学交相辉映。

【祖冲之】(429—500) 中国南北朝时期数学家、科学家。字文远。范阳郡道县(今河北涿源)人。虽原籍北方,但几代祖先都在江南做官且通晓历法。祖父掌管土木工程建筑,父亲也学识渊博。他从小有机会接受家传科学知识,青年时代进入专门研究学术的华林学省学习研究。一生中先后任过多处官职。他在天文历法、机械、数学等方面卓有成就。他计算出圆周率 π 的真值在 3.1415926(朏数)和 3.1415927(盈数)之间,分数形式的 π 值为:约率 $22/7$ (≈ 3.14),密率 $355/113$ (≈ 3.1415929),是当时世界上最先进的成就。他和儿子祖暅共同圆满解决了球体积的计算问题,得到正确的球体积公式,提出“刘祖原理”,比后来西方所称的“卡瓦列里(1598—1647)原理”早 1000 多年。他们父子所著的算书《缀术》被列为《算经十书》之一,唐代为算学必修课本之一,可惜书已失传。祖冲之创制了《大明历》,最早把岁差引进历法,采用 391 年 144 闰的闰法,并作出几项新规定,开辟了历法史新纪元。他设计制造了水碓磨、铜机指南车、“千里船”等多种巧妙的机械设施。他还精通音律,写过小说。据各种史籍记载,他的著述很多,可惜绝大部分已失传。

【祖暅】(约 5—6 世纪) 中国南朝数学家、天文学家。字景烁。祖冲之之子。继承父愿,三次上书政府建议采用《大明历》,终如所愿。他修著完成《缀术》一书,并自著有《漏刻经》、《天文录》等书,可惜也多失传。祖冲之父子为中国古代数学及科技发展做出了杰出贡献。

【王孝通】中国唐代数学家、天文学家。唐初任历算博士,曾任通直郎太史丞,并参加修历工作。著有《缉古算经》,被列为《算经十书》之一。它是中国及世界现存古算书中、最早提出三次方程代数解法的著作。他对三次方程系数的称谓:实、方、廉、隅与刘徽开立方术注文是相一致的,可以认为是对《九章算术》开立方术的发展。他在利用勾股术进行几何计算方面也有创新。他的成就解决了当时工程中出现的大量问题,在学术上和应用上都有重要价值。他用几何方法和技巧建立了三次方程的各项系数,但限于当时的数学水平,未能给出列方程的一般方法,对方程计算也仅取其正根。他的著作词意深奥,不易通晓,以致后世出现大量评注之作。但他的成就仍为宋、元时期“天元术”的发展奠定了基础,在中国古代数学史上占有重要地位。

【张遂】(683—727) 中国唐代数学家、天文学家。魏州昌乐(今河南南乐)人。他自幼学习历象,21 岁出家,取名一行,后世亦称僧一行。以学识渊博闻名。曾跟随善无畏、金刚智等研究佛经和天文数学,参加翻译《大日经》的工作。亲自起草并主持编制了《大衍历》,这一历法比唐代已有的其他历法都精密,行用近百年。他在《大衍历》中创造“步日躔术”,利用自变量不等间距二次内插法,推广了

刘焯的“关于太阳运行不等速”的内插法公式。还使用了具有正弦函数性质的表格和含有三次差的近似内插公式。他还改进制造天文仪器以观测天象,发起并主持了大规模的全国天文大地测量,实现了世界上最早的子午线实测,比西方最早的子午线实测早 90 年。

【贾宪】中国宋代数学家。曾著有《黄帝九章算法细草》九卷及《算法斠古集》二卷,惜均已失传。在杨辉的著作《详解九章算法》及其附录《九章算法纂类》中记录了贾宪的部分工作,得以知道他在任意高次方程数值解法方面所作出的卓越贡献。他创造了“立成释锁平方法”、“增乘开平方法”、“贾宪立成释锁立方法”、“增乘开立方方法”。“释锁”是宋元数学家开方或解数字方程的代用名词,“立成”是唐以后推算各种数据时所用的算表的通称。贾宪的方法就是利用算表进行开方计算。他的算表是指数为正整数的二项式—— $(X+a)^n (n=1,2,\cdots,6)$ 的展开式各项系数的排列,称为“开方作法本源”图。这一发现比西方所称的“帕斯卡三角形”早 600 多年。他的“增乘开方法”也比现在流行的鲁菲尼-霍纳方法早 750 年,演算步骤基本相同。

【沈括】(1033—1097) 中国北宋科学家、政治家。字存中。自幼随父走南闯北,见多识广。父亲为官吏,母亲精通文理,对他早年学习有一定作用。12 岁入学堂,博览群书,接触社会较多。先后任过主簿、县令、司理参军、司天监、军器监、安抚使等许多官职。在各种任职期间,治理水土,编校书籍,研究天文历法,改进天象观测仪器、计时仪器,主持编修《奉元历》,研究军事知识,改进兵械,推行新法,改革盐钞,修正役法,抗击外族侵扰,做出了大量有声有色的政

绩。1087 年移居润州(今镇江)的梦溪园私宅,倾毕生余力,潜心著述。1088 年前后完成不朽科学巨著《梦溪笔谈》。现存传本 26 卷,《补笔谈》3 卷,《续笔谈》1 卷,共 30 卷。这部内容丰富的学术著作有关科学技术的约 200 多条,内容涉及数学、物理、化学、天文、地质、地理、气象、工程技术、生物、医学等多种学科。数学方面,他在《梦溪笔谈》卷十八《技艺》中,创立了二阶等差级数求和的一种方法“隙积术”以及求弓形的底和弧长的一种方法“会圆术”。他将数学知识广泛应用于天文、历法、工程、军事中,得出不少重要结果。在物理学方面,他最早发现了地磁偏角的存在,阐述凹镜成像原理,研究共振规律等。在地质学方面,他利用地形研究水的浸蚀作用;从岩石中的生物遗迹,推知冲积平原形成的过程。在医学方面,他精究药用植物与医学,辑成《良方》十卷。此外,他首先提出石油的命名;发明木屑面糊溶蜡的地势模型法;对当时的活字印刷术、炼钢、炼铜等最新的科学技术做了全面细致的阐述。他的这部巨著不仅是中国古代科学史的杰作,也是世界科技资料库中的一份宝贵遗产。他是中外学者公认的卓越科学家,被李约瑟尊称为“中国科学史上的坐标”。

【秦九韶】(约 1202—1261) 中国南宋数学家。字道古。普州安岳(今四川安岳)人。父亲秦季樵曾为南宋官吏。他早年随父迁居,先后向太史、“隐君子”等人拜师求学。后曾在江南多处地方做官,饱经战争忧患。因政界内部倾轧,补贬梅州(今广东省梅县),不久便死于任所。他自幼聪颖,涉猎广泛,“星象、音律、算术以至营造等事无不精究”。1244 年,母亲故去,他为母守孝期间,潜心研究著述,于 1247 年完成杰作《数书九章》,南宋时称为《数学大略》或《数术大略》,明代又称为《数学九章》。全书共 18 卷,81 题,分为 9 类,每

类 9 题。内容涉及卜卦、天文历法、田亩、测量、运输、税收、仓储、营建、军事、市场交易等多种实际问题,体现了秦九韶“设为问答以拟于用”的思想。他创造的求解一次同余式组的“大衍求一术”以及发展完成的求高次方程正根的“增乘开方法”,在世界数学史上占有崇高的地位。他的同余式组的解法比西方由高斯建立的解决同样问题的理论早 554 年;他的高次方程求正根的方法比西方完全相同的“霍纳法”早 572 年。此外,他还独立发现了与古希腊海伦公式等价的三角形求积公式;改进了线性方程组解法,使运算更加迅捷简便。《数书九章》不仅是一部数学专著,还翔实地反映了南宋的社会经济状况,成为非常有价值的历史资料。中外学者对秦九韶和他的杰作《数书九章》做了大量的研究工作,发表了许多重要研究成果。

【李冶】(1192—1279) 中国金元时期数学家。字仁卿,号敬斋。真定府栾城县(今河北栾城)人。生于大兴城(北京),童年到元氏县(今河北元氏县)求学,“自幼喜算术”。1230 年考中“词赋进士”。曾到钧州(今河南禹县)居官两年。因战乱而放弃功名,元世祖忽必烈多次召见均辞官不受。40 岁后便潜心从事数学研究。先后隐居于崞山(今山西崞县)桐川及元氏封龙山,著书讲学,自得其乐。生平著述颇多,惜多已失传,他的得意之作《测圆海镜》及《益古演段》作为重要的数学遗产而流传到今。他在临终时曾说:吾平生著述,死后可尽燔去,独《测圆海镜》一书,虽九九小数,吾常精思致力焉,后世必有知者”。他于 1248 年著成《沿圆海镜》十二卷,系统讲述“天元术”,发展了高次方程的建立和表示方法,开创了中国的符号代数学,所论内容均与勾股形、内切圆,旁切圆、直径等几何图形有关。无论在内容上或方法上对传统数学都有较大的扩充与发展。

1259 年,他在前人蒋周的《益古集》的数学成果基础上,写成了《益古演段》三卷,普及天元术,内容大多涉及平面图形间的面积关系,用天元术和等积变换求解,相当于现代的代数和方程解法。清代学者阮元称《测圆海镜》是“中土数学之宝书”,李善兰称它是“中华算书实无有胜于此者”。

【杨辉】中国南宋末数学家。字谦光。生卒年不详,约生活于 13 世纪。钱塘(今浙江杭州)人。虽然有关他的生平资料极少,但他编著的大量数学著作尚有多种流传迄今。使后人得以对当时数学的发展情况有较多了解,对早期已失传的一些数学著作可略知一二。据记载,他所编著的数学书共 5 部 21 卷:《详解九章算法》12 卷(1261)、《日用算法》2 卷(1262)、《乘除通变本末》3 卷(1274)、《田亩比类乘除捷法》2 卷(1275)、《续古摘奇算法》2 卷(1275)。他在研究高阶等差级数的“垛积术”以及组合数学的“纵横图(幻方)”方面有独到的见解和成就。他尤为重视数学中计算方法的简捷和应用性,主张“须责实用”的思想。他不仅是著述甚丰的数学家,还是杰出的古代数学教育家。他的著作多是为普及数学知识和推广简捷的计算方法而编著,还制定了“习算纲目”这样的明确实用的数学教学计划。他提倡循序渐进、熟读精思、以简见繁、举一反三、触类旁通的教学和学习方法。他编写的算书广泛引证古代数学典籍,使数学史上大量宝贵资料得以保存,具有很高的科学价值。他的数学教学思想和教学方法,至今仍有重要的参考作用。

【朱世杰】(1300 前后) 中国元代数学家。字汉卿,号松庭。生卒年不详,约生活于 12 世纪末 13 世纪初。寓居燕山(今北京附近)。曾长期以数学研究和数学教育为业游学四方。著有《算学启蒙》三卷

(1299)和《四元玉鉴》三卷(1303)传世。他的学术活动时间大约在13世纪末20年到14世纪初20年之间。《算学启蒙》从简单的四则运算出发,逐步深入到高次开方、天元术等高深内容,体系完整,可算是当时较好的一部数学启蒙教科书,曾流传至朝鲜和日本。《四元玉鉴》重在阐述高次方程组(四元以内)解法、高阶等差级数求和、高次内插法等当时较高深的数学问题。在这几方面都获得了当时领先于世界数学水平的成果。他的可推广为任意高次的招差法公式,比西方早400余年。清代罗士琳在《畴人传·续编》中评论他说:“汉卿在宋、元间,与秦道古(秦九韶)、李仁卿(李冶)可称鼎足而三。道古正负开方,仁卿天元如积,皆足上、下千古,汉卿又兼包众有,充类尽量,神而明之,尤超越乎秦、李之上。”他和秦九韶、李冶、杨辉一起被称为宋、元数学四大名家。由于他们的努力、使宋、元时期的中国数学在许多方面都居于世界前列。

【王恂】(1235—1281) 中国元代天文学家、数学家。字敬甫。中山唐县(今河北唐县)人。幼年随父学习数学、天文、后到邢台附近的紫金山师从刘秉忠(忽必烈的谋臣)学习天文历法和数学,与郭守敬同学。元世祖至元十三年(1276),他奉命与郭守敬一道组织太史局(1279改称太史院),以数学家身份主持改历工作,任太史令,承担天文观测和计算方面的任务。1280年编成《授时历》新历。他虽无专著传世,但在编制新历过程中,多次使用三次内插法,并制造了三次差分表;第一次研究了球面上弧与弧的关系;把高次方程用于历法研究;这些突出成就在当时世界数学发展中也居于领先地位。后人曾赞誉他为“算术冠一时”的数学家。

【项名达】(1789—1850) 中国清代数学家。原名万准,字步莱,号

梅倡。浙江仁和(今杭州)人,祖籍安徽歙县。1816年中举人,考授国子监学正。1826年中进士,改署知县,未就职。曾与友人研讨数学多年,担任过余杭苕南书院主讲(1837前),后在紫阳书院执教。1846年冬休职还家。主要数学著作有《勾股六术》1卷(1825)、《三角和较术》1卷(1843)、《开诸乘方捷术》1卷(1845)、合刻为《下学^⑥算学三种》印行。前两种内容浅显易懂,对勾股形、平面三角形、球面三角形的各边及其和、差的互求关系,做了较系统的分类总结。后一种提出了幂指数为 $1/n$ 的二项式定理以及用逐次逼近法开 n 次方的递推公式。在他晚年又集中精力撰写了《象数一原》6卷(1849),主要论述三角函数幂级数展开式问题,推广了明安图和董祐诚的结果,得到两个新公式,还推导出椭圆周长公式等。由于年老病重,他的最后著述未能完成。遵其嘱托,友人戴煦于1857年补写完成,并为椭圆求周术补作图解1卷,故现传本《象数一原》共7卷。项名达用初等方法求得椭圆周长的级数表达式,与用近代椭圆积分所得完全相同,实为难能可贵。

【赵友钦】中国宋、元时期数学家。字子公。鄱阳人。生卒年不详,约生活于13世纪后期。原为宋宗室子弟,南宋亡前后隐遁埋名,研究数术、天象之类,世人称为缘督先生。著有《革象新书》5卷。其中第五卷“乾象周髀”篇内记载了他关于圆周率的研究。他继承了刘徽的割圆思想和方法,从圆内接正方形算起,依次计算出正 4×2^n ($n=0,1,2,\dots,11,12$)边形的一边之长(最多为 $4 \times 2^{12}=16384$ 边形)。证明了 $\pi=355/113$ 的精确性,并与以往历代数学家所求的 π 值比较,认为取 π 值为 $355/113$ 最为准确。他指出了圆内接正多边形的分割“…初之小方,渐加渐展,渐满渐实。角数愈多而其为方者不复为方而为圆矣”。正是刘徽割圆极限思想的通俗解释。

【王文素】中国 16 世纪数学家。结合当时的社会需要介绍商业数学,著有《新集通证古今算学宝鉴》(1524)简称《算学宝鉴》,内容十分丰富包括河图洛书、纵横图、筹算术、珠算术等。受杨辉的影响,书中记载了许多纵横图式。王文素对纵横图作过研究,虽然大部分图式与杨辉的雷同,但也有一些独立的成果。当时珠算术已兴起,但王文素尚未充分认识到珠算术的优越性,只叙述介绍了一些珠算加减法口诀如“一起四作五”。“无一去五下还四”等。《算学宝鉴》主要介绍筹算方法,给出了筹算的除法口诀,如一归的撞归,起一口诀等。王文素与程大位同时代,都对实用算术的发展做出过贡献,但其影响远不如程大位。

【吴敬】中国明代数学家。字信民,号主一翁。浙江仁和(今杭州)人。生卒年不详,约生活于 15 世纪。曾任浙江布政使司的幕府,掌管全省田赋和税改的会计工作。以善算而闻名当地。1450 年撰成《九章算法比类大全》10 卷,包括商业实用算术、各种开方术等 1000 多个问题。首次给出国外流行的格子乘法,称之为“写算”乘法。书中还有“不用算盘,至无差误”等句,是迄今所发现的较早提及算盘的数学著作之一。他在中国古算的普及和广泛应用于生产、生活实践方面做了重要工作。

【程大位】(1533—1606) 中国明代数学家、珠算家。字汝思,号宾渠。安徽休宁人。少年时博览群书,尤其对书法和算学有兴趣。一生未做官。20 岁起在长江中、下游一带经商,精于计算,并遍访名师、搜集数学古籍。40 岁左右退居还家,潜心算学研究,60 岁时(1592)著成《直指算法统宗》17 卷(简称《算法统宗》),并详述编写

经过。书中详细记载了珠算盘的定位方法和加减乘除口诀,最早提出了用珠算开平方和开立方的方法,珠算口诀已趋完备,系统总结了珠算术的理论。这是中国古算书中印行数量最多、流传和影响最广的一部著作。明朝末年,曾传入朝鲜、日本及东南亚各地,对传播和普及珠算算法,起到重要作用。书中还收录了宋元以来各种数学书目 51 种、指算速算法“一掌金”、多种纵横图(即幻方)、及由伊斯兰国家传来的“铺地锦”乘法(格子乘法)等。1598 年,他对这部书进行删节,取其切要部分,另编为《算法纂要》4 卷,在安徽屯溪刊行。

【徐光启】(1562—1633) 中国明末科学家。字子先,号玄扈。上海徐家汇人。生于上海,卒于北京。40 岁以前曾以教书为业。41 岁时(1603)入天主教会,认为天主可“补益王化,左右儒术,救正佛法”。40 岁时(1604)在北京参加会试中进士,入翰林院。后几经迁升,临终前官至文渊阁大学士。他是中国介绍和吸收欧洲科学技术知识的先驱和积极推动者,认为科学技术可富国强兵,有益民用。他在数学、天文学、农学等方面均做出了重要贡献。1607—1612 年间,他与利玛窦(1552—1610)合作相继译出《几何原本》前 6 卷(1607)、《测量法义》(1607)、《简平仪说》(1611)、《泰西水法》(1612)等著作,最早将欧几里得几何学及其严密的逻辑体系和推理方法、西方的三角学和测量术等引入中国。他还亲自撰写了《测量异同》和《勾股义》,进行中西比较和借鉴。1613 年后,他的注意力转向农学,编著了《农政全书》60 卷(1625—1628),包括历代农书的精华、欧洲一些先进的科学知识及他本人的研究成果,这是一部促进中国农学发展的重要著作。1629 年,他受皇帝之命组织历局,修正历法,领导了明代第一次重大的历法改革。数年后编译出

137 卷的巨著《崇祯历书》，其中有 105 卷由他亲自定稿，系统介绍西方天文历法和三角学，对推动中国天文学和数学的发展起了重要作用。

【年希尧】(? —1738) 中国清代数学家。字允恭。广宁(今辽宁北镇)人。一生在多处为官。对科学颇有兴趣，主要研究数学和医学。他在宫廷供职时，曾向为清宫绘画的意大利画家郎世宁学习欧洲透视画法，专心研究，著成《视学》二卷(1729)。后继续研究，补充附图，1735 年又出修订版。这部书是在西方透视学知识影响下完成的，是第一部介绍西方画法几何学的著作。但年希尧结合中国绘画传统，多举中国器皿、建筑物等图样为例，加入了许多自己独创的研究成果，有些成果是国外后来才解决或充实完善的。书中图例丰富精美，并清楚讲述了绘制直观图(立体图)和平面图(投影图)的原理及方法。所刊刻的数学著作还有《测算刀圭》、《三角法摘要》、《面体比例便览》等，论及三角、三角对数、对数等知识。

【明安图】(约 1692—1765) 中国清代数学家、天文学家、地理测绘学家。字静庵。蒙古族正白旗人。少年时为官学生，后被选派至钦天监，专门学习天文、历法和数学。结业后，在钦天监工作达 50 余年，曾任时宪科五官正，后升任钦天监监正。曾亲往新疆测绘经纬度，完成了准噶尔地区 and 天山南部地区的测量与绘图任务，为后来《乾隆内府舆图》的绘制奠定了重要的基础。他在天文考测、推算方面曾担负重要责任，参与编著了《律历渊源》、《历象考成》、《历象考成后编》、《仪象考成》等天文、历法著作。他花费 30 余年时间，对法国传教士杜德美(1668—1720)向中国学者介绍的三个无穷级数公式(π 的无穷级数公式，正弦和正矢的幂级数展开式)进行刻苦

钻研,终于用“割圆连比例法”证明了这三个公式,同时还得出有关弦、弧、矢和半径的相互关系的另外六个级数公式,成功地将中国传统数学知识与当时传入的西方数学知识结合起来。他的数学成就总结在《割圆密率捷法》一书中。在他生前完成了一部分手稿,后由其学生陈际新等人继续研究整理,于 1774 年全部完成。他的成就对清代的无穷级数研究有很重要的影响。

【汪莱】(1768—1813) 中国清代数学家。字孝婴,号衡斋。安徽歙县人。出身贫寒,幼年受教育不多。青少年时勤于钻研、刻苦自学、博通经史、天文、数学及乐律等知识。成年后以教书为业,曾与李锐、焦循为友,共同探讨学术问题。1807 年在北京考取八旗官学教习,1809 年选授石埭县(今安徽省石台县)儒学训导。曾在国史馆纂修《天文志》和《时宪志》。著有《衡斋算学》7 册,《衡斋遗书》9 卷,合刻刊行。另有《递兼数理》、《参两算经》等著作。他在数学方面的主要功绩是提出高次方程正根的存在和数目问题及判别条件,发展了中国古代的方程理论。他系统讨论了球面三角形六种基本问题的求解,得出了有解和无解的条件。在组合问题、勾股问题、P 进制、高阶等差级数求和等方面均有创见。曾设计制造过浑天、简平、一方、勺漏等天文仪器。

【梅文鼎】(1633—1721) 中国清初数学家、天文学家。字定九,号勿庵。安徽宣城人。少年时随父梅士昌及私塾老师罗王宾学习天文知识,29 岁向倪正学习《大统历算交食法》,42 岁后专心致力于天文数学的研究。曾在北京等地教书数年。60 多岁时回到家乡,继续研究天文数学,直至去世。据他自撰的《勿庵历算书目》(1702)称,他著有天文数学著作 70 余种,其中数学著作 20 余种。流传下

来的以《梅氏历算丛书辑要》62卷(1745,梅穀成作序)最完备。他一方面坚信中国传统数学“必有精理”,不遗余力研究、发掘。同时主张“去中西之见,以平心观理”、“法有可采何论东西,理所当明何分新旧”,对当时传入的西方数学进行全面、系统的整理和移植,并有所创新。他的著作有释义,有理论,有解法,有应用,范围几乎涉及当时所能研究的各个领域。当时慕名向他求教的学者很多,康熙皇帝亦曾于1705年三次召见他,向他请教天文数学。他在当时被誉为“国朝算学第一”。

【梅穀成】(1681—1763) 中国清代数学家。字玉汝,号循斋,又号柳下居士。安徽宣城人。梅文鼎的孙子。曾任左都御史之职,参加过编著《律历渊源》100卷(1723)的工作。其中《数理精蕴》53卷,主要讨论数学问题,是当时的一部初等数学百科全书。书中将西方传入的“代数学”译作“借根方”。当时中国传统数学中的天元术失传已久,无人通晓。梅穀成在深入研究借根方之后,又重新细读郭守敬的《授时历草》和李冶的《测圆海镜》等书,发现“天元一术”就是“借根方”,于是作《赤水遗珍》(1761),附录在《梅氏历算丛书辑要》中作为卷六十一,说明了两种方法及理论的相同实质,使天元术重显于世。他还自著有《兼济堂历算书刊缪》1卷,《柳下旧闻》16卷。他增删《算法统宗》11卷,并重编了《梅氏历算丛书辑要》六十二卷。

【李善兰】(1811—1882) 中国清代数学家、教育家。原名心兰,字竞芳,号秋纫,别号壬叔。浙江海宁县硖石镇人。9岁读《九章算术》,14岁学欧几里得《几何原本》前6卷。以后又买回李冶的《测圆海镜》、戴震的《勾股割圆记》等算书研读。常与嘉兴等地的数学

家顾观光、张文虎、汪日桢、戴煦、罗士琳、徐有壬等人研讨学术。30岁左右即已在数学、天文方面造诣精深。1845年前后,他在中国传统数学理论的基础上进行研究,独立深入到解析几何和微积分学的领域,发表了研究成果——尖锥求积术。尖锥术理论主要见于他的《方圆阐幽》、《弧矢启秘》、《对数探源》三种著作中。他在数学研究中还有垛积术和素数论等重要成果。主要见于《垛积比类》和《考数根法》等著作中。1867年,他在南京出版《则古昔斋算学》,汇集他20多年的著述13种24卷,约15万字。其中有《方圆阐幽》、《弧矢启秘》、《对数探源》、《垛积比类》、《四元解》、《麟德术解》、《椭圆正术解》、《椭圆新术》、《椭圆拾遗》、《火器真诀》、《对数尖锥变法释》、《级数回求》、《天算或问》。此外,还著有《考数根法》、《粟布演草》、《测圆海镜解》、《九容图表》刊行。未刊行著作还有《造整数勾股级数法》、《开方古义》、《群经算学考》、《代数难题解》等。他任北京同文馆天文算学总教习时,审定了《同文馆算学课艺》、《同文馆珠算金旻》等数学教材,并从事数学教育10余年,直至去世,培养了一大批数学人才,是中国近代数学教育的鼻祖。他还曾翻译、合译了大量西方数学、数理哲学、物理学、植物学等近代科学著作,创译了许多至今仍沿用的科学名词,如“代数”、“函数”、“方程式”、“微分”、“积分”、“级数”、“植物”、“细胞”等。

【阮元】(1764—1849) 清代数学家。字伯元,号云台,江苏仪征人。阮元博览群书,酷爱历法算术,除自己考注外,又组织名家对天文、数学及技术书籍进行整理和注疏,其最主要的功绩是组织编纂《畴人传》一书。《畴人传》以人为纲,用传记体裁写出自上古传说人物黄帝下迄清初的天文学家和数学家的生平事迹和科学工作,搜集了各个时期天文、数学史料,表彰和评论了专业人才的卓越成就,

对于批判地接受科学文化遗产,形成科学研究的风气有其积极意义。从乾隆乙卯(1795)至嘉庆己未(1799),历时四年编成,共46卷,书中记载了从黄帝到嘉庆4年已故天文学家和数学家共270余人,其中有数学著作传世者近50人。书中还附录了明末以来传入中国的天文、数学书中涉及到的外国天文学家、数学家和来华传教士共41人。由阮元主编的《畴人传》开系统阐释天文算学史之先河,被誉为天文算学史的开山鼻祖。

【李俨】(1892—1963) 中国数学史家、铁路工程师。原名禄骥,字乐知。福建闽侯人。1912年入唐山路矿学堂学习。1913年考入陇海铁路局,历任实习生、测量员、助理工程师、总段长、工务处长、副总工程师等职,为中国的铁路建设事业,辗转工作40余年。1955年调入中国科学院,任历史研究所研究员。1957年出任中国科学院自然科学史研究室主任,直到逝世。曾任中国科学院社会科学哲学部学部委员。他是中国数学史研究的学科奠基人。早在1911年前后,就开始了中国数学史的整理和研究工作,毕生搜集了大量中国古算书,去世后由他的家属捐赠给中国科学院自然科学史研究所,成为研究中国数学史的宝贵资料。1917年发表第一篇数学史论文《中国数学史余录》。1920年前后发表重要论著《中国数学源流考略》。在长期的业余研究中,先后发表论文百余篇,其中重要者汇为《中算史论丛》1—5集(1954—1955)。另有专著《中国算学史》(1937初版,1955修订)、《中国数学大纲》(1958)、《中国古代数学史料》(1954初版,1963修订)、《中国十三、十四世纪中国民间数学》(1957)、《计算尺发展史》等。其他方面尚有著作,数学方面的《微积分学初步》(1936)、工程学方面的《铁道定线法》(1938)等。

【钱宝琮】(1892—1974) 中国数学史家、数学教育家。字琢如。浙江嘉兴人。早年于嘉兴秀水县学堂肄业。1907 年入苏州铁路学堂学习,1908 年考取浙江省官费欧美留学生,到英国伯明翰大学土木工程系学习,1912 年毕业回国。先后在苏州中等工业学校(1912)、南开大学(1925)、中央大学(1927)、浙江大学(1928)任数学教员、副教授、教授等职。1956 年调入中国科学院,任自然科学史研究所研究员,直至去世。他是中国数学史研究的学科奠基人之一。早在 1912 年回国从事数学教育时,就开始了中国数学史和天文学史的研究,1921 年开始发表数学史论文,毕生发表论文数十篇,去世后整理为《钱宝琮科学史论文选集》(1983),由科学出版社出版。他还陆续发表了若干专著,如《中国算学史》上册(1932)、《古算考源》(1933)、《中国数学史话》(1957)、《算经十书》(校点,1963)。此外,他主持编写了《中国数学史》(1964)、《宋元数学史论文集》等,并担任《科学史集刊》(1958 创刊)的主编。他在数学史研究及培养专业研究人员,在数学教育等方面均做出了突出成就。

【严敦杰】(1917—1988) 中国天文学史家兼中国数学史家。生于浙江嘉兴。字季勇。其父母早年双亡,中学毕业多年后,在中央大学数学系工读一年多。并长期在重庆、成都、上海、天津等地业余从事中国科学史研究;中华人民共和国成立后,历任中国科学院自然科学史研究所研究员、所长、《自然科学史研究》主编等职。共发表科学史论著 160 余篇,计 300 多万言,遍及科学史各个领域,尤其在中国天文学史及中国数学史等方面多所建树。与中国数学史专家李俨、钱宝琮在中国科学史界形成鼎立而三的局面。遗著《中算史论丛》第六集、第七集、第八集、第九集、第十集,作为中国数学史专家李俨巨著《中算史论丛》第一至第五集之续,成为中国数学史

领域上的经典宏著。

【毕达哥拉斯】(约前 560—前 480) 古希腊哲学家、数学家、天文学家、音乐理论家。生于萨摩斯(今希腊东部小岛),卒于梅塔蓬图姆(今意大利半岛南部塔兰托)。早年游历埃及、巴比伦等地,接受古代流传下来的天文学和数学知识。后定居于意大利半岛南部的克罗托内,在那里广收门徒,组织了一个政治、宗教、学术合一的秘密团体。这个团体后在政治斗争中遭破坏,他逃到塔兰托,终被杀害。他的团体组织严密,是带有浓厚宗教色彩的学派。他们将一切发明都归之于学派的领袖,且秘而不宣,只是当保密的教条被放弃后,才出现一些公开讲述这个学派教义的著作。后世学者对毕达哥拉斯作了大量研究,他的思想和学说才逐渐为人们所知。他宣称数是宇宙万物的本源,企图用抽象的数来解释一切,通过研究揭露数的奥秘来探究宇宙的真理。他以发现“毕达哥拉斯定理”(即中国勾股定理)而著称于世。他非常注意数与图形、算术与几何的紧密联系,发现了完全数、亲和数、形数(三角、五角、六角等数)、毕达哥拉斯数。他阐明了单弦的乐音与弦长的关系,创造了一整套音乐理论,成为音乐理论的鼻祖。他的学派还发现了五种正多面体(四、六、八、十二、二十面体),并证明正多面体只限于这五种。通过勾股定理,他们发现不可通约量,后来由无理量的发现而引起的数学史上所谓第一次危机,导致了数域的扩大。他们首创地圆说,认为日、月、五星都是球体,浮悬在太空中。毕达哥拉斯死后,这个学派还继续存在两个世纪之久。他们的思想和学说,对希腊文化有巨大影响。

【欧多克索斯】(约前 400—前 347) 古希腊数学家、天文学家。生

于尼多斯(今土耳其西南角)。曾向阿尔希塔斯学习几何,向柏拉图学习哲学,对数论和音乐也很有兴趣,并学过医学。后游历埃及和小亚细亚,钻研了天文学,在基齐库斯(今马尔马拉海南岸)创立了一个学派。后期定居故乡尼多斯,从事讲学和著作,直至逝世。他是古希腊最杰出的学者之一。在数学方面,他的地位仅次于阿基米德。他的成果被欧几里得作为《原本》中 V VI XII 卷的主要基础内容。他创立了比例论,开创了运用分理法作出非常严密的实数定义的先河,处理了不可通约量即无理量问题,推动了数论和几何学的发展。他发现的公理及所证明的命题,被公认为是近代极限论的先驱。他完善了穷竭法,被尊为微积分学的创始人之一。他最早研究球面天文和描述星座,倡导同心球理论来解释日月星辰的运动,这些天文学理论对古代和中世纪的欧洲产生了较大影响。他在地理、医学、法律等方面也有贡献。他的著作很多,可惜没有一部保存下来。

【欧几里得】(约前 330—前 275) 古希腊数学家。早年在雅典受教育,熟知柏拉图的学说。公元前 300 年左右,受托勒密王(前 364—前 283)之邀,到埃及统治下的亚历山大城工作,长期从事教学、研究和著述。他写过不少数学、天文、光学和音乐方面的著作,而以巨著《原本》最闻名于世。《原本》原有 13 卷,后人又补充 2 卷。这本著作的原稿早已失传,现存的是公元 4 世纪末西翁的修订本和 18 世纪在梵蒂冈图书馆发现的希腊文手抄本。这部西方世界现存最古老的数学著作,为 2000 年来用公理法建立演绎的数学体系树立了最早的典范。德摩根曾说,除了《圣经》,再没有任何一种书像《原本》这样拥有如此众多的读者,被译成如此多种语言。从 1482 年到 19 世纪末,《原本》的各种版本竟用各种语言出了 1000 版以上。中

国最早的中文译本是 1607 年利玛窦、徐光启合译的《几何原本》前 6 卷,1857 年伟烈亚力、李善兰合译了后 9 卷。

【阿基米德】(约前 287—前 212) 古希腊数学家、力学家。生于西西里岛的叙拉古。早年曾在亚历山大受教于欧几里得的学生,并一直同亚历山大的学者保持密切联系。在当地著述、教学,最后死于入侵的罗马士兵手中。他最著名的成就之一是发现了流体静力学的基本原理——物体在流体中减轻的重量等于排去液体的重量。这一原理总结在他的名著《论浮体》中,以“阿基米德原理”著称于世。他的《阿基米德方法》遗作中所表述的思想方法已具有近代积分论的思想,完全可以誉之为近代积分学的先驱。他将熟练的计算技巧和严格证明融为一体,将抽象的理论和工程技术的具体应用紧密结合,获得许多重要发明。他有许多重要的学术著作,流传下来的还有:《论球与圆柱》、《圆的度量》、《劈锥曲面与旋转椭圆体》、《论螺线》、《抛物线求积》、《平面图形的重心》、《数沙者》等。他在数学方面的研究成果卓著,后人评价极高,常把他和牛顿、高斯并列为有史以来贡献最大的三大数学家。

【阿波罗尼奥斯】(约前 262—前 190) 古希腊数学家、天文学家。生于小亚细亚南岸的珀加。年青时求学于亚历山大的欧几里得的门徒,后来长期在那里教学和研究。其间曾访问过帕加马(今土耳其西海岸)王国,在那里新建的大学和图书馆工作过。他与欧几里得、阿基米德并称为古希腊亚历山大前期的三大数学家。他总结了前人在圆锥曲线论方面的成果,加入自己的杰出创造,撰成巨著《圆锥曲线论》8 卷,将圆锥曲线的性质网罗殆尽,使后人几乎没有插足的余的,在阿拉伯和西欧曾长期被奉为经典之作。直到 17 世

纪的帕斯卡和笛卡儿才在这一研究领域有新的突破。据记载,他还著有《论切触》、《平面轨迹》等许多论著。他还研究过大数问题,几何基础问题,“无序无理数”及求 π 值等问题。在天文学方面,他证明了求简单本轮系统中行星留点的法则,成功地将几何学应用于天文。

【丢番图】希腊数学家。生卒不详,约生活于公元 3 世纪前后。他的活动年代主要是根据 11 世纪一位学者的信件提到丢番图在三世纪中叶的某些学术交往来确定的。另外根据四世纪希腊文选上一首短诗(或墓志铭):“丢番图的一生,幼年占 $1/6$,青少年占 $1/12$,又过了 $1/7$ 方结婚,5 年后得子,子先父 4 年而卒,仅为父寿之半”。由此推知他享年 84 岁。他的著作最重要的是《算术》,序言中记载应用 13 卷,但现存仅前 6 卷希腊文手抄本,及后来发现的 4 卷阿拉伯文译本。这本著作主要讲数的理论,大部分内容可归入代数学。他的特点是使问题的求解完全脱离了几何形式,在希腊数学中独树一帜,对后来的阿拉伯数学,文艺复兴时期的意大利数学乃至整个欧洲数学提供了发展源泉。他还有一本存有残篇的著作是《多角数》。其他著作完全失传。他在代数方面作出了巨大贡献,有人称他为“代数学之父”。为了纪念他的功劳,现在对于具有整系数的不定方程,当只考虑整数解时,就称这类方程为丢番图方程。类似地还有丢番图逼近、丢番图分析等名称。

【婆罗摩笈多】(约 598—约 665) 印度天文学家、数学家。活动于印度中部印多尔之北乌贾因地区,属乌贾因学派。628 年著成《婆罗摩修正体系》一书,全书 24 章,第 12 章、第 18 章专论数学。其中《算术讲义》研究三角形、四边形、算术运算规则、二次方程等,涉及

到零和负数。《不定方程讲义》研究一阶和二阶不定方程,获得了二阶差分内插公式、有理数勾股数公式及二次不定方程的解等成果。其余各章也都涉及不少数学问题。他在 655 年还完成过另一本天文学著作。他对几何作图、几何计算、代数专业名词、代数符号、代数式运算等方面均有研究和创见。他的著作约于 766 年被人带到阿拉伯曼苏哈里发统治时期的巴格达,在皇室的支持下译成阿拉伯文,对阿拉伯的天文学和数学产生了一定的影响。

【婆什迦罗第二】(1114/1115—约 1185) 印度数学家、天文学家。生于印度南部的比杜尔。长期在古代印度文化中心乌贾因的天文台担负领导工作。他是 12 世纪印度最有成就的数学家和天文学家。他的代表作是用散文形式撰写的《天文系统极致》,其中包括两部重要的数学著作《丽罗娃提》和《根的计算》,分别讲述算术和代数。他的著作中包括了 12 世纪时印度记数法,度、量、衡、面积、货币等各种单位换算,包括自然数、分数、负数在内的八种基本运算,方程、不定方程、线性方程组的一些解法、面积、体积等一些几何计算,甚至有三角函数值的精确表达式等。从总体内容看,他的著作代表了 1000—1500 年间印度数学的最高成就,同时在采用缩写文字和一些记号来表示未知数和运算符号方面为代数学的发展做出了重要贡献。他的著作中还记载了许多天文观测,描述了在天文研究中使用的数学方法及天文装置,对天文学也有一定的贡献。公元 600 年左右印度还有一个同名的叫婆什迦罗的数学家,称为婆什迦罗第一。

【花拉子米】(约 780—850) 阿拉伯中世纪数学家,阿拔斯王朝第五代哈里发的司书官,是阿拉伯数学初期的代表人物。他曾从印度

学者的天文表里作了摘要,编辑了阿拉伯最古的天文表,校对了托勒密的天文表。他还编著了阿拉伯国家算术和代数的最古书籍,覆盖了算术、代数、天文、几何、历法等方面。花拉子米第一个用阿拉伯文叙述了十进位制记数法及其运算法则,包括“0”符号及0在十进位置制记数法中的应用及乘法性质。讨论了分数表示和分数开方的情形。花拉子米最伟大的成就是代数学。“代数学”一词就是源于他的数学著作。花拉子米的著作系统地讨论了六种类型的一次及二次方程的解法,给出了几乎所有的根(没考虑零根),其中包括用配平方法来解的方法,给出了正根,类似于一元二次方程求根公式,也指明了今天称为“判别式”的必须非负。每种解法详尽而系统,他还对这几种方程给出几何证明。最后讨论一般形式的方程,认为通过“还原”和“对消”两种变形后,所有一次、二次方程均能化为六种基本型方程之一。这两种变换形成现在解方程的两种基本变形:移项与合并同类项,从此方程解法成为代数学的方向,也是代数的基本特征。因此,花拉子米获得了“代数学之父”的称号。

【纳速拉丁·图西】(1201—1274) 中世纪阿拉伯天文学家、数学家、哲学家。出生于波斯的图斯(今伊朗东部霍腊散省)。曾在图斯学习,后游学于内沙不尔、巴格达等地,是颇有声誉的学者。他曾供职于纳赛尔王朝(伊朗东部),著《纳赛尔的道德》。蒙古旭烈兀汗攻灭纳赛尔王朝后,他受到蒙古统治者的重用,于1259年,在都城马腊格主持建造了一座规模宏大的天文台。一时四方学者云集,使这里成了阿拉伯世界的又一科学中心。他使用当时最先进的仪器进行天文观测,积10余年观测结果,编制了著名的《伊尔汗历》。在数学方面,他的主要成就是探讨了平行公理问题(开欧洲同类问题讨论的先声);撰写了《横截线原理书》,把三角学从天文学的附属地

位中独立出来;撰写了《论四边形》,把平面三角学和球面三角学系统化。他还翻译注释过欧几里得、阿基米德、阿波罗尼奥斯、托勒密等许多著名数学家、天文学家的著作。

【斐波纳契】(约 1170—约 1250) 意大利数学家。12—13 世纪欧洲数学界的代表人物。生于比萨,卒于比萨。早年随经商的父亲到过北非,向一位阿拉伯教师学习计算。后到埃及、叙利亚、希腊、西西里、法国等地游历、学习,熟知不同国家在商业上的算术体系。经比较分析,确信印度—阿拉伯数码体系最方便、适用。在 1200 年前后回到比萨,潜心研究写作。1202 年完成了他的重要著作《算盘书》(1228 年修订)。全书共 15 章,介绍印度—阿拉伯数码及东方诸国(主要是阿拉伯国家)的算术与代数知识。这本著作流传颇广,对把阿拉伯数学介绍到欧洲起了重要作用。他的著作中有关不定方程的解法和中国《孙子算经》中解法一样。他的“兔子问题”导致“斐波纳契数列”的广泛研究,不仅作为初等数学中引人入胜的趣题,而且其相关理论已应用于数论、运筹学,优化理论等许多方面。他的另外几种著作是《几何实用》(1220)、《平方数书》(1225)、《精华》(1225)等。

【卡尔达诺】(1501—1576) 又称卡当。意大利医生、数学家、占星术家。生于帕维亚,卒于罗马。早年学习过古典文学、数学和天文学。后专攻医学,1526 年获医学博士,并在米兰等地行医,闻名全欧。1534 年在行医同时开始任教,讲授数学、天文学、希腊文和辩证术。1539 年加入米兰医学协会。1543 年任帕维亚大学医学教授。1562 年受聘为波伦亚大学数学教授。1571 年定居罗马,直至去世。他是有名的怪异学者,智力超群,但自负轻信,嗜赌博,喜算命,曾

当过教皇的占星术士,也进过监狱。他一生共有 200 多种著述,内容涉及数学、天文、占星术、医学、哲学、音乐等多种学科。他在数学方面造诣极高。1539 年出版两本算术书,是他多年的教学总结。1545 年发表著名的《大术》一书,首次公布了一般三次代数方程的解法,这是他向数学家塔尔塔利亚求教的结果,曾发誓永不泄密,因失信而引起与塔尔塔利亚的一场争论。后来三次方程求根公式从“卡尔达诺公式”流行于世,而塔尔塔利亚却名不见经传。但《大术》中确有卡尔达诺的独特创造,他最早认真讨论了虚数,给出表示虚数的符号和运算法则。此外,他根据多年赌博经验写成《论赌博》一书,系统提出概率计算,形成概率论的早期探索,但发表很晚(1663)。他对代数方程论、流体力学、医学等许多方面都有贡献。

【韦达】(1540—1603) 法国数学家。生于法国东部地区的普瓦图,卒于巴黎。学习法律任律师,曾为议员,数学是其业余爱好。他被誉为 16 世纪最大的代数学家。他研读过许多著名数学家的著作,特别重视研究丢番图等人的在数学中使用符号的思想。他是第一个有意识地、系统地使用字母于数学的人。他写了许多代数学著作,如《分析方法入门》(1591)是最早关于符号代数的著作。他的名著《论方程的识别与订正》(1591 著,1615 出版)是方程论发展中的一个重要标志。他发现了有名的代数方程根与系数的关系——“韦达定理”。他的工作成果为近代代数学的发展奠定了重要基础,推进了方程论的发展。他还准确地预言了未来将出现一种运用符号的关于量的演绎科学。他对三角学、几何学、天文学也有研究,曾出版过三角学著作,设计改进历法等工作,在战争中为政府破译对方密码,赢得很高声誉。

【德扎格】(1593—1662) 法国数学家。射影几何学创建者之一。生于里昂,卒于同地。曾任军事工程师和建筑师。与数学家梅森、笛卡儿等有交往。1636年出版《论透视截线》小册子,开始论及透视问题。1639年出版《试图处理圆锥与平面相交情况初稿》,书中术语怪异,不易理解,当时新兴的解析几何对人们具有更大的吸引力,致使这部重要著作很快被遗忘。直到1845年沙勒偶然发现了这部书的手抄本,才重新引起数学家们的普遍重视,把它列为中世纪纯粹几何学的经典著作。书中引入了无穷元素,讨论极点和极线、透射、透视等问题,为射影几何奠定了坚实的基础。他所发现的“德扎格定理”(两三角形对应顶点连线共点,则对应边交点共线)是全部射影几何的基本定理。

【笛卡儿】(1596—1650) 法国哲学家、数学家、物理学家、生理学家、解析几何学奠基人之一。生于法国土伦,卒于瑞典斯德哥尔摩。出身贵族家庭,襁褓丧母,自幼体弱,早年在学校读书时,校长特许每天早晨在床上读书思考,养成了“晨思”的习惯,一直保持到晚年。1612年到巴黎的普瓦捷大学攻读法学,4年后获博士学位。1618年从军,到过荷兰、丹麦、德国。1621年回国,正值法国内乱,又去荷兰、瑞士、意大利旅行,1625年返巴黎。1628年,移居荷兰,从此得到了较为安静、自由的学术环境,潜心研究哲学数理及天文、物理、化学、生理等许多领域,埋头著述20多年。1649年冬,应邀为瑞典女王克里斯蒂娜(1626—1689)讲课,因生活习惯被破坏,数月后患肺炎逝世。(16年后,遗骨运回巴黎)。他的贡献是多方面的,尤其在哲学及数学方面有独到的见解。如强调科学的目地在于“造福人群”;反对经院哲学,主张“系统的怀疑”方法;提出“我思故我在”的原则,成为他的哲学中第一条原理。他强调使人“成为自然

的主人和统治者”。他主张唯理论,他的数学思想与哲学思想有着极为密切的联系。他把几何学的推理方法或演绎法应用到哲学上,并且明确宣称,科学的本质是数学。他把物质运动的概念作为自然科学的哲学基础之后,就把运动带进了数学,在数学和其他自然科学里就有了辩证法。有许多重要著作,其中《方法论》(1637)是一部文学和哲学的经典著作,后面有三篇著名的附录:《折光学》、《论大气现象》、《几何学》。《几何学》是他唯一的数学著作,却确立了他在数学史上的崇高地位。他明确表述了解析几何的思想,标志着解析几何学的诞生。他还对微积分的创立起到了重要的推动作用。

【费马】(1601—1665) 法国数学家。生于法国南部博蒙—德洛马涅,卒于卡斯特尔。父经商,他自幼有良好的学习条件。在太学习法律,以律师为业,30岁时任图卢兹议会议员,直到去世。他博览群书,精通数国文字,掌握多门自然科学知识,特别热爱古典文学。数学只是业余爱好,但在数论、解析几何、概率论等方面都有重大贡献,被誉为“业余数学家之王”。他性情淡薄,为人谦逊敦厚,公正廉明,生前不愿发表作品。去世后,他的很多论述遗留在旧纸堆中,书页空白处或在给朋友的书信里。他的儿子将这些内容汇集整理,编成《数学论集》两卷,在图卢兹出版(1679)。他对于近代数论的研究在欧拉之前几无人可与之匹敌。著名的“费马大定理”(不可能有满足 $X^n + y^n = z^n$, $n > 2$ 的正整数 X, y, z, n 存在)激起后来历代数学家的兴趣,而至今尚未得到普遍证明。他独立于笛卡儿发现了解析几何的基本原理。由于提出了求曲线的切线及其极大、极小点的方法而被认为是微分学的创始人之一。他还是17世纪兴起的概率论的开拓者之一。他提出光学的“费马原理”,给后来的变分法研究以极大的启示。

【沃利斯】(1613—1703) 英国数学家。生于肯特郡,早年就读于剑桥大学神学系,业余自修数学。成为当时著名数学家之一。早年研究了一些古代数学家的著作,并翻译出版了一些古希腊数学家的名著。他的《无穷算术》中,计算了相当于某些代数函数和定积分,并得到关于 π 的无穷乘积的表达式:

$$4/\pi = \int_0^1 \sqrt{(1-X^2)} dX$$

$= (1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdots) / (2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots)$ 这是数学史上第一个无穷连乘积的例证,具有重要意义。他还著有《圆锥曲线》、《普通数学或算术大全》、《代数》、《力学,或论几何运动》等书。其中在算术、代数、微积分、几何等方面都有独特的创造,成为 17 世纪英国数学思想的代表人物之一。

【帕斯卡】(1623—1662) 17 世纪法国数学家和物理学家,曾被誉为“阿基米德与牛顿两者间的中间环节”。帕斯卡在学术气氛浓厚的环境中长大,自小便表现出对数学的浓厚兴趣与卓越才华。13 岁时精通欧几里得《原本》,并于同年偶然得到所谓的帕斯卡三角形。他发现该图的巧妙结构,研究了该数阵的性质,写成《三角阵算术》、从而奠定了组合论、数学归纳法的基础。帕斯卡三角形对应于牛顿二项式定理展开式各项系数,但晚于中国北宋贾宪三角约 600 多年。

帕斯卡在射影几何中给出了帕斯卡定理等一系列结论,还与费马共同建立了概率论和组合论的基础。

帕斯卡对摆线(旋轮线)进行了深入研究,由此得到一系列曲线绕轴旋转确定立体重心的结论,从而刺激了莱布尼茨对微积分

的工作。

帕斯卡曾制造出历史上第一台加减法计算器,能进行六位数的计算,为后来计算器的发展奠定了基础。

帕斯卡在物理学、流体静力学等领域也表现出杰出的才能,还是法文散文大师和神学辩论家。只可惜他一生为宗教信仰所累,断断续续地研究数学,又断断续续地回到宗教冥想之中去,苦行僧似的自我折磨使他只活了 39 岁。

【关孝和】(约 1642—1708) 日本数学家。生于群馬县藤冈,卒于江户(今东京)。出身于武士家庭,长期在江户为贵族家府掌管财赋,直到 1706 年退职。他是日本传统数学——和算的奠基人,也是关氏学派的创始人,在日本被尊为算圣。他生前出版的著作仅有《发微算法》(1674),去世后又由其弟子编纂出版了一部遗稿《括要算法》(1712)。他的其他主要著作是在关氏学派内部以抄本形式秘传。他的主要数学成就是:在代数中改进了由中国传入的天元术方法,开创和算独特的笔算代数;使由中国传入的高次数字方程解法为和算家所掌握;建立了行列式概念及其初步理论;发现方程正负根存在的条件及与牛顿迭代法类似的解法等。此外,他在幻方、垛积、“圆理”(径、弧、矢间关系的无穷级数表达式)等方面均有研究。他还写过一些天文历法方面的著作。他的思想由他的弟子建部贤弘(1664—1739)等人继承和应用,对日本数学的发展产生了重要影响。

【牛顿】(1642—1727) 英国数学家、物理学家、天文学家和自然哲学家。生于英格兰林肯郡之南约 13 公里的伍尔索普小村子里,卒于伦敦肯辛顿区,葬于伦敦威斯敏斯特教堂。少年就读于农村小

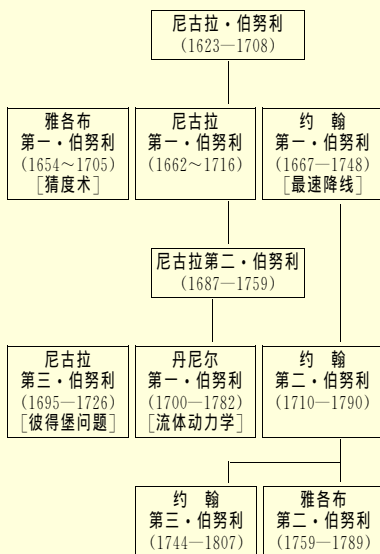
学,并曾务农多年。但他发奋图强,学习刻苦,1661年6月,以优异成绩考入剑桥大学三一学院。1665年获文学士学位。由于伦敦流行鼠疫,波及剑桥,大学临时停办,他返乡两年。在乡间他终日思考各种问题,运用他的智慧和数年的知识积累,来制定科学创造的蓝图。他平生的三大发明(流数术(微积分)、万有引力和光的分析)都发轫于这一时期。1667年回剑桥后当选为三一学院院委,1668年获硕士学位。他多年受教于巴罗教授,此时又协助巴罗编写讲义,撰写微积分和光学的论文,得到巴罗的高度评价。1669年,巴罗坦然宣称牛顿的学识已经超过自己,当年便将“卢卡斯教授”的职位让给牛顿。一时传为佳话。牛顿担任这一职务直到1701年。1696年他任皇家造币厂监督,并移居伦敦,1699年成为厂长。1703年当选为英国皇家学会主席。1705年被女王安妮(1665—1714)封为爵士。他晚年潜心于自然哲学和神学。在数学方面,他的最卓越的贡献是创建微积分。早在1665年他的手稿中就开始有“流数术”的记载。1669年,他写成第一篇微积分论文《运用无穷多项方程的分析》,并交皇家学会备案(1711年出版)。他的正式的流数术著作《流数术方法和无穷级数》于1671年完成。1676年,他又写成他的第3篇重要的微积分论文《曲线求积术》(后来作为《光学》一书的附录发表于1704年)。1687年,他在哈雷的敦促和帮助下发表了巨著《自然哲学的数学原理》。这部著作从作为力学基础的定义和公理(运动定律)出发,将整个力学建立在严谨的数学演绎基础之上,不仅深入地运用了他本人创造的分析工具,也是他的微积分学说的第一次正式公布。他在代数学领域也有一系列重要发现,如 n 次代数方程根的 m 次幂和的公式,实系数方程虚根成对的证明等。此外,他在数论、解析几何、曲线分类、变分法、概率论等分支也有涉及。他在力学、光学、天文学等许多自然科学领域都有巨大贡

献,被奉为最伟大的科学家之一。但他的天才常被夸大到神化的程度,而忽略了他长期刻苦努力的一面。他在晚年信奉上帝。

【莱布尼茨】(1646—1716) 德国数学家、哲学家、自然科学家。生于莱比锡,卒于汉诺威。他自幼丧父,但作为莱比锡大学伦理学教授的父亲,给他留下了丰富的藏书,他的母亲很有学识和远见,从小送他进莱比锡最好的学校学习,使他自幼受到良好的家庭及学校教育。他从小就学习用多种语言表达思想,并表现出超常的哲学天赋。14 岁时对逻辑学产生兴趣,常提出自己的独立见解。1661 年入莱比锡大学学习法律,又曾到耶拿大学学习几何,接触了伽利略、培根、霍布斯、笛卡儿等人的科学和哲学思想。1666 年他在纽伦堡阿尔特多夫大学取得法学博士学位。他当时写出的论文《组合的艺术》已含有数理逻辑的早期思想,后来的一系列工作使他成为数理逻辑的创始人。1667 年后,他投身于外交界,有机会游历欧洲各国,接触数学界名流,尤其是与惠更斯的交往,激起了他对数学的兴趣。他曾制作了一台能作乘法的计算器,是继帕斯卡加法器(1642)之后,计算工具的又一进步。1673 年出访伦敦时,他把这台机器献给了英国皇家学会,还曾送一台复制品给中国的康熙皇帝,可惜目前在故宫已找不到这台机器。1676 年,他到了汉诺威,任公爵处顾问及图书馆馆长。此后 40 年,他常居汉诺威,直到去世。他曾创建勃兰登堡科学协会(后改为柏林科学院),并担任主席。他虽卷入过各种政治斗争,但始终未中断科学研究。他的研究领域极为广泛,涉及到逻辑学、数学、力学、地质学、法学、历史学、语言学、生物学以及外交、神学等方面。他在数学领域最重要的贡献是与牛顿各自独立地创立了微积分学。牛顿建立微积分主要是从运动学的观点出发,而莱布尼茨则从几何学的角度去考虑,特别和巴罗的微

分三角形有密切关系。他的第一篇微分学论文于 1684 年发表在《学艺》杂志上,是世界上最早的微积分文献。他所创造的微积分符号至今仍在高等数学领域中广泛使用。此外,他在组合分析、代数行列式、曲线族的包络等理论方面也都有重要发现。他系统地阐述了二进制记数法,并把它和中国的八卦联系起来。在哲学方面,他倡导客观唯心主义的单子论,并含有辩证法的因素,1714 年写成《单子论》,综述了他的哲学观点。

【伯努利家族】(17—18 世纪) 瑞士巴塞尔的大家族,祖孙三代,仅数学家就有 10 多位。原籍比利时安特卫普,1583 年遭宗教迫害,迁往德国法兰克福,后定居巴塞尔。主要成员及世系见后图。



其中最重要的是雅各布第一·伯努利、约翰第一·伯努利和丹尼尔第一·伯努利。

雅各布第一·伯努利初学神学,后受笛卡儿、沃利斯等人著作启发,转向数学。1687 年任巴塞尔大学教授,直至去世。他是莱布尼茨的朋友,对微积分有深刻了解;最早系统地使用极坐标,对悬

链线、双纽线、对数螺线等有深入研究和成果。巨著《猜度术》(1713)的出版是组合数学及概率论史的一件大事,书中给出的“伯努利数”应用极广,“伯努利定理”则是“大数定律”的最早形式。

约翰第一·伯努利初学医,同时研习数学。他也是莱布尼茨的好友。1691年到巴黎时,曾作为洛必达的私人教师,现在求不定式极限的“洛必达法则”实际是约翰的功绩。1705年,他接替哥哥雅各布任巴塞尔大学教授。他于1696年提出的“最速降线”问题,后来引起变分法的产生。

丹尼尔第一·伯努利生于荷兰格罗宁根,卒于巴塞尔。他和哥哥尼古拉第三·伯努利于1725年同被邀请到彼得堡,尼古拉曾提出著名的概率论方面的“彼得堡问题”,可惜第二年便在彼得堡死去。丹尼尔则受邀成为彼得堡科学院数学教授(25岁)。他最早的论著是解决“吕卡提方程”(1724)。他在概率论、偏微分方程等数学领域均有贡献。在物理学方面,他于1738年出版了《流体动力学》。此外,他在1733年回到巴塞尔后,还教授解剖学、植物学和自然哲学。

【欧拉】(1707—1783) 瑞士数学家。生于瑞士巴塞尔,卒于俄国彼得堡。他生于牧师家庭,父亲欧拉也是一位数学家,父亲希望他学神学,他最感兴趣的却是数学。1720年入巴塞尔大学,受到约翰第一·伯努利的格外赏识,给以特别指导,两年后便获得这所大学的学士学位,此后第二年又获得硕士学位。18岁时,放弃神学而专攻数学。他同约翰·伯努利的两个儿子尼古拉第三·伯努利和丹尼尔·伯努利成为亲密的朋友。1727年,在伯努利两兄弟的推荐下,他受彼得堡科学院的邀请到了俄国。1733年,丹尼尔成为彼得堡科学院的数学教授,时年26岁。28岁时,因工作过度致使右眼失明,年近60岁时,左眼也逐渐失明了。但他从未停止工作。他从19岁起

就开始著作直到 76 岁。他写下许多书籍和论文，以致出版他的全集竟耗时半个世纪。1907 年，瑞士自然科学会就开始为出版《欧拉全集》而筹备整理，计划是 72 卷，从 1911 年起出版，共 3 辑，第一辑数学，共 29 卷，直到 1956 年才出齐。他是 18 世纪数学界最杰出的人物之一，也是数学史上最伟大的数学家之一。至今几乎每一个数学部门，都可以看到欧拉的名字，在几乎所有数学的最重要分支中，都有他的开创性的贡献。他不仅对数学本身做出了巨大贡献，而且把数学用到了几乎整个物理领域。他的研究从数学直到行星运动、刚体运动、热力学、弹道学、人口学，这些工作和他的数学研究相互推动。在他完全失明后的 17 年中，他还口述著作了几本书和 400 篇左右的论文。直到他生命的最后一刻，他还写出计算天王星轨道的要领。他留给后人丰富的科学遗产。其中，分析、代数、数论占 40%，几何占 18%，物理和力学占 28%，天文占 11%，弹道学、航海学、建筑等其他问题占 3%。他被科学史家们列为人类有史以来贡献最大的四位数学家（阿基米德、牛顿、欧拉、高斯）之一。

【拉格朗日】（1736—1813） 法国数学家、力学家、天文学家。生于意大利西北部的都灵，卒于法国巴黎。曾在都灵大学学习。父亲有意让他专攻法律，他却有志于自然科学，先是对古典几何学感兴趣，后又被分析学所吸引，18 岁便在分析学领域有所成就。19 岁时，他被任命为都灵炮兵学校的数学教授，不久又被选为柏林科学院的通讯院士。1757 年，他参与创建了都灵科学协会（都灵科学院前身），他在协会的科技会刊上发表了大量数学、力学论文，内容涉及变分法、概率论、微分方程、振动弦问题、最小作用原理等方面，为这些学科的发展做了开创性的工作。由于他在欧洲所享有的很高声望，1766 年，他受普鲁士王之邀，到柏林科学院工作达 20 年之久。

在这期间，他完成了《分析力学》（1788）一书，这是牛顿之后的一部重要的经典力学著作。他还在代数学的方程论、数论以及分析学、微分方程等数学领域做出了大量重大成果。1786 年，腓特烈（1712—1786，1740—1786 在位）大帝去世后，他认为普鲁士人的排外主义已使科学研究受到阻碍，便接受法王路易十六（1754—1793）的盛情邀请，到巴黎定居，直至去世。在这 27 年中，他曾出任法国米制委员会主任，又先后在巴黎高等师范学院和巴黎综合工科学学校任数学教授。他写出了《解析函数论》（1797）和《函数计算讲义》（1801）两大分析巨著。近百余年来，数学领域的许多新成就可以直接或间接地溯源于拉格朗日的工作。他在数学史上被认为是对分析数学的发展产生全面影响的数学家之一。

【蒙日】（1746—1818） 法国数学家、化学家、机械理论学家。生于法国东部的小镇博讷，卒于巴黎。是法国大革命时期（拿破仑时代）学术界的领导人物。早年就学于法国阿登省梅济耶尔军事学院，1768—1783 年在该院任数学教师。1780 年定居巴黎，1792—1793 年任海军与殖民部长。1794 年成为巴黎理工科大学教授。他参与巴黎综合工科学学校的创办（1795），以后成为该校的领导人直到 1815 年。拿破仑远征埃及时，带领一批专家学者随行，蒙日是其中之一（1798—1801）。他在数学方面的研究涉及微分方程、微分几何、解析几何、画法几何等领域。他开创了用几何来解释和分析运动的方法。他把微积分应用于曲线和曲面的研究，他在三维微分几何方面取得的成就超过了欧拉。由于他和欧拉、拉格朗日的工作，使解析几何成为一个独立的数学分支。法国大革命前后，军事上筑城术的需要促使画法几何产生，蒙日的研究使这一学科独树一帜。只是由于涉及军事机密，他的著作《画法几何》在这一学科创立之后 30 年

才得以出版。他在物理、化学、冶金、机械等方面也有研究和成果。他是一位优秀教育家，在他的直接教导和影响下，一大批几何学家成长起来，并对近代综合几何学产生了重大影响。

【拉普拉斯】（1749—1827） 法国数学家、天文学家。生于法国诺曼底的博蒙一个农民家庭，卒于巴黎。因家境贫寒，靠别人资助入学。曾就读于博蒙军事学校，并担任过该校的数学教员。18岁时到巴黎，经过一番努力和周折，他受到著名科学家达朗贝尔的赏识，被介绍到巴黎陆军学校任数学教授。1785年当选为法国科学院院士。1795年任巴黎综合工科学学校教授。后又在高等师范学校任教授。1816年被选为法兰西学院院士，1817年任该院院长。他的研究领域广泛，涉及天文、数学、物理、化学等方面的许多课题。1768—1778年，他撰著了许多关于积分学、数理天文学、宇宙论、概率论等方面的论文，在学术界初具声望。1778年之后，他在天体力学和概率论方面获得了许多重要研究成果。1796年，他发表了天文学巨著《宇宙体系论》，探讨了太阳系的起源，提出著名的星云假说，因为康德（1724—1804，德国哲学家）也发表过类似的假说，后来就通称为“康德—拉普拉斯星云说”。从1799年起，他陆续发表了5卷16册巨著《天体力学》（1799—1825），汇集了他在天文学上的几乎全部发现。他推广了许多力学定律，将其广泛地应用于运动和天体的研究。书中对位势理论作了重点阐述，对物理学的许多分支的发展有深远的影响。1812年，他出版了《分析概率论》一书，这是概率论方面的一部内容丰富的奠基性著作，总结了以往的研究成果，增加了概率论在一些应用上的新内容，系统叙述了概率论的基本定理，书中大量运用了拉普拉斯变换、生成函数等许多数学工具。此外，他在数学方面还发现了解二次、三次、四次方程的方法；研究了复变

函数的积分和有限差分;推广了函数展开成级数的拉格朗日定理等,对数学发展做出了重大贡献。1805—1827年,是拉普拉斯学派形成的阶段,他们在物理学领域投注了主要精力——毛细管作用、热学理论、粒子论光学、声音的速度等方面的成果,促进了科学的数学化。

【勒让德】(1752—1833) 法国数学家。生于巴黎,卒于巴黎。18岁毕业于马扎林学院。1782年以弹道学研究方面的论文获柏林科学院奖。1783年当选为巴黎科学院助理院士,两年后升为院士,1787年成为伦敦皇家学会会员。1795年当选为法兰西研究院常任院士。1813年继任拉格朗日在天文事务所的职位,直至1833年去世。他与拉格朗日、拉普拉斯被并称为法国数学界的“三L”。他的研究主要涉及数学分析、初等几何、数论及天体力学等方面。他是椭圆积分理论的奠基人之一,发表了《椭圆函数论》(2卷,1827—1832)等大量论著。他的《几何原理》是一部初等几何教科书,书中将几何理论算术化、代数化,说理透彻、简明易懂,在欧洲作为权威教科书达一个世纪之久。他的另一力作《数论》对数论进行了较全面的论述。他还在大地测量理论、球面三角形理论、最小二乘法等方面有重要贡献。此外,他还对高等几何学、力学、天文学、物理学等问题有过论述。

【傅里叶】(1768—1830) 法国数学家。物理学家。生于法国中部奥塞尔城,卒于巴黎。出身于裁缝家庭,9岁父母双亡,沦为孤儿。后被一主教收养,送入地方军事学校读书。虽学习成绩优异,因出身微贱,未能从军,愤而进入修道院。21岁时被召回奥塞尔当数学教师。27岁时任巴黎理工科大学讲师。30岁时随拿破仑远征埃及,

曾当过埃及学院的秘书，并兼管军中外交、政治和科学研究。1801年回法国后受任伊泽尔地区的行政长官。在长达14年的任职期间，荣获功臣与学者的双重桂冠。1808年受封为男爵。他是法国分析学派的公认代表，突出贡献是热传导理论的研究。他给出了热传导的方程及在不同边界条件上的积分法，提出求解这种方程的变量分离法。并在此基础上，将函数表示为由三角函数构成的级数和，并详细讨论了系数构成法则，从而建立起著名的“傅里叶级数”理论，开创了近代数学的一大分支。接着又由系统地运用三角积分而提出了“傅里叶积分”他为求解偏微分方程的边值问题而提供的“傅里叶级数法”，极大地推动了微分方程理论的发展，特别是数学物理等应用数学的发展；其次傅里叶级数拓广了函数概念，极大推动了函数论的研究以及其他纯粹数学领域的研究。他还从事过代数学等其他数学分支的研究。拿破仑倒台后，傅里叶境况不佳，后移居巴黎。1817年任科学院院士。1822年成为科学院终身秘书。他还是英国皇家学会外籍院士（1827）和彼得堡科学院荣誉院士（1829）。他深信数学是解决实际的最卓越的工具，认为“对自然界的深刻研究是数学最富饶的源泉”。

【高斯】（1777—1855） 德国数学家、天文学家、物理学家。历史上伟大的数学家之一。生于不伦瑞克，卒于格廷根。他出身在贫苦的工匠家庭，祖父务农，父做短工，舅舅是个很有才能的纺织工匠，对高斯的知识启蒙有很大影响。高斯从小就显示出很高的数学才能。1792年在不伦瑞克公爵的资助下入卡罗琳学院学习，1795年又入格廷根大学，1799年因证明代数基本定理而获博士学位。从1807年直到1855年去世，一直担任格廷根大学教授兼格廷根天文台台长。他的数学成就遍及各个领域，在数论、代数学、非欧几何、微分几何、

超几何级数、复变函数论、椭圆函数论等方面均有一系列开创性贡献。他一生发表了 155 篇论文，但还有大量创作没有发表出来。他总是等到作品十分成熟时才发表，往往省略分析和思考的过程，以完美无瑕的结果出现。以致使有些学者认为，如能及早发表，其真知灼见对后辈会有更大的启发，进入新的领域也会更快。他十分注重数学的应用。在对天文学、大地测量学、电磁学的研究中，他发明和发展了最小二乘法、曲面论、位势论等。他的《算术研究》是他对数论研究的总结，这本著作奠定了近代数论的基础，是数学史上为数不多的经典著作之一。他在《曲面的一般研究》中，全面阐述了三维空间中的曲面的微分几何，提出了内蕴曲面理论，将微分几何大大推进一步，并决定了这一学科发展的基本方向。他生前关于椭圆函数理论、非欧几何等方面未发表任何公开文章，但从他的遗稿发现，他是这些领域的重要开创者之一。他还发表过关于级数研究、天体运动、磁学及力学方面的论著。

【彭赛列】(1788—1867) 法国数学家、力学家。生于法国梅斯，卒于巴黎。1807—1810 年就读于巴黎理工大学，是蒙日的学生，后又到梅斯炮兵工程兵学院学习并任工程教官。1812 年随拿破仑军队远征俄国，战败被俘，囚禁在萨拉托夫伏尔加河畔的监狱里。1814 年获释回国。在俘虏营的两年中，他回忆、思考了所学过的数学，利用取暖的木炭，在墙上作图，探索几何学问题，创立了射影几何学，他对同囚的技术人员讲解了他的新思想。回国后，1815 年担任过梅斯兵工厂工程师，后任梅斯炮兵工程兵学院教授(1825—1838)，巴黎理学院教授(1838—1848)，巴黎理工大学校长(1848—1858)。1834 年被选为巴黎科学院士，1835 年成为“国防委员会”的成员，1851 年被选为彼得堡科学院通讯院士。1822 年他写成《论图形的射影性

质》一书，是他在俘虏营中的研究成果的总结，这本书奠定了射影几何学的基础。他的工作对 19 世纪的数学有重大影响。1824 年，他关于水轮效率的论文获得巴黎科学院的力学奖金，对力学的研究做出了贡献。

【柯西】（1789—1857） 法国数学家、力学家。生于巴黎，卒于巴黎附近的索镇。他出身于高级官员家庭，父亲曾任法国参议院秘书长，他从小就受到良好的教育。1805 年在巴黎理工大学学习，1807 年入巴黎桥梁公路学院就读。1809 年成为工程师，后在运河、桥梁、海港等工程部门工作。1813 年回到巴黎，任教于巴黎理工大学，1816 年由于他在数学和物理方面的杰出成就而取得教授职位，并被任命为法国科学院院士，以后又担任为巴黎大学理学院和法兰西学院的教授。1830 年，七月革命后，他拒绝宣誓效忠新君主，因此失去了所有的职位。他出走到瑞士的弗里堡，不久被任命为都灵大学的数学物理学教授。1833 年到布拉格，应法国流亡国君的要求，作他私人教师。1838 年回巴黎，仍因拒绝宣誓效忠而再次失去受聘机会。1848 年二月革命，他才又成为理工大学教授。1852 年政变后，恢复帝制及宣誓仪式，但允许柯西可以例外。他的一生著作甚丰，至少出版过 7 部著作和 800 多篇论文。数量之巨，仅次于欧拉。他的成就遍及数学分析、复变函数、误差理论、代数、几何、微分方程、力学、天文学等诸多领域。他的名字联系着许多概念和定理。他在数学方面最重要的贡献在三个领域：微积分学、复变函数、微分方程。他是经典分析的奠基人之一，他为微积分学所奠定的严格基础推动了整个分析学的发展。他是现代复变函数理论的创建人之一，也是他在数学领域最具特色的杰出工作。他在微分方程问题上提出的两个基本问题开创了这一领域研究的新局面。他弄清了弹性理论的基础。

本数学结构，为弹性力学奠定了严格的理论基础。他的著作朴实无华，有思想、有创见。他所发现和创立的定理和公式，往往是一些最简单、最基本的事实。因而，他的数学成就影响广泛，意义深远。

【罗巴切夫斯基】（1792—1856） 俄国数学家。生于俄国下诺夫哥罗德（今高尔基城），卒于喀山。1807年入喀山大学学习，1811年获硕士学位并留校工作。1816年起先后任教授助理、非常任教授，教授。还曾任物理数学系主任（1820—1821，1823—1925）、喀山大学校长（1827—1846）等职。1846年后，任喀山学区副督学，直至去世。他是非欧几何的创始人之一。他在试图证明欧几里得几何第五公设（即平行公设）的过程中，发现了一种逻辑上无矛盾且与“绝对几何”（即与平行公设无关的几何学）不相冲突，但又与欧几里得几何不同的新的几何体系。他称这种几何体系为“虚几何学”。从1826年他在喀山大学公开发表自己的新学说起，直到1855年（逝世前一年），他始终坚持不懈地广泛宣传自己的新的理论发现。可惜在他生前，他的学说未曾受到应有的重视，甚至受到嘲弄和打击。直到他去世后不久，高斯对他的学说予以肯定，他的思想才逐渐被普遍理解和接受。他的学说对几何学和整个数学的发展都起了巨大的作用。他在无穷级数论（特别是三角级数）、积分学、概率论等方面，也做了出色的工作。此外，他还是一位杰出的教育家和管理者。他创立了喀山数学学派和喀山数学教育学派。他曾经凭藉他卓越的组织才能和教育才能把喀山大学从混乱不堪的状态中挽救出来。

【阿贝尔】（1802—1829） 挪威数学家。生于挪威西南的芬岛，卒于弗鲁兰。他家境贫困，父为乡村牧师。13岁入教会学校学习。15岁时得年轻数学教师霍尔姆伯（1795—1850）细心指导，自学许多

数学著作。1821年，在一些教授的资助下，入克里斯蒂安尼亚大学。在学校中，他几乎全是自学，并从事研究，在数学上取得了很大成绩。他一生中得到霍尔姆伯很多帮助。包括资助他到欧洲大陆求职，在他去世后为他校订出版论文集（1839）等。1824年，他用法文写成了“一般五次方程用根式不能求解”的证明论文，但未受到数学家们的重视。1825年，他到德国柏林，认识了克雷尔，这是对他事业有极大帮助的第二个人。他在克雷尔创办的数学刊物上连续发表了许多文章，可惜当时均未受到重视。1826年，他到巴黎，交给法国科学院一篇关于椭圆积分的论文，仍未获得反响。他只好返回柏林，克雷尔为他谋求教授职位，又未成功。1827年，他贫病交迫地回到挪威，靠作家庭教师维生。直到他去世前不久，才因他在《克雷尔杂志》上发表的出色论文开始受到欧洲大陆学术界的注意，人们才认识到他的价值。1828年，四名法国科学院院士上书挪威国王，请为阿贝尔提供合适的科学研究位置。次年4月6日，年仅26岁的阿贝尔就在贫病交困中郁郁去世。柏林大学任命他为数学教授的通知，在他死后两天才到达。荣誉和褒奖在他去世后才接踵而来。1830年他获得法国科学院大奖。他是椭圆函数论的创始人之一。他在交换群、二项级数的严格理论、级数求和等方面都有巨大的贡献。他是近代数学发展的先驱者。可惜他的论文的价值如同他本人一样，没有被社会和学术界及时认识和重用。

【鲍耶】（1802—1860） 匈牙利数学家。生于克劳森堡，在父亲指导下，13岁即掌握微积分学，并应用于力学。16岁进入维也纳皇家工程学院学习，上学期间，即潜心于平行线研究，企图证明欧几里得《原本》的第五公设。毕业后仍然继续研究，并决心创立新的几何学。虽然遭到其父的劝阻，但终于1823年得到了非欧几何的基本

原理。鲍耶把自己的观点，写成《空间的绝对几何学》，请求他父亲协助发表，经过 4 年的努力，于 1832 年才在他父亲一部初等数学书中按“附录”发表了。为了取得权威人士的评论，他父亲把“附录”寄给好友高斯，但没有引起高斯的重视，这无疑是对鲍耶的一次打击；他还致力于复数理论的研究，1873 年，莱比锡征询对虚数理论的文章，他寄去一篇很出色的论文，并预言哈密顿的所谓“三维复数”表示法，这项工作又未得到应有的公允评价，因而使他感到与罗巴切夫斯基的遭遇相似时，他心灰意冷，从此再也没有发表过任何一篇数学学术论文。直到 1894 年，匈牙利数学物理学会才在他的坟墓上竖立起他的石像，并把他的“附录”列入科学经典著作，藉以纪念他对非欧几里得几何学创始的功绩，使之与世长存。

【狄利克雷】（1805—1859） 德国数学家。生于迪伦，早年就读于法兰西学院和巴黎理学院，1826 年回德国，1828 年任柏林大学特别教授，1839 年提升为正式教授，1831 年被选为普鲁士科学院院士。1855 年受聘于格廷根大学，任教授之职。在数论方面，他对高斯的《算术研究》进行了研究，并有所创新，对费马大定理，他给出当 $n=14$ 时，无整数解的证明；还探讨了二次型、多项式的因子、二次和双二次互反律等等问题；还开创了解析数论的研究。在 1863 年出版了他的遗著《数论讲义》，其中包含了他在数论方面的许多成果。在分析方面，他先后发表了《关于三角级数的收敛性》、《用正弦和余弦级数表示完全任意函数》，其中进一步发展了傅里叶级数的理论，并提出新的单值函数概念，还提出所谓“狄利克雷函数”、所谓“狄利克雷积分”等。他还在位势论、热学、磁学、数学物理等方面也有一些创造。

【伽罗瓦】（1811—1832） 法国数学家。在中学时，得到老师沙里的指导，显示出非凡的数学才华，1829年进入巴黎高等师范学校学习，因为支持1830年法国七月革命而被捕入狱，出狱后，由于政治及爱情的原因，1832年卒于决斗，年仅21岁。伽罗瓦在数学上引进了许多新概念，并开辟了数学的新领域——群论。他在生前，曾向巴黎科学院多次投稿，论述他的新发现，但是，不是遭到遗失，便是被认为“不可理解”。他在决斗的前夕，给他的朋友夏瓦利叶写信，请求他把论文公诸于世。直到1846年，即死后14年，伽罗瓦的论文《论方程和根式可解性条件》才发表于《纯粹与应用数学杂志》上。到1870年，一些法国数学家论著发表后，对伽罗瓦的理论作了评论之后，才引起广大数学家的注意，从此，伽罗瓦的工作才完全被理解，他在数学上的地位才得到公认。

【魏尔斯特拉斯】（1815—1897） 德国数学家、分析大师。在柯西、波尔查诺等人工作的基础上，重新考察数学分析中最基本的概念，把分析的基础归结为对实数理论研究，在戴德金、康托尔的努力下，创立了实数理论。魏尔斯特拉斯的分析算术化工作帮助人们摆脱了依赖几何直观的陋习，用不等式刻画极限过程，给出了沿用至今的极限“ ϵ - δ ”定义，他还找到处处不可微的连续函数的例子。这一“病态函数”的发现刺激了对函数的研究，进而产生了实变函数论。他在中学执教了15年，业余钻研数学，年近40才成为职业数学家，他研究数学并运用于教学，许多成果发现于自编讲义或听众笔记中，出版的著作不多，但非常重要。他桃李满天下，莱夫勒出其门下，而第一位女教授：索菲娅·柯瓦列夫斯卡娅（1850—1891）也是他力排陋俗亲自辅导成才的。被人们誉为“现代分析之父”，其光辉业绩使他于1868年被选为法国科学院和柏林科学院的

院士，成为 19 世纪最有影响的分析学家。

【黎曼】(1826—1866) 19 世纪德国数学家。19 岁在哥廷根大学攻读哲学和神学时转攻数学，后来他师从狄利克雷、雅可比、高斯，获博士学位。33 岁时被聘为教授。此时已病魔缠身，39 岁便死于肺结核。他短暂一生发表的作品较少，但高质量的工作却揭开了 19 世纪数学新时期的序幕，对当时最重要的分析、几何与数论领域都做出了划时代贡献，并产生了深远影响。其博士论文《单复变函数的一般理论的基础》是关于复变函数一般理论的名著，确定了复变函数的几何理论基础，蕴含了拓扑学的本质，是近代数学史上重要贡献之一。黎曼在数学分析方面，讨论了连续与可微可积的关系，提出黎曼积分，及其存在的充分必要条件，研究级数收敛与发散性，对整个数学产生了深远影响。在几何基础方面，黎曼对欧氏几何和非欧几何作了纵贯古今的概括，从而统一了非欧几何兴起后的各种几何理论，开辟了微分几何发展的新途径，黎曼几何还为爱因斯坦的广义相对论提供了最合适的数学工具。黎曼在数论、函数论、微分几何、代数几何等方面都有贡献，并写过物理学方面的论文。

【戴德金】(1831—1916) 德国数学家。早年进行代数学、数论的研究，亦曾论及无限思想。这一思想被康托尔所发展，构成了势和序数理论的基础。戴德金是 19 世纪进行数学分析基础严密化工作的数学家，他提出“分划”理论定义了有理数、无理数，进而得到实数，建立起完整的实数理论，标志着数学分析基础的完成。他的代表作是《连续性和无理数》、《数的性质和意义》。他针对当时对有理数的模糊认识，用崭新的方法定义无理数。他发现了直线连续性的本质表现。戴德金应用“分划”定义了有理数和无理数，并将二者

合称为实数，实数的连续性由此得证，数学史上把 1872 年——戴德金发表《连续性与无理数》一书的时间，称为数学分析基础完成的一年。建立了实数概念用其连续性后，戴德金又用分划方法定义实数的算术运算，建立起实数的加法和乘法结合律和交换律等。最后发现：凡有理数的所有运算和各种性质，实数全都满足。戴德金还在《数的性质与意义》中用集合论思想给出了整数理论。

【康托尔】（1845—1918）德国数学家，因集合论被誉为对 20 世纪数学发展影响最远的学者之一。康托尔师从于魏尔斯特拉斯，以数论的工作获得博士学位。他在研究函数表达为三角级数唯一性的判别中提出集合理论，他用有理数的基本序列来定义实数，与戴德金，魏尔斯特拉斯各自从不同角度出发独立地建立了实数理论，他们的工作既把经典分析置于严密的理论基础上，也为近代分析（点集拓扑和泛函分析）的产生和发展开辟了道路。康托尔首先提出了“无穷集合”新概念，第一次把集合作为数学对象，定义集合的交与并，引入“一一对应”原则来衡量两个集合等价和具有相同的基数。认为如果一个集合能跟它的一部分构成一一对应，则该集合是无穷的，从而证明了无穷之间的差别。他的一系列论文创立了以研究集合论的一般性质为对象的数学分支——集合论，成为集合论的奠基人。集合论思想得到了一些数学家的支持与捍卫，也受到另一些数学家的非难与围攻。康托尔历经 20 年艰辛，终于获得了世界公认。人们把 1873 年 12 月 7 日称为集合论的誕生日，集合论促进了许多新的数学科目的产生和发展，在科学界享有崇高的声誉。

【克莱茵】（1849—1925）德国数学家。师从普吕克，研究各种几何学和变换群的关系，他以统一的观点用变换群对几何学作了分类，

即几何学是图形相对于某种变换群的不变性质的研究，被称为“爱尔兰根纲要”（1872），从而把各种几何作了完美的统一。克莱茵总结古希腊几何三大问题，于1895年给出了简单而明晰的证法，从而彻底了结了两千多年的历史悬案。克莱茵数学兴趣极为广泛，对几何、数论、群论、不变量理论、代数、函数论、方程论等方面均有独到的见解。克莱茵另一贡献就是使哥廷根大学成为数学界的朝圣地，他采取一系列措施，网罗和吸引了世界上第一流的数学家，也培养了一大批数学天才人物，20世纪初格廷根所体现出的那种数学研究的风格是一种为所有数学研究部门一直追求的理想，数学研究的真实基础在于理论与实际的恰当联盟之中。克莱茵还是数学教育改革的积极倡导者和热情实施者，他的数学教学的要求受到社会学术界和国家的重视，对现在的数学教育仍有一定的借鉴作用。

【庞加莱】（1860—1934） 法国数学家。他的创新精神和卓越技巧，使他涉猎了纯数学与应用数学的许多领域并做出了重大贡献。他的工作涉及到数学、物理、电学、光学、弹性学、位势理论、动力学、流体力学、相对论和天体力学等。在分析学方面他创立了自守函数的理论。自守函数是周期函数的推广，是在克莱茵群变换下不变的函数，庞加莱应用自守函数代数系统的线性微分方程；建立双曲几何相容性的模型，使自守函数发展成为数学中一个内涵丰富的分支。庞加莱另一贡献是创立了微分方程的定性理论。为研究描述行星运动及行星和卫星轨道稳定性的微分方程有无周期解而开辟了一条崭新的途径；他研究了微分方程的周期性和渐近解，对积分曲线的奇点进行分类等。庞加莱讨论发散级数的本质定义，研究有用的发散级数——称为渐近级数，他把渐近级数理论应用到解微分方程和解天体力学问题中去，并获得了成功。庞加莱是最早系统探讨几何图

形组合理论的人，因而被公认为组合拓扑的奠基者。他提出了代数拓扑的基本概念——同调，在拓扑领域还留下了一些重要的猜想。庞加莱在天体力学、数学物理方面都有许多重要贡献，他还为数学的传播和普及作过不懈的努力。

【希尔伯特】（1862—1943）德国数学家，在代数数域、积分方程、变分法等领域中都取得了辉煌的成果。1899年，他的《几何基础》一书用近代观点建立了欧几里得几何公理化体系，开创了数学公理化的研究方法，并很快推广到数学领域的许多分支，公理化的处理方法被认为是开创了数学研究的新时代。希尔伯特为格廷根学派达到鼎盛时期建立了不朽的功勋。1900年巴黎国际数学家大会上，希尔伯特应邀作大会发言，他在国际数学家面前作了题为“数学问题”的著名演讲，把“问题”看成数学发展的动力，列出了在新世纪里应去努力解决的23个问题，为人们树立了把握全局，抓住关键，推出研究数学主方向问题的范例，他的报告向数学工作者展示了新世纪的前景，为数学史留下了一份珍贵的文献，对20世纪数学发展产生了深刻的影响。希尔伯特是卓越的研究者，也是一位优秀的教师，他具有优良的领导才能，继克莱茵之后把哥廷根数学研究推向全盛。希尔伯特对物理学保持浓厚的兴趣，希望对物理学进行整理，以达到完善、统一的形式和概括、精炼的陈述目的。他坚信物理问题是数学研究的重要实际背景，认为这两门学科的结合是许多成果产生的源泉。他的工作推动了不同领域学科间的学术交流，也促进了数学的发展。

【阿达玛】（1865—1963）法国数学家。他在《关于波的传播的讲义》（1903）中，把特征理论推广到任意阶偏微分方程。并赞同集合

论新理论的作用。1897 年第一次国际数学家大会上指出超限数理论在分析中的重要作用，使得在测度论与拓扑学方面开展了进一步的应用。由于变分法的原因，阿达玛开创了泛函的研究，并做出了成绩。阿达玛对数学基础寄予浓厚的兴趣。对当时数学的逻辑状态提出过批评，赞同康托尔超越数存在的证明，保证所有实数组成集合的存在。阿达玛在行列式方面也有研究成果。他在数论方面应用整函数理论证明了素数定理，研究整函数的目的就在于证明这个素数定理，素数定理是解析数论的中心问题之一。阿达玛是法国 19 世纪函数论王国的代表人物，从一个侧面刺激了布尔巴基学派的形成与发展。

【嘉当，E】（1869—1951） 法国数学家。生于法国伊泽尔。1891 年毕业于巴黎高等师范学校，1894 年以《关于有限维连续变换群的构造》获得博士学位。先后任蒙彼利埃大学、里昂大学、南锡大学、巴黎大学教授等职。嘉当在李群、微分形式、积分不变式、联络几何等许多数学领域内做出了很重要的贡献，例如 1894 年他奠定了李群代数的理论基础，1896 年研究了线性组合代数，1898 年对单李代数进行了完全的分类，20 世纪初期，研究了无限维李群，还研究了群的拓扑性质，并提出一般联络的微分几何概念，他的联络思想对微分几何的发展产生了深远的影响。由于他在许多领域中有卓越的贡献，于 1931 年当选为法国科学院院士，1947 年成为伦敦学会的会员，还获得罗巴切夫斯基国际奖，也得到法国科学院多次奖励。

【罗素】（1872—1970） 英格兰人，20 世纪影响最大的哲学家之一。著述丰富，涉及数学、物理学、教育、社会等领域。早年学欧几里得几何时，对公理，对数学前提表示怀疑，从而决定了他研究的兴

趣——数学基础。他首先完成了《论几何基础》的论文，(1895)以康德哲学精神处理非欧几何，认为这些几何中某些公理是先验的。开始认定数学是逻辑的高度发展形式。1900年接触到皮亚诺符号逻辑体系后，发现这是他寻求多年可以进行逻辑分析的工具，准备把这一工具推广到关系逻辑上。他与数学家兼哲学家怀特海（1861—1947）一起探讨数学基础问题，合作撰写《数学原理》三大卷。这是20世纪初数理逻辑最主要的成就。罗素在本世纪初发现的集合论悖论“罗素悖论”曾予数学界以极大冲击，开始了数学发展史上的第三次危机，促使数学家们重新审视数学的基石——集合理论，这一悖论直到他发展“类型论”才解决。《数学原理》试图把整个数学归结为逻辑，以逻辑的“原始概念”与“原始命题”为出发点，从命题演算开始，通过类和关系的理论，建立自然数系，从而构成全部数学的基础。罗素代表了本世纪初数学基础大讨论中的逻辑主义学派，通过“类型论”排除集合论悖论，刺激了数理逻辑学科的发展，把符号逻辑和古典数学紧密结合起来，也开辟了20世纪以来数学基础和数理逻辑的独特的研究道路。

【勒贝格】（1875—1941）法国数学家，致力于研究各种病态函数，导致了积分学的革命，也创立了实变函数学科。勒贝格潜心研究三角级数以及测度和积分，1902年发表《积分、长度和面积》一文，系统地阐述其测度和积分的思想，他的成名著作是：《论三角级数》（1903），《关于积分法和原函数研究的讲义》（1904）。勒贝格在波莱尔（1871—1956）开闭区间测度思想的影响下，更一般化地引出勒贝格可列可加测度，然后定义一种积分把 $f(X)$ 定义区间 $[a, b]$ 分为若干个勒贝格可测集，作积分和。经典积分下的积分和如果不收敛，用分可测集的方法就可能收敛。于是按黎曼积分不可积的

函数，按勒贝格意义却变得可积了，这是积分学上的一个巨大的突破。勒贝格积分用到三角级数研究上，容易得到许多重要定理，也改进了那时函数可展为三角级数的充分条件。推广了导数概念后，微积分中的牛顿-莱布尼茨公式也有了相应的新结论。实变函数论由此诞生。勒贝格顶住了一些人对他处理病态函数的反对，努力宣传他的思想。到本世纪 30 年代，勒贝格测度论、积分论已成熟，并在概率论，谱理论，泛函分析等方面获得了广泛的应用。

【布劳威尔】（1881—1966） 拓扑学著名学者，荷兰人，20 世纪初数学基础大论战中直觉主义学派继克罗内克、庞加莱之后的代表人物。布劳威尔把数学思维理解为一种创造性程序，数学必须受到数学直觉的限制，即在直观地给定的自然数列上，仅用有限次构造性的方法建立数学。他主张：不承认排中律，不准用反证法证明命题为真。一个命题的存在性被证明，必须被证明可以用有限的步骤构造出来，只证明该条款的不存在导致矛盾是不充分的。排中律只可用于有限集合而不能用于无限集合。是由于逻辑定理滥用于数学的无穷集合导致矛盾、悖论的，而构造法建立的数学可避免悖论的出现。尽管按布劳威尔等人的观点，古典数学中很大一部分内容不能构造，不能称为“数学”，但直觉主义观点受到了许多数学家的重视，对数学基础研究起到了很大的推动作用。布劳威尔本人是通过函数论问题对拓扑学感到兴趣，他在拓扑学中提出了连续变换的单纯逼近法，不动点定理，扩展 Jordan 曲线定理等，从而取得在拓扑学方面的卓越成就。

【诺特】（1882—1935） 德国籍犹太女数学家，在哥廷根大学受业于克莱茵、希尔伯特、闵可夫斯基等人，获爱尔兰根大学数学博士

学位，是当时世界上屈指可数的女数学博士，他开始研究几何图形在满秩线性变换下的不变性，进而在微分不等式方面得到重要结果，被希尔伯特邀请到哥廷根工作。诺特对代数学的环和理想子环的抽象理论作了极为深刻的研究，写成《环和理想论》，成为抽象代数现代理论的开创者之一。她关于环的理论在代数几何中有各种应用，她与其学生，代数学家范德瓦尔登一道，按公理体系奠定了近世代数的现代基础，开拓了数学的新世界，从根本上改变了代数学的面貌，也深深地影响了布尔巴基学派的结构思想。在希特勒驱赶犹太人的淫威之后，诺特被迫离开哥廷根，受聘于布林马尔学院和普林斯顿高等研究所，两年后死于外科手术。爱因斯坦称她为：“自妇女受高等教育以来最重要的富于创造性的数学天才”，韦尔说：“她是一位伟大的数学家……也是历史曾经产生过的最伟大的女性之一”。诺特在代数学领域中的贡献，使她无愧于“抽象代数之母”。

【巴拿赫】（1892—1945） 波兰数学家，泛函分析奠基者。其《抽象集合上算子及其在积分方程上的应用》（1922）是本世纪重要论文之一。1932年他的名著《线性算子理论》标志着泛函分析的成熟。巴拿赫建立起具有不用内积定义的范数空间，给出空间三组公理，这样得到的 Banach 空间比 Hilbert 空间更为一般。他能对其空间证明许多人们熟悉的事实，然后定义算子，讨论关于积分方程抽象形式的解的定理。后来引入 B 空间的伴随空间，伴随算子等。泛函分析后来广泛运用于量子力学等领域。巴拿赫是波兰数学学派的中坚人物，他和史坦因豪斯（1887—1972）一起在里沃夫领导了泛函分析研究中心，与华沙的点集拓扑、集论、数学基础、数理逻辑研究互相呼应。里沃夫泛函分析研究中心的工作是波兰数学的骄傲，也是波兰数学学派最后一个黄金时期。二次大战时，巴拿赫在一个预防

伤害病的研究所里喂养虱子，战后不久即死于肺癌。

【柯尔莫哥洛夫】(1903—) 苏联数学家、概率论公理化的创始人。先在函数论方面作了大量研究，如三角数论、测度论、集论、积分论，后在概率论方面成功地证明了大数定律的必要和充分条件，他推广了切比雪夫不等式，提出了柯尔莫哥洛夫不等式，创立了具有可数集状态的马尔可夫链理论。他最著名的工作是在概率论公理化方面，他的工作标志着概率论发展进入了一个新的阶段，为随机过程的研究提供了必要的基础，被载入数学发展的史册。他建立起来的第一流的概率学派为苏联数学赢得了光辉的荣誉。柯尔莫哥洛夫还把概率论和数理统计用到实践中去，统计检验产品质量，创立统计流体力学以解决湍流力学问题，他的研究体现了学科交叉渗透的特点，与近代科学技术发展相适应。

【冯·诺伊曼】(1903—1957) 美国数学家。生于匈牙利。早年以集合论和数学基础的工作著称，二次大战中参与同反法西斯战争有关的各项科学计划，担任过制造原子弹的顾问。他的科学足迹遍及纯粹数学、应用数学、力学、经济学、气象学、理论物理学、计算机科学及脑科学、他的成就相当于30年科学发展史的概要。他集中研究纯粹数学，涉及到集合论公理系统、元数学、冯·诺伊曼代数算子环等，解决了希尔伯特第五问题，对量子力学加以公理化。1940年他由纯粹数学家转为应用数学家，并应召参与许多重要军事科学计划和工程项目，帮助设计了原子弹的最佳结构，研究空气动力学，转向航空技术。二战后期，他开始计算机研究，在电子计算机逻辑体制中引入代码，编制各种程序，把崭新的科学思想付诸实践，是第一台电子计算机 *ENIAC* 诞生的催生师。现代计算机许多基本设

计中都带有他的思想标记。冯·诺伊曼还创立了对策论，抛弃传统的经典力学方法处理经济问题，而代之以新颖的策略思想和组合工具。晚年则致力于自动机理论，意识到计算机和人脑机制的某种类似，为人工智能研究打下了基础。

【嘉当，H】（1904— ） 当代法国数学家。法国数学家 E. 嘉当之子。生于法国的南锡，1926 年毕业于高等师范学校，1928 年获得博士学位，1936 年任斯特拉思堡大学教授，1940 年任巴黎大学教授，并成为法国布尔巴基学派代表人物之一。研究工作涉及数学的很多分支，例如多复变函数论、自守函数论、拓扑学、同调代数、李群、代数几何、数理逻辑学等。他的著作有《同调代数》（1956）、《微分学、微分形式》（1967）等，1980 年荣获沃尔夫奖。1985 年曾应邀到上海来参加中国数学会成立 50 周年年会，并作了题为《皮尔—费马方程》的学术报告。

【哥德尔】（1906—1978） 奥地利逻辑学家。早年攻读哲学，后转攻物理和数学。由于偏爱精密性和严格性，对他后来的发展有重大意义。集合论悖论出现后的数学界试图对集合论的使用加以限制；罗素提出类型论区分集合，策梅罗和弗兰克尔提出 ZF 选择公理系统，对集合论公理化，试图在不损害数的前提下解决危机。就在希尔伯特自信数学中没有不可知时，哥德尔提出了他的第一、第二不完全定理，并郑重宣告：数学家们企图在某种限制之后，期望一个广阔又严格基础之上放心研究的想法是不可能实现的。他的论文引起了数学家们的极大兴趣，说是把整个数学形式化的主张是注定要失败的，数学中一定有不可知，不可证，数学真理大河中有绝对的不可知，而不是受认识能力，知识限制的不可知。逻辑主义和形式主义

的原则不能贯彻到底，从此，数学家只使用集合论的一些最简单的结果，而把更深入的数理逻辑和数学基础问题留给了专门的数理逻辑学家。

【图灵】（1912—1954） 英国数学家。早年兴趣集中在“可计算数”上，他的理论奠定了计算机科学理论的基础。二次大战时，图灵奉召到英国外交部通讯部所属的密码学校从事破译工作，他领导的数学家、语言学家和计算人员共同研制了一种快速计算机，能高速分析密码——各种可能的组合。图灵的理想计算机的思想导致了世界上第一台数字式专用“巨人”电子计算机的研制成功，也为二次大战的最后胜利建立了不朽功勋。大战结束后，图灵致力于研制大型电子计算机，写出了计算机总体设计方案，包含了仿真系统、子程序和子程序库、错误自检系统、机器自动编译程序等。图灵在机器智能方面做出了许多开创性的工作。并论述了智能机器的可能性，以他特有的理论彻底性对包括智能计算机在内的所有机器作了严密的分类，把数学计算机分为“有组织的”和“无组织的”，两大类。图灵一生的工作覆盖了几个重要领域：数理逻辑、群论、破译密码机、计算机、机器智能，并做出了巨大的贡献，他还对与生命起源有密切关系的“形态发生”的化学理论进行了可贵的探索。他的独创性和预见性愈来愈受到人们的敬佩。

【康托罗维奇】（1912— ） 俄罗斯数学家。生于彼得堡，1930年毕业于列宁格勒大学，1934年任教授，1935年获得博士学位。1964年被选为苏联科学院院士。其数学上的主要贡献有：（1）他是最优化数学方法的创始人之一。1939年，他提出最优化生产计划理论的论点，并发表《生产组织与计划中的数学方法》一书，具有划时代

的意义，为创立线性规划奠定了理论基础。从而形成新兴的经济数学这门交叉学科。(2) 他是计算数学创始人之一。1936 年发表了《高等分析近似方法》，是计算数学奠基著作之一。(3) 在泛函分析方面，他研究了一类半有序空间，先后发表了《半序空间中的泛函分析》、《赋范空间的泛函分析》，并把泛函分析的理论应用到计算数学中，从而建立了解算子方程的方法。在程序设计、函数论、数学物理、微分方程、积分方程、变分法等许多方面都有所贡献；并于 1949 年获得苏联国家奖，1959 年获得列宁奖，1975 年获得诺贝尔奖。

【埃尔朗】丹麦工程师。其生平事迹不详。研究服务系统中的排队现象的随机规律，称之为排队论。虽然 1907 年第一篇有排队论的论文《电话呼叫数与等待时间》是约翰森所写，但第一篇从数学角度出发、严格论述排队问题的论文，埃尔朗于 1909 年所写，他不但把打电话的排队现象作数学处理，他还以“等待时间”为描述服务系统中的指标，并提出“排队论”的名称。到 20 世纪 60 年代，排队论得到很大的发展，排队论不但成为最优化、控制论的基础，也成为计算机设计的理论基础。

【鲁宾逊】(1918—1974) 美国数学家。生于德国，早年求学于耶路撒冷、巴黎等地，后在耶路撒冷、伦敦、多伦多等地工作。1962 年移居美国，并加入美国国籍。主要从事于数理逻辑、代数学等方面的研究，他的最大的研究成果，是运用数理逻辑的概念与方法，将无穷大、无穷小引进微积分，建立在严格的理论基础之上，从而创立了所谓“非标准分析”，著有《非标准分析》、《模型论和元数学导引》，他的著作对数学产生了一定的影响，虽然非标准分析已应用到

一些学科之中，但在数学界，则对其评价不一。

【算经十书】 中国由汉至唐近千年间出现的 10 部数学著作。唐高宗显庆元年（656），天算家李淳风等人校订并注释了 10 部数学书，做为国子监算学馆的教材，通称“十部算经”，它们是：《周髀算经》、《九章算术》、《海岛算经》、《孙子算经》、《夏侯阳算经》、《张丘建算经》、《缀术》、《五曹算经》、《五经算术》、《缉古算经》。北宋元丰七年（1084）秘书省刊刻数学书时，《夏侯阳算经》与《缀术》均已失传，由于唐中叶的一部实用数学书卷首引录了“夏侯阳曰：夫算之法，约省为善……”一段内容，便被误以为原本而刊刻。南宋鲍澣之翻刻这些数学书时（1213），又加入了他发现的曾在唐初作为国子监算学馆辅助教材的《数术记遗》，于是形成了后人所称的《算经十书》。它们既是除新发现的竹简《算数书》外目前仅存的汉唐时期的数学著作，也基本上代表了这一时期中国数学的发展水平。

《周髀算经》原名《周髀》，约成书于公元前 100 年前后，是一部主张盖天说的天文学著作，其中涉及的数学知识有：相当复杂的整数与分数四则运算，计算等差数列通项及公差的公式，一次内插法，勾股原理一般形式的明确表述，开平方，勾股测量等。三国时（3 世纪）赵爽为之作注，在“勾股圆方图注中”，利用著名的弦图给出了勾股定理简捷而严格的证明，并导出了一元二次方程的求根公式及 20 多条与勾股有关的命题。在“日高图注”中利用出入相补原理证明了传统的重差公式，为中国古代测量学中的这一基本公式奠定了理论基础。

《九章算术》中算经典，大约成书于东汉初期，采用问题集形式，收集了 246 道与生产实践有密切关系的应用问题，每题有问有答有术，共 202 术。按问题的性质与解法，分为方田，粟米、衰分、少

广、商功、均输、盈不足、方程、勾股九章，总结了秦汉数学成就，为中国传统数学的发展树立了光辉的典范。《九章算术》论述了平面图形面积算法，系统而完整的分数算法，粮食交换比率和四项比例算法、配分比例算法，已知方形面积立体体积求边长之算法，几何体体积算法，较复杂的配分比例算法，盈亏算法，线性方程组解法，正负数加减法则，不定方程，勾股原理及应用测量法等。魏晋间刘徽注《九章》在理论上创造许多数学原理，并严加证明，然后论证，应用到各种算法中去，形成中国传统数学的理论体系。

《海岛算经》刘徽著，受《九章》勾股章测望问题启发依据重差术撰写测望九问，在测量技术方面发展两次测量为三次、四次测量。

《孙子算经》三卷，作者不详，约成书于东晋南北朝时期，问题集形式，包含有“鸡兔同笼”，“物不知数”等典型问题，分别是带余除法，一次同余式问题的早期成就。

《张邱建算经》三卷，约成书于五世纪或北魏时代，有涉及一次不定方程组的“百鸡术”及等差数列各种法及理论。

《五曹算经》，北周（6世纪）甄鸾撰，分为田曹、兵曹、集曹、仓曹、金曹五卷，是为地方官吏编写的实用算术书，内容浅显。

《五经算术》二卷，甄鸾撰，解释儒家经典及其古注中的数字计算，内容浅显。

《数术记遗》一卷，卷首题“汉徐岳撰，北周汉中郡守、前司隶，臣甄鸾注”，但原文极为简略，必须依赖甄鸾所作的详细注释才能理解，其中又出现了一些时代较晚的典故和用语，许多学者认为它实际上是甄鸾自撰自注的，但也有学者认为确是徐岳所作，迄今尚无定论。书中较突出的内容是记载了中国古代的大数进位制度、记数方法与计算器械，其中“珠算”被认为与后世的算盘有某种渊源关系，“九宫算”明确记载了三阶幻方，在“计数”条下还仿照《张邱

建算经》编制了两个百鸡问题。

《缉古算经》唐初王孝通著，依据几何图形和性质建立高次方程来求解土木建筑问题，并记载有乘法公式及复杂的求积公式。

唐初作为国子监算学馆教材的《夏侯阳算经》本是四五世纪时的一部数学书，已经失传，现传本是中唐时的一部实用数学书，作者韩延，约成书于 770 年前后，全书 3 卷，卷首保存了原本《夏侯阳算经》的部分内容，约 600 字左右，包括对筹算记数制度、乘除算法及分数算法的详细叙述。此外，书中以大量篇幅记载了唐代筹算乘除简便算法的内容。

【九章算术】中国古代影响最深远的一部数学经典，全书 9 卷，共收入 246 个数学问题，多数内容与当时的生产及生活实践密切相关，但也包括一些较为纯粹的数学理论问题，每题大致由问（问题）、答（答案）、术（解题方法或过程）三部分组成，共有 202 术，有的一题一术，有的则是一题多术或多题一术。这些问题与解法意义远远超出了解答具体问题的范围，实际上是以具体问题及其解法为模型，广泛地阐发中国古代数学的独特理论与方法，其首要特点是构造性与程序化，形成了一个较为完整的算法体系。全部问题按其性质和解法分属九章：第一章方田，38 问 21 术，主要论述平面图形（田亩）的面积算法，以及系统而完整的分数算法。第二章粟米，46 问 33 术，主要论述四项比例算法，并给出了 20 种谷物的交换比率。第三章衰分，20 问 22 术，配分比例。第四章少广，24 问 16 术，开平方、开立方，以及由圆面积、球体积反求其直径。第五章商功，28 问 24 术，各种立体图形的体积计算及其在土木工程中的应用。第六章均输，28 问 28 术，较复杂的配分比例。第七章盈不足，20 问 17 术，盈不足术（双假设法）的基本原理并用以解决各种复杂的应用

问题。第八章方程，18问19术，线性方程组解法及正负数运算法。第九章勾股，24问22术，勾股定理，勾股恒等变形，勾股数及测量问题。书中的数学内容按现代的观点可归为以下几类：

(1) 数的概念与运算。自然数的概念与运算在书中被做为熟知的内容。以此为基础，首先在方田章对分数作了系统而完备的处理，包括分数概念，约分、通分、比较分数大小、四则运算、求平均分数等。在约分术中使用了更相减损术，即辗转除法。少广章中有意识地运用了最小公倍数。同时，对在开平方、开立方中遇到的开方不尽数已经给予专门的注意，称这种情形为“不可开”，可以看作对无理根数认识的初步阶段。在方程章，由于使用直除法（即加减消元法）求解线性方程组，不可避免地导致了正负数的加减运算，并给出明确的运算法则，在解题中实际上也已涉及正负数的乘除运算。

(2) 比率算法与代数。率的概念与比率算法是全书的核心内容之一，几乎渗透到全部篇章。书中以粟米、衰分、均输三章专门叙述比率概念与算法，今天所说的正比、反比、复比、连比、配分比例等内容都统一地归结为今有术（四项比例）与衰分术（配分比例）两大基本算法，并从而衍生出各种算法。许多今天看来不属于比例范围的题目也用今有衰分之术处理，有些题目则涉及等差、等比数列。

盈不足术是中国古代数学中求解应用问题的一种别开生面的方法，在《九章》中设有专章。为解决复杂的线性问题及一般非线性问题，通过两次假设试算，将其化为特定的盈亏类数学模型，用固定的演算程序求解，其原理即现在的弦位法。

少广章开方术，开立方术给出了整齐划一、代数意义十分明显的开方程序，其原理可直接用以求解二次及三次方程的正根，并可自然地推广到开高次方及求解一般高次方程的正根。勾股章第20题

则叙述了建立一个一元二次方程的过程。方程章专门阐述线性方程组解法,给出了统一的演算程序,所用直除法即今天的加减消元法。

(3) 面积与体积计算。方田章给出了矩形、等腰三角形、直角梯形、等腰梯形、圆、圆环、环缺等面积的正确公式,又给出了弓形与球冠形面积的两个粗糙的近似公式。商功章给出了多种棱柱、棱锥、棱台、拟柱体以及圆柱、圆锥、圆台体积的正确公式,少广章开立圆术由球体积求其直径,等价于一个粗略的球体积公式。

(4) 勾股及测望理论。勾股章首先明确叙述了勾股定理,然后通过一系列问题阐发了其应用及多种勾股恒等变形。第15、16题给出了在勾股形中所作正方形边长及内切圆直径的正确公式,第14及21题给出了与今天完全一致的整勾股数一般公式。这一章还收入了不少勾股测望问题,是相似勾股形比例性质的应用。

《九章算术》在中国古代数学的发展史上居有至高无上的地位,在世界数学史上也是屈指可数的传世名作之一。首先,它是中国自先秦至汉代数学发展的系统总结,标志着中国初等数学理论体系的建立,这既体现在中国传统数学作为一个构造性、机械化算法体系的确立上,也体现在传统数学的大致分类与基本格局的确立上,其次,它对中国数学的发展产生了极为深远的影响。它通过数学问题阐述数学理论与方法的形式,它对数学内容的取舍、分类与侧重,它的构造性、机械化、形数结合、注重应用的数学思想,都对后世数学的发展与数学著作的编纂产生了明显的规范作用,它所涉及的内容及关心的主题成为后世大量数学成果的源头。第三,它记载了多项在世界上遥遥领先的杰出成就,并对世界数学的发展产生了重要影响。它的系统而完备的分数算法,整齐划一、便于推广的开方程序,体系严整、求解迅速的线性方程组理论,均比其他国家的相近成果早了千年以上,比例算法与正负数运算也早了500年以上,书

中许多著名问题曾传入印度、阿拉伯甚至辗转传入欧洲，其中盈不足术传入欧洲后曾是长期占支配地位的算法。因此，正如许多学者已经指出的，以《九章算术》为代表的机械化算法体系，与以希腊《几何原本》为代表的公理化演绎体系，各擅其长，东西辉映，同为世界数学发展的两大源泉。鉴于《九章算术》在中国数学史上的极端重要性，历史上为之作注的学者为数众多，早期最重要的注释者有魏晋时刘徽、南北朝宋齐间祖冲之父子、唐初李淳风。祖氏父子的注释已失传，而刘徽的《九章算术注》则是中国传统数学中最富理论创见的杰作。

【数书九章】南宋数学家秦九韶数学专著。共 18 卷约 20 万字，列 81 个应用题，分为大衍、天时、田域、测望、赋役、钱谷、营建、军旅和市物 9 类，每类 9 题。《数书九章》在数学内容上多所创新，能道前人所未道。主要有“大衍总数术”——一次同余式组解法和“正负开方术”——高次方程的数值解法，代表了中国乃至世界中世纪数学的最高成就。秦九韶对一次同余式组的解法，比德国数学家高斯于 1801 年发表的同样结果早 500 余年，是中国数学史中数论领域的卓越成就，有人认为“中算之长于西算者，以大衍求一与四元消去法为尤著”。代数方面的高次方程数值解法亦直到 5 个世纪之后才有英国人霍纳的类似的方法，而秦九韶的方法在质量上还胜一筹。另外，《数书九章》还改进了联立一次方程组解法。用互乘对减法，即矩阵法求解。《数书九章》研究勾股测量问题及“三斜求积”题，即已知三角形边之长为 a 、 b 、 c 求面积 A ，秦九韶公式为：

$$A = \sqrt{\frac{1}{4} [a^2b^2 - (\frac{a^2+b^2-c^2}{2})^2]} \text{ 与古希腊海伦公式等价。}$$

【测圆海镜及益古演段】金元时期数学家李冶著，天元术的代表作。李冶（1192—1279）字仁卿，号敬斋。《测圆海镜》（1248）共 12 卷，书前给出了一些定义和 692 条几何命题，收集 170 问，都是勾股形的切圆问题。虽属几何，但李冶是用天元术设定未知数、布列方程、解方程的步骤来处理的，所以实际上是用代数方法解几何题的著作。卷首“圆城图式”，在一个大勾股形中作一内切圆，然后再作一些特殊线，构成 15 个与之相似的勾股形，由此推出各种容圆问题的求法，在几何方面是《九章》勾股容圆问题的发展，在代数上，是对前人工作的引深与提高，为解决设定未知数及布列方程之困难迈出了关键的一步。书中用代数方法解几何问题，利用乘法消去分母，使分式方程化为整式方程，利用乘方消去根号，使根式方程化为整式方程；并作了半符号代数式的尝试。《益古演段》（1259）共收集 64 问，大多数是平面图形之间的面积关系，处理问题的方法与《测圆海镜》相同，亦用天元术解决几何问题，只在筹式排列的次序上、语言文词上与《测圆海镜》略有差异。

【杨辉算法与详解九章算法】杨辉，字谦光，南宋钱塘人，著述甚丰，今流传者有 5 种 21 卷，即《详解九章算法》12 卷（1261 年），《日用算法》2 卷（1262 年），《乘除通变本末》2 卷（1274 年），《田亩比类乘除捷法》2 卷（1275 年），《续古摘奇算法》2 卷（1275），后 3 种世称《杨辉算法》。《详解九章算法》传本已散乱，原序称，此书取《九章》246 问中 80 问进行详解，另添 3 卷，即图 1 卷，乘除算法 1 卷，算类 1 卷。今者仅在盈不足、勾股及算类 3 卷。“详解”体例分 3 项：“解题”即解释题意，名词术语、文字校勘和对题目的译论等内容；“细草”包括图解与算草，又称“图草”；“比类”则选取《九章》之外与原题算法相同或可比附的例题，作对照分析。杨辉的

“算类”是将《九章》全部题目按解题方法由浅入深、重新分为乘除、分率、合率、互换、衰分、叠积、盈不足、方程、勾股等九类。《杨辉算法》中的《乘除通变本末》讨论乘除算法、及加减、求一、九归诸术。《田亩比类乘除捷法》发展了《详解九章算法》方田章内容，题目与例证更切合实际，同时批评《五曹算经》中的一些错误。《续古摘奇算法》则是杨辉搜集“诸家算法奇题及旧刊遗忘之文”编辑成书，其中保存了许多珍贵史料。卷上论纵横图，卷下说《海岛》，都有很高的科学价值。刘益的“正负开方术”，贾宪的“增乘开方法”与“开方作法本原”幸得杨辉引用，方不致失传，因而极其珍贵。杨辉著作除联系实际注重应用和重视数学普及教育强调“乘除”法外，还包含有一些独到的数学成就。杨辉创“纵横图”之名，并在前人的基础上系统化整理了13幅纵横图及各种连环衍发图，略述了某些纵横图的构造方法，说明他对纵横图构成规律已有所发现和概括，他的著述是关于幻方亦即组合数学历史研究的珍贵资料。杨辉还继沈括隙积术之后研究垛积术，对高阶等差级数做了更进一步的探讨。

【算学启蒙和四元玉鉴】元代数学家朱世杰著。朱世杰字汉卿，号松庭，燕山人。著《算学启蒙》3卷（1299），分20门，共259问。又著《四元玉鉴》3卷（1303），24门，共288问。《算学启蒙》内容包括乘除、面积、体积、垛积、盈不足、差分、方程，最后一门是“开方释锁”即天元术，由浅入深，是一部数学入门书。《四元玉鉴》为论垛积术及四元术之杰作，卷上有“四象细草”四问。卷中“茭草形段”，“如象招数”和卷下“果垛迭藏”三门中，集中记述了垛积术和内插法方面的问题，给出了三角垛和四角垛这两个基本的垛积系统以及在这两个垛的基础上产生的岚峰垛系统和四角岚峰

垛。这些组合学方面的成就是惊人的，其中三角垛的性质和贾宪三角的密切关系。书首的“右法七乘方图”比贾宪三角增添了一些数字和连线，使连线上数字规律阐释无遗。书中还利用三角垛结果建立四次内插公式，由书中的二次、三次内插公式看，朱世杰能导出一般的内插公式。同样方法处理级数在西方最早是莱布尼茨（1673）或法国穆顿（1618—1694），但都比朱世杰的成果要晚 300 多年。《四元玉鉴》最伟大的成就是“四元术”，即由“天元术”——一元代数学发展而成的四元代数学，四元高次方程和布列和解法。在前人一元、二元、三元的基础上，推广到天地人物四元，以直观形象给四元高次方程一种固定的表示法，在天元式记法影响下用四个方向代表四元，其中包括有某些难以统一的位置表示。四元方程的解法是依次变形“消”去未知数最后得到一个一元的高次方程而解。朱世杰的“消”法是中国数学史上一项杰出的成就。西方由方程 $f(X \cdot y) = 0$, $g(X \cdot y) = 0$ 消去一个未知量的方法是法国数学家别朱（1730—1783）于 1764 年给出的方案，直到 1779 年方正式发表于《代数方程的一般理论》中，已在朱世杰后 400 余年。《四元玉鉴》使得中国数学在四元术方程及其解法方面达到了高峰，在世界数学史上，也占有一定的地位。

【九章算法比类大全】明代数学家吴敬著，是当时商业数学著作的代表，成书于 1450 年。共分 10 卷，卷首部分列举了大数、小数的记法、度量衡单位及其进率，还有整数分数的四则运算等。正文从第一卷到第九卷，是 1000 多个应用问题解法的汇编，第十卷专述开方。吴敬可能从杨辉的著作中间接了解《九章算术》的内容。认为《九章》包容极广，一切应用问题均出于其中，所以把全部问题按《九章》名义分为方田等九类，编成九卷。各卷先给出摘抄的“古问”，

再结合当时实际附加新的应用问题，称为“比类”。“比类”问题中，包括利息计算、商品交换、就物抽分、合伙经营等大量商业算题，反映出数学在商业资本发展过程中的作用。《九章算法比类大全》中还介绍了一种“写算”乘法，即后来程大位《直指算法统宗》中的“铺地锦”，这是当时印度、阿拉伯、欧洲等地非常流行的一种算法。

【算法统宗】明代以算盘为计算工具的应用数学书，程大位（1533—1606）著。全书共 17 卷，有 595 个应用问题，绝大多数问题是从传本数学书中摘录的。在编写的体例和内容上都受吴敬的《九章算法比类大全》的影响。其特色是所有数学计算都使用珠算。这是明代商业数学发展、珠算法普及其流行时期数学的实用性表现。书中载有珠算口诀，并举例说明如何按口诀在算盘上演算。其中开平方、开立方的珠算方法是程大位首先提出来的。《算法统宗》是一本非常受欢迎的应用数学书，流传极为广泛，明清两代不断翻刻改编，“风行宇内”，学算之人“莫不家藏一册”，影响之大在中国数学史上是罕见的。后来程大位为了普及珠算术，把《算法统宗》一书加以改编，缩写为《算法纂要》四卷，于 1598 年出版。这两部书对朝鲜、日本等邻国数学的影响也是非常深远的。

【几何原本】欧几里得《原本》的汉文译本。16 世纪意大利教士利玛窦来华，与徐光启合译欧几里得《原本》。他们根据德国数学家克拉维斯的拉丁文本欧几里得《原本》（15 卷）翻译了前 6 卷，1607 年出版。前 6 卷包括了欧几里得平面几何的全部内容，其中也夹有克拉维斯收集的对欧氏《原本》的解释及增添的例题。又借用汉语中“几何”一词来表示量，把它看成是研究量的科学，因而定书名为《几何原本》。如今“几何”已成了一门重要数学分支的专有名词。

《几何原本》的内容虽然只是古希腊的几何知识，也不是《原本》博大精深的全部，但对于具有独特传统的中国数学来说是极其令人惊异的。这种从定义、公理出发，按照形式逻辑的要求编排内容的形式，以及演绎推理的证明方法，与中国传统数学是截然不同的。《几何原本》文字通顺，语句流畅，错误极少。徐光启还在翻译中精心研究创设了许多绝妙的译名，不仅在中国沿用到今，而且还影响了朝鲜、日本等国。如：点、直线、曲线、平行线、角、三角形、四边形等。欧几里得《原本》的后半部翻译工作是 250 年后中国数学家李善兰和传教士伟烈亚力合作译出的，出版于 1858 年，至此《原本》方为完璧。

【梅氏历算全书辑要】清代数学家梅文鼎（1633—1721）的数学著述，共 80 多种，200 余卷。康熙壬寅春季（1722）魏荔彤收集已刻及未刻稿编为《梅氏历算全书》，共 29 种 74 卷（包括他人 2 卷），由杨作枚“为之订补疏剔”，于雍正元年（1723）出版。乾隆二十六年（1761），梅文鼎之孙梅穀成以《梅氏历算全书》“校仇编次不善而名为全书亦非实录”，另行编次，更名为《梅氏丛书辑要》，共收 23 种计 60 卷。包括了梅文鼎的主要数学著作，如：《筹算》、《笔算》、《度算释例》、《平三角举要》、《弧三角举要》、《方程论》、《几何补编》、《环中黍尺》、《堑堵测量》、《勾股举隅》、《方圆幂积》11 种，涵盖了初等数学的各个分支。梅文鼎的《笔算》改变《同文算指》笔算中的横写法为竖写法，除法演算步骤较为简捷而乘法比较复杂。书中还包含有定位如小数点的记法。梅文鼎几何方面对立体几何研究做出了贡献，讨论了方灯、圆灯等正多面体的变形和正多面体互容问题，是在《原本》后几卷尚未译出时独立得到的。他还结合天文历法深入研究了三角学，其《平三角举要》、《弧三角举要》是我国

最早的平面三角和球面三角的著作，书中系统而全面地介绍了三角知识。《堑堵测量》及《环中黍尺》对球面三角公式作了补证，创造性地利用投影原理，把球面三角问题转化为平面问题。《方程论》论述多元一次方程组的分类，解法和应用，包含有辅助元或代换法的思想，并讨论了互乘对减法相减次数的最小值。

【则古昔斋算学】清代数学家李善兰的算学文集。李善兰（1811—1882），字壬叔，号秋纫，浙江海宁人。是清代后期数学的代表。数学成就集中于尖锥术、垛积术、数根术等方面，创造了传统数学研究的新水平，同时又能兼收并蓄，成一家之言，在微积分，组合数学、数论和级数等领域很有见地，颇多独创。他的缜密思考和精辟论述影响了整一代人。其数学著作于1867年汇刻出版，书名为《测古昔斋算学》，包括13种：《方圆阐幽》（1845），《弧矢启秘》（1845）、《对数探源》（1845）、《麟德数术解》（1848）、《对数探源》、《垛积比类》、《火器真诀》、《对数尖锥变法解》、《级数回求》、《天算或问》及研究椭圆的三种。《方圆阐幽》概括了李善兰的“尖锥术”方法。尖锥术是在垛积术和传统的极限思想的基础上创立的，相当于微积分的方法，是中算史上微积分的萌芽。《弧矢启秘》结合尖锥术讨论弧矢关系，得到割圆函数级数展开的若干表达式。《对数探源》则由尖锥术给出自然对数的展开式，并以此造表。

《垛积比类》是李善兰对中国数学史上垛积术集大成的著作，其体例清晰、逻辑严明，但所有公式均未给予证明。书中得到相当于费玛组合恒等式，斯特灵数结果。《垛积比类》的研究方法和内容都属于现代组合数学的范围，这是一部现代组合数学出现前的杰作。

【畴人传】阮元主编，由李锐、周治平、钱大昕、凌廷堪、谈泰、焦

循等人集体合作完成的。历时 4 年（1795—1799），共 46 卷。书中记载有从黄帝到嘉庆四年已故天文学家和数学家共 270 余人，其中有数学著作传世者近 50 人。书中还附录了明末以来传入中国的天文、数学书中涉及到的外国天文学家、数学家和来华教士共 41 人。《畴人传》编成后，杨辉、朱世杰等人的著作相继发现，而《畴人传》未列出。所以罗士琳（1789—1853）作《续畴人传》6 卷（1840）。在《续畴人传》中，罗士琳叙述当代学者的生平事迹和学术成就比较详实，列传后的评论也比《畴人传》公允。1886 年，钱塘诸可宝又作《畴人传三编》7 卷，收入 129 人的传记。1889 年，黄钟骏作《畴人传四编》11 卷，续补 275 人，另有西方人 153 人，妇女 8 人。《畴人传》及其续编共收入 922 人的传记，对了解中国古代天文学家和数学家的生平事迹提供了不少方便。但对畴人事迹研究不够，仅是一些史料的堆积，对前人的成就也缺少评价，是数学史研究的早期准备工作。

【莫斯科纸草书】 现存最古老的埃及数学文献之一，大约是公元前 2000—前 1800 年的作品。长 544 厘米，宽 8 厘米，包括 25 个数学问题，1893 年它被一位俄国学者郭列尼舍夫从埃及购得，1912 年被收藏于莫斯科美术博物馆，因而得名。纸草也称莎草，是一种多年生草本植物，盛产于尼罗河三角洲。古埃及人把它的茎从纵面剖开得到一些窄长的草片，由于其中含有粘汁可以互相粘合，压平晒干后就成了可供书写的纸草纸。用这种纸写成的书就被称为纸草书，由于容易毁坏，只有极少数保存至今。莫斯科纸草书中包括了一些简单的数字计算和面积、体积计算，还出现了这样的问题：“某长方形的面积为 12，其宽是长的 $\frac{3}{4}$ ，求其长和宽”。“某直角三角形的一

个直角边是另一个直角边的 $2\frac{1}{2}$ 倍，其面积为 20，求这两边之长”。又如“把一个面积为 100 的正方形分成两个小正方形，使得其中一个的边长是另一个的四分之三”。即相当于 $X^2 + y^2 = 100$, $X : y = 1 : 3/4$ 。其作法是：作一个边长为 1 的正方形，取其边的 $3/4$ 为另一正方形的边，两个正方形的总面积为 $1\frac{9}{16}$ ，求出 $1\frac{9}{16}$ 的平方根为 $\frac{5}{4}$ ，求出 100 的平方根为 10，用 10 除 $\frac{5}{4}$ ，得 8 就是其中一个正方形的边长，乘以 $3/4$ ，即得另一正方形边长为 6，在今天人们也许会认为这是属于代数学范围的二次方程问题，但埃及人解这类问题时所用的方法是他们早已熟知的试位法，其原理是比例，并不真正涉及代数学方法。书中第 10 题计算一个篮子的表面积，对其具体内容的理解有几种意见，一种是把它理解为一个半球形，直径为 $4\frac{1}{2}$ ，其结果与今天的结果完全一致。另一种意见认为篮子的形状是一个半圆柱，直径与高都是 $4\frac{1}{2}$ ，所计算的是这个半圆柱的侧面积，且不包括平面部分，计算过程与前面的相同，恰好也是完全正确的。最引人注目的是第 14 题，其计算过程应用公式 $V = \frac{1}{3}h(a^2 + ab + b^2)$ 计算了一个正四棱台的体积。

在古代，正四棱台体积的计算是相当困难的问题，除了巴比伦人在大致相同的年代曾给出过一个粗糙的近似公式外，世界其他各民族都很晚才得出这一公式（例如在中国，它最早的明确记载是在《九章算术》中），因此，埃及人以对称形式给出的这一结果确实是早期几何计算中一项最出色的成就。他们很可能是在建造金字塔的过程中熟悉它的，因而它被数学史家们一语双关地称为“最伟大的埃及金字塔”。

【阿默斯纸草书】藏于英国伦敦博物馆的古代埃及纸草数学书，由阿默斯抄写故名阿默斯纸草书，亦以购买者的名字命名为莱因特纸草书。苏格兰的埃及学家莱因特于 1858 年，在埃及买到一卷古代埃及的数学纸草书，长 550 厘米，宽 33 厘米，包括 85 个数学问题及其解答，是一位名叫阿默斯的抄写员在公元前 1650 年左右抄写的。阿默斯纸草书记载了丰富的算术与几何知识，包括整数与分数运算、比例、配分比例、等差数列、面积与体积计算等，并涉及大量的食品分配、谷物贮藏、田亩面积、动物饲养等实际问题，可能是当时的一部实用数学手册。古埃及人习惯于使用分子是 1 的单位分数，除 $\frac{2}{3}$ 之外，其他分数均要化为单位分数或它们的和。阿默斯纸书一开头便给出了一张将形如 $\frac{2}{2n+1}$ ($n=2, 3, \dots, 50$) 的分数化为单位分数之和的分解表，利用它可以方便地把分数在 5 与 101 之间的其他分数化为单位分数之和，也可用以进行相应的整数除法运算，一般的除法运算书中第 21、24、25 题是极好的例子：

$$(21) \quad 4 \div 15 = \frac{1}{5} + \frac{1}{15}$$

$$(24) \quad 19 \div 8 = 2 + \frac{1}{4} + \frac{1}{8}$$

$$(25) \quad 16 \div 3 = 5 + \frac{1}{3}$$

古埃及人熟练地运用了比例算法，书中有一类“堆算”问题常常被认为是一元一次方程，实际上书中解这类问题用的“试位法”仍是以比例算法为基础。书中已经熟悉等差数列的简单性质；第 79 题常常被作为埃及人掌握了等比数列的证据。阿默斯纸草书中叙述了如何去计算正方形、长方形、直角或等腰三角形、直角或等腰梯形以及能分割成这些图形的田地面积。第 50 题利用公式 $S = (\frac{8}{9} \cdot d)^2$

计算圆形土地的面积，作者把圆的直径分为 9 等份，并且建立了边长为这样的等份的 8 个正方形，所得正方形被认为与这个圆同样大小，也就是表示圆面积为 $(\frac{8}{9}d)^2$ ， d 为圆的直径。相当于 $\pi=3.1605$ ，误差小于 0.6%。有时他们用 3.2 作为圆周率值。这在当时是一项杰出的成就，第 48 题实际上给出了其计算原理，书中也给出一些体积算法包括圆柱和半球的体积。此外书中还有一些问题讨论了金字塔的陡度，即金字塔斜面上任一点距其底边的水平距离与其高度的比值，实际上就是金字塔的斜面与底面的成二面角的余切。

【原本】欧几里得著。原名为《ΣτοΙΧητα》（希腊文），指基本要素。翻译到中国后定名为《几何原本》。《原本》总结了在此之前数学的发展成就，用严密完整的逻辑体系统率材料。它的基本结构是选取少量原始概念和不需证明的几何命题，作为定义、公理、公设、使之成为全部几何学的逻辑依据，然后运用逻辑推理证明其余的命题，从而得到一系列的几何定理，定理的编排由简到繁，这被称为公理化结构。《原本》共 13 篇，现传本有 15 篇，其最后两篇是在流传过程中由后人补加的。第一篇用 23 个定义提出了点线面圆和平行线的原始概念，接着是 5 个公设和 5 个公理，然后用 48 个命题讨论了关于直线和由直线构成的平面图形的几何学，如三角形、平行线、平行四边形等。第二篇讨论几何代数，第三篇论圆，第四篇论圆的内接外切图形并讨论正多边形，第五篇讨论关于量的比例问题，第六篇是比例理论的几何应用，如相似形等。第七至第九篇为数论内容，第十篇最为艰深，重点是对形如 $\sqrt{\sqrt{a} \pm \sqrt{b}}$ （ a, b 为有理线段）的无理量的所有 25 种可能的形式进行分类，以及毕达哥拉斯数，辗转相

除法等，第十一、十二、十三篇为立体几何和穷竭法，正多面体等。欧几里得《原本》堪称历史上第一部内容丰富，论证精彩的数学典籍，包括了现代中学数学课本的几乎全部内容，而且深深地影响了后世数学的发展，其演绎结构成为西方数学的典范，是人类文化的宝贵财富之一。《原本》成书后出了 1000 多版，印刷量仅次于《圣经》，是当之无愧的科学经典。

【巴克沙利手稿】1881 年在印度西北的巴克沙利村出土了一部用白桦树皮写的残简，数学史称之为“巴克沙利手稿”。据考证，它是公元四五世纪的遗物。巴克沙利手稿中保存有印度数码的发展痕迹，以往的数学记载中用类似“0”符号来表示的零，在巴克沙利手稿中用一个点来表示。

巴克沙利手稿中还记载有：

$$\sqrt{A} = \sqrt{a^2 + r} = a + \frac{r}{2a} \text{ 并 具 体 给 出 } \sqrt{2} = 1 + \frac{1}{3} + \frac{1}{3 \cdot 4} - \frac{1}{3 \cdot 4 \cdot 34},$$

相当于 1.414215686，表明对开方不尽问题有相当好的近似公式。

巴克沙利手稿表现出高超的代数水平，印度数学家解方程 $aX = b$ 时常用的“单假设法”可在此中见到。

对于问题 $aX + c = b$ ，则不用单假设法求解，巴克沙利手稿的方法是先假设一数 g ，按题意算得 $ag + c = b'$ ，然后计算与题设结果之差 $b - b'$ ，那么所求数是 $X = \frac{b - b'}{a} + g$ 。

巴克沙利手稿中还有相当于如下式的线性方程组：

$$\begin{cases} X_1 + X_2 = 16 \\ X_2 + X_3 = 17 \\ X_3 + X_4 = 18 \\ X_4 + X_5 = 19 \\ X_5 + X_1 = 20 \end{cases}$$

其解为 $X_1=9$, $X_2=7$, $X_3=10$, $X_4=8$, $X_5=11$ 。说明当时印度数学家已掌握了特殊类型的线性方程组解法。