考研数学复习宝典(理工类)

毕志伟　叶　鹰　编

华中科技大学出版社

内 容 提 要

本书依据全国硕士研究生入学统一考试数学大纲编写。对大纲中提到的所有概念、结论等知识点，进行了点拨式的归纳和阐述，尤其是对概念的联系和区别进行了整理；将考试要用到的各类公式，集中归纳在相应的章节；对常见的考试题型给出了实例和应对措施。本册为理工类分册。

本书可作为考研复习阶段的同学作为随身手册使用。也可作为低年级本科生学习高等数学的参考书。

前　　言

这是一本为正在刻苦复习的学子们提供的以最新研究生入学考试数学大纲为标准编写的考研数学备考手册。本手册充分考虑到考生的实际情况和临考心态，内容完整而精练，细节丰富而实用，体系科学而简明，查阅快捷而方便。对于正在学习大学数学课程的低年级大学生来说本书也有较大的参考价值。

如果你觉得教科书的内容太多，自己对概念的理解不太到位，定理的使用也有点盲目，公式分散而孤立因而比较难找的话；如果你觉得已经看完的研究生考试复习书还是太厚，题目多而杂，知识的系统性和要领还是不太清楚的话；如果你觉得自己需要整理一下所学的知识，检查复习的完整性，系统地归纳基本公式、常见题型和解题思路方法的话——请使用这本为以上的需求所设计的小册子吧！

本手册每章包括三部分：大纲要求、知识点·点拨、常见题型·应对。

本手册既适合于进入复习阶段总结期的临考冲刺者，也适合于开始复习或准备复习的起跑者。毕竟不走弯路、提高效率是大家的共同目标。

本书由一直在教学一线的资深教师集体策划并编写。近十年来，他们参加了每一年硕士研究生数学

入学考试的评卷和分析工作,担任了华中科技大学数学系的考研辅导教学任务。如果这本凝结了编者多年教学经验的小册子能够提高考生的复习效率的话,那将是编者最期望、最快乐的事情。

编　者

2005 年 5 月

于华中科技大学

目　　录

第1篇　微　积　分

第2篇 线 性 代 数

第3篇　概率论与数理统计

第1篇　微　积　分

第1章　函数、极限、连续

1.1　函　　数

大纲要求

理解函数的概念，掌握函数的表示法，会建立简单的应用问题的函数关系. 了解函数的有界性、单调性、周期性和奇偶性. 理解复合函数及分段函数的概念，了解反函数及隐函数的概念. 掌握基本初等函数的性质及其图形，了解初等函数的概念.

知识点·点拨

【常量·变量】在某个考察过程(时间或空间过程)中保持不变的量称为常量，发生变化的量称为变量. 微积分课程中只考虑实数变量.

【变域】在考察过程中，变量所取得的数值的集合(通常是区间)称为变域.

【邻域】包含点 x_0 的开区间 $(x_0-\delta, x_0+\delta)$ 称为 x_0 的一个邻域，其中 $\delta>0$ 称为该邻域的半径.

【左邻域·右邻域】区间 $(\alpha, x_0]$ $(\alpha<x_0)$ 称为点 x_0 的

左邻域;区间$[x_0,\beta)(x_0<\beta)$称为点x_0的右邻域.

【去心邻域】从点x_0的邻域中去掉点x_0后的集合称为点x_0的去心邻域.

【一元函数】设x与y是两个变量,D是一个非空数集,f是一个联系着x与y的对应规则.如果对每个$x\in D$,依据规则f,总有惟一的数值y与之对应,则称y或f为x的函数,记作$y=f(x)$;称D(有时记作D_f)为此函数的定义域,x与y分别为函数的自变量与因变量.

〖点拨1〗与x对应的y必须是存在且惟一的.

〖点拨2〗定义域及对应规则是函数的两要素.

【值域】设函数$y=f(x)$的定义域是D,则以下数集W称为函数的值域:

$$W=\{f(x)\mid x\in D\}.$$

【函数的相等】两个函数f与g相等的充分必要条件是它们的定义域相等,并且对相同的自变量x,函数值也相等,即$f(x)=g(x)$.以下两个函数

$$y=\sin x\quad(-\infty<x<+\infty),$$

$$x=\sin y\quad(-\infty<y<+\infty)$$

是同一个函数.使用什么字母表示不是重要的.

【自然定义域】对使用数学公式及数学运算表示的函数$f(x)$,称使得$f(x)$有意义的全体实数组成的集合为该函数的自然定义域.如果没有指明$f(x)$的定义域,则默认其定义域为自然定义域.

【函数的奇偶性】设函数$f(x)$的定义域D关于原点对称,则称$f(x)$是:(1)奇函数,若$f(-x)=-f(x)$,

$x\in D$;(2)偶函数,若 $f(-x)=f(x)$,$x\in D$.

【奇偶函数的性质】在函数作图、积分计算等问题中常用到函数的奇偶性,其主要性质如下.

〖几何特点〗奇函数的曲线 $y=f(x)(x\in D_f)$ 关于原点对称,如果 $0\in D_f$,则 $f(0)=0$;偶函数的曲线 $y=f(x)$ 关于 y 轴对称,如果 $f(x)$ 在 $x=0$ 处可导,则 $f'(0)=0$.

例　如果奇函数在 $(-\infty,0)$ 上单调增加及下凸,则它在 $(0,+\infty)$ 上单调增加且上凸;如果偶函数在 $(-\infty,0)$ 上单调增加及下凸,则它在 $(0,+\infty)$ 上单调减少且下凸.

〖四则运算〗函数的奇偶性经过四则运算后的变化如下:

f	g	$f\pm g$	$f\cdot g$ 及 f/g	$f(g(x))$
奇	奇	奇	偶	偶
奇	偶	不确定	奇	偶
偶	偶	偶	偶	偶
偶	奇	不确定	奇	偶

〖导函数〗设函数满足可导条件,则奇函数的导函数是偶函数,偶函数的导函数是奇函数.

〖原函数〗设 $f(x)$ 连续,则奇函数 $f(x)$ 的所有原函数是偶函数,而偶函数 $f(x)$ 的原函数中却只有一个(通过原点的,例如 $\int_0^x f(t)\mathrm{d}t$)是奇函数.

【函数的周期性】设函数 $f(x)$ 的定义域是 D. 若有正常数 T,使得当 $x\in D$ 时 $x+T\in D$ 且 $f(x+T)=f(x)$,则称 $f(x)$ 是周期函数,T 是 $f(x)$ 的一个周期.

〖点拨〗可以验证,$2T,3T,\cdots$ 也是 $f(x)$ 的周期,因此周期函数有无限个周期.

【基本周期】如果 $f(x)$ 有一个最小的周期 T_0,则称 T_0 是 $f(x)$ 的基本周期.

例　$\sin x,\cos x$ 都以 2π 为基本周期;狄利克雷函数没有基本周期,因为每个正有理数都是其周期.

【狄利克雷函数及其性质】以德国数学家狄利克雷名字命名的函数. 用来说明一些重要的函数性质,定义如下:

$$D(x)=\begin{cases}1, & x\text{ 为有理数},\\0, & x\text{ 为无理数}.\end{cases}$$

此函数具有以下基本性质:

(1) $D(x)$ 是分段函数,不是初等函数;

(2) $D(x)$ 是周期函数,以每个正有理数 r 为周期,因为 $D(x)=D(x+r)$;

(3) $D(x)$ 是一个处处有定义的有界函数,也是一个处处不连续的函数;

(4) $D(x)$ 在任何有限区间 $[a,b]$ 上的定积分不存在.

【周期函数的性质】周期函数的性质主要有以下几条.

〖复合运算〗设 T 是 $f(x)$ 的周期,$a>0$. 则 T/a 是复合函数 $\varphi(x)=f(ax+b)$ 的周期.

〖和函数〗当两个周期函数的周期 T_1,T_2 有最小公倍数 T 时，它们的和函数是周期函数，周期为 T. 而当 T_1,T_2 无最小公倍数时，和函数不一定是周期函数. 例如 $f(x)=\sin x$ 以 2π 为周期，狄利克雷函数 $D(x)$ 以正有理数 r 为周期，由于 2π 与 r 无最小公倍数，其和 $f(x)+D(x)$ 不是周期函数.

〖导函数〗可导的周期函数 $f(x)$ 的导函数还是周期函数. 事实上，于 $f(x+T)=f(x)$ 两边同时对 x 求导，得出 $f'(x+T)=f'(x)$. 这说明 $f'(x)$ 以 T 为周期.

〖原函数〗一般地，周期函数 $f(x)$ 的原函数 $F(x)$ 不一定是周期函数，例如，$f(x)=1+\cos x$ 是周期函数，但是其原函数 $F(x)=x+\sin x$ 却不是周期函数. 但是，如果周期函数 $f(x)$ 在周期区间 $[0,T]$ 上的定积分 $\int_0^T f(t)\mathrm{d}t=0$，则它的原函数都是周期函数. 事实上，设 $F(x)$ 是 $f(x)$ 的原函数，令 $h(x)=F(x+T)-F(x)$，则会有 $h'(x)=0$，$h(0)=0$，故 $h(x)=0$.

〖有界性〗在 $(-\infty,+\infty)$ 上连续的周期函数是有界函数.

例如 $f(x)=\sin x^2$ 不是周期函数，因为它的导函数 $f'(x)=2x\cos x^2$ 是 $(-\infty,+\infty)$ 上连续的无界函数，从而不是周期函数，进而推出 $f(x)$ 不是周期函数.

〖定积分〗设 $f(x)$ 是以 T 为周期的在 $(-\infty,+\infty)$ 上连续的周期函数，a 为任意实数，则有

$$\int_0^T f(x)\mathrm{d}x=\int_a^{a+T} f(x)\mathrm{d}x.$$

【函数的单调性】设函数 $f(x)$ 在区间 I 上有定义，$x_1,x_2\in I$.

(1)若 $x_1<x_2$ 时恒有 $f(x_1)\leqslant f(x_2)$(或恒有 $f(x_1)\geqslant f(x_2)$)，则称 $f(x)$ 为 I 上的单调增(或单调减)函数. 单调增函数与单调减函数统称为单调函数.

(2)若 $x_1<x_2$ 时恒有 $f(x_1)<f(x_2)$(或恒有 $f(x_1)>f(x_2)$)，则称 $f(x)$ 为 I 上的严格单调增(或严格单调减)函数. 严格单调增函数与严格单调减函数统称为严格单调函数.

【函数的有界性】设函数 $f(x)$ 在区间 I 上有定义. 若存在常数 B，使得对每个 $x\in I$，有 $f(x)\leqslant B$(或 $f(x)\geqslant B$)，则称 $f(x)$ 在 I 上有上界(或有下界)，且称 B 为 $f(x)$ 在 I 上的一个上界(或下界). 若存在 $M>0$，使得对每个 $x\in I$，$|f(x)|\leqslant M$，则称 $f(x)$ 在 D 上有界；否则称 $f(x)$ 为 I 上的无界函数. 既有上界又有下界的函数是有界函数.

【可逆函数】若自变量取不同值时，对应的函数值也不相同，则称该函数为可逆函数.

〖点拨〗严格单调函数必为可逆函数，但反之不然.

【反函数】当函数 $y=f(x)$ 可逆时，可以定义从函数 $f(x)$ 的值域 W_f 到定义域 D_f 之间的一个函数 f^{-1}：$f^{-1}(y)=x, y\in W_f, f(x)=y$. 称之为函数 $f(x)$ 的反函数.

〖点拨〗通常，从方程 $y=f(x)$ 中解出 $x=f^{-1}(y)$

即得到所求的反函数表示式.

【反函数的图形】在同一个坐标系中函数 $y=f(x)$ 的图形与其反函数 $y=f^{-1}(x)$ 的图形关于直线 $y=x$ 对称,但是与函数 $x=f^{-1}(y)$ 的图形却是相重合的,如图 1-1-1 所示.

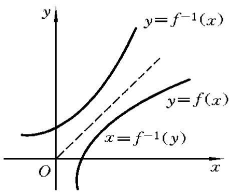


图 1-1-1

【基本初等函数】包括幂函数、指数函数、对数函数、三角函数及反三角函数.

【初等函数】由基本初等函数经有限次四则运算、有限次复合运算构成的用一个数学算式表示的函数.

常见题型 · 应对

历年考卷中较少专门考察函数知识,但是作为微积分的研究对象,函数的基本知识贯穿在许多问题当中,因此必须熟练地掌握本章的知识点.

【求函数的定义域问题】首先依据 $f(x)$ 中出现的基本初等函数性质写出关于自变量的不等式组.例如:对于偶次根式函数如 $\sqrt{u}$ 要求 $u\geqslant 0$;对于对数函数

如 $\ln u$ 要求 $u>0$;对于分式函数如 $\frac{v}{u}$ 要求 $u\neq 0$;对于反三角函数 $\arcsin u$ 或 $\arccos u$ 要求 $|u|\leqslant 1$. 然后对写出的不等式进行求解.

例 函数 $y=\frac{\ln(x+1)}{\sqrt{x-1}}$ 的定义域是( ).

A. $\{x|x>-1\}$　　B. $\{x|x>1\}$

C. $\{x|x\geqslant -1\}$　　D. $\{x|x\geqslant 1\}$

解 首先写出函数中出现的基本初等函数的定义域:$x+1>0$ 以及 $x-1>0$,联立即知正确选项是 B.

【简单函数的复合运算】将函数 $f(x)$ 及 $g(x)$ 按照指定顺序代入即得复合函数 $f(g(x))$ 或 $g(f(x))$.

【复合运算的逆问题】给出函数 $g(x)$ 及复合之后的函数 $f(g(x))$,求 $f(x)$.

〖点拨〗可以采用换元法或凑变量法来求得函数 $f(x)$.

例 设 $f\left(\frac{1}{x+1}\right)=\frac{1+x}{2+x}$,求 $f(x)$.

解1 换元法. 令 $t=\frac{1}{x+1}$ 得 $x=\frac{1}{t}-1$,将其代入条件并化简,得 $f(t)=\frac{1}{1+t}$.

解2 凑变量法. 因

$$f\left(\frac{1}{x+1}\right)=\frac{1+x}{2+x}=\frac{\left(\frac{1}{x+1}\right)^{-1}}{1+\left(\frac{1}{x+1}\right)^{-1}},$$

故

$$f(x)=\frac{x^{-1}}{1+x^{-1}}=\frac{1}{x+1}.$$

【分段函数的复合运算】给出两个函数 $f(x)$，$g(x)$，**其中至少一个**（**通常是** $f(x)$）**是分段函数，求复合函数** $f(g(x))$.

〖点拨〗将函数 $g(x)$ 代入 $f(x)$，注意定义域必须同时作相应的变化.

例 1　设 $f(x)=\begin{cases}1+x, & x\leqslant 0,\\ 4-x, & x>0,\end{cases}$ $g(x)=1-x$，求 $f(g(x))$.

解　$f(g(x))=\begin{cases}1+g(x), & g(x)\leqslant 0\\ 4-g(x), & g(x)>0\end{cases}$

$=\begin{cases}1+1-x, & 1-x\leqslant 0\\ 4-(1-x), & 1-x>0\end{cases}$

（写入 $g(x)$）

$=\begin{cases}2-x, & x\geqslant 1,\\ 3+x, & x<1.\end{cases}$

（整理定义域）

例 2　设 $f(x)=\begin{cases}1+x, & x<0,\\ 1, & x\geqslant 0,\end{cases}$ 求 $f(f(x))$.

解　$f(f(x))=\begin{cases}1+f(x), & f(x)<0\\ 1, & f(x)\geqslant 0\end{cases}$

$=\begin{cases}1+f(x), & x<-1\\ 1, & x\geqslant -1\end{cases}$

$=\begin{cases}2+x, & x<-1,\\ 1, & x\geqslant -1.\end{cases}$

1.2　极　　限

大纲要求

理解极限的概念，理解函数的左极限与右极限

的概念,以及函数极限存在与左、右极限之间的关系.掌握极限的性质及四则运算法则,掌握极限存在的两个准则,并会利用它们求极限,掌握利用两个重要极限求极限的方法.理解无穷小、无穷大的概念,掌握无穷小的比较方法,会用等价无穷小求极限.

知识点·点拨

【数列】一串以自然数为下标排序的实数

$$x_1, x_2, \cdots, x_n, \cdots$$

称为无穷数列或数列,记作$\{x_n\}$或x_n.

【数列的单调性】若$n \geqslant 1$时有$x_{n+1} \geqslant x_n$(或$x_{n+1} \leqslant x_n$),则称$\{x_n\}$为单调增(或单调减)数列;将不等号换作严格不等号时便得到严格单调增(或减)的数列的定义,它们统称为单调数列.

【数列的有界性】若有常数A(或B),使$x_n \geqslant A$(或$x_n \leqslant B$)($n \geqslant 1$)成立,则称A(或B)是$\{x_n\}$的一个下界(或上界).既有上界又有下界的数列称为有界数列.

〖点拨1〗数列若有一个上界,则便有无限个上界;下界情况类似.

〖点拨2〗有界的数列不一定能取得其上界或下界的值.例如,$\left\{\frac{1}{n}\right\}$有下界0,但是数列的每一项都不会等于0.

【无界数列】或者无上界,或者无下界的数列为无界数列.

【收敛数列】设有常数a,使得任给正数ε,都有自然数N,使$|x_n - a| < \varepsilon$($n > N$)成立,则称数列$\{x_n\}$是收敛

数列，a 是其极限. 或说 $\{x_n\}$ 收敛于 a，记作 $\lim\limits_{n\to\infty}x_n=a$.

【发散数列】不收敛于任何实数的数列为发散数列.

例　当 $n\to\infty$ 时，$\sqrt[n]{n}$，$\left(1+\dfrac{1}{n}\right)^n$ 是收敛数列；而 $\sin n$，$\sqrt{n}$，$(-1)^n$ 则是发散数列.

【收敛数列与有界数列】收敛数列必为有界数列，但反之不然，如摆动数列 $(-1)^n$，它是发散的有界数列.

【发散数列与无界数列】无界数列必为发散数列.

【数列极限存在的两个准则】

〖单调有界准则〗单调增加有上界或单调减少有下界的数列为收敛数列.

〖夹逼准则〗设 $a_n\leqslant x_n\leqslant b_n\,(n>1)$，则当数列 a_n 和 b_n 都收敛到同一实数 l 时，数列 x_n 也收敛到 l.

【数列极限四则运算法则】设 $x_n\to A$，$y_n\to B$，则

$$x_n\pm y_n\to A\pm B,\quad x_ny_n\to AB,$$

$$x_n/y_n\to A/B\quad (B\neq 0)$$

【子列】从数列 $\{x_n\}$ 中选出的保持其原来位置次序而编排的无穷数列称为 $\{x_n\}$ 的子列. 如 $\{x_n\}$ 的奇数项子列 $\{x_{2n-1}\}$，偶数项子列 $\{x_{2n}\}$.

【用子列刻画敛散性】$\{x_n\}$ 收敛到 $a\Leftrightarrow$ 它的每个子列也收敛到 a. 于是，当两个子列分别收敛到不同的值，或者有一个子列发散时，数列 $\{x_n\}$ 发散.

【奇偶数列的作用】$\{x_n\}$ 收敛到 $a\Leftrightarrow\{x_n\}$ 的奇数项子列和偶数项子列都收敛到 a.

【收敛数列的保号性】设 $\lim\limits_{n\to\infty}x_n=a$，若 $a>b$，则当 n 足

够大(指大于某个自然数)之后,$x_n>b$.

【收敛数列的比较性质】设数列$\{x_n\}$,$\{y_n\}$都收敛,若$x_n\leqslant y_n$,则$\lim\limits_{n\to\infty}x_n\leqslant\lim\limits_{n\to\infty}y_n$.

〖点拨〗换作严格不等号时上述结论不成立.亦即在$x_n<y_n$情形下,不能保证$\lim\limits_{n\to\infty}x_n<\lim\limits_{n\to\infty}y_n$.例如$x_n=1/n$,$y_n=2/n$.

【两个重要的数列极限】

$$\lim_{n\to\infty}\sqrt[n]{n}=1,\quad \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=\mathrm{e}.$$

【$x\to+\infty$时函数的极限】设有常数l,若对任给正数ε,都有实数X,当$x>X$时,成立不等式:$|f(x)-l|<\varepsilon$,则称l是函数$f(x)$当$x\to+\infty$时的极限,记作

$$\lim_{x\to+\infty}f(x)=l\quad 或\quad f(+\infty)=l.$$

【$x\to-\infty$时函数的极限】设有常数l,若对任给正数ε,都有实数X,当$x<X$时,成立不等式:$|f(x)-l|<\varepsilon$,则称l是函数$f(x)$当$x\to-\infty$时的极限,记作

$$\lim_{x\to-\infty}f(x)=l\quad 或\quad f(-\infty)=l.$$

【$x\to\infty$时函数的极限】设有常数l,若对任给正数ε,都有正数X,当$|x|>X$时,成立不等式:$|f(x)-l|<\varepsilon$,则称l是函数$f(x)$当$x\to\infty$时的极限,记作

$$\lim_{x\to\infty}f(x)=l.$$

【$x\to x_0^+$时函数的极限】设有常数l,若对任给正数ε,都有正数δ,当$0<x-x_0<\delta$时,成立不等式:$|f(x)-l|<\varepsilon$,则称l是函数$f(x)$当$x\to x_0^+$时的极限,记作

$$\lim_{x\to x_0^+} f(x)=l \quad \text{或} \quad f(x_0^+)=l.$$

【$x\to x_0^-$ 时函数的极限】设有常数l,若对任给正数ε,都有正数 δ,当 $0<x_0-x<\delta$ 时,成立不等式:$|f(x)-l|<\varepsilon$,则称 l 是函数 $f(x)$当 $x\to x_0^-$ 时的极限,记作

$$\lim_{x\to x_0^-} f(x)=l \quad \text{或} \quad f(x_0^-)=l.$$

【$x\to x_0$ 时函数的极限】设有常数l,若对任给正数ε,都有正数 δ,当 $0<|x-x_0|<\delta$ 时,成立不等式:$|f(x)-l|<\varepsilon$,则称 l 是函数 $f(x)$当 $x\to x_0$ 时的极限. 记作$\lim\limits_{x\to x_0} f(x)=l$.

【函数的收敛与局部有界】若$\lim\limits_{x\to x_0} f(x)=a$,则函数$f(x)$在 x_0 的某个去心邻域$(x_0-\delta,x_0)\cup(x_0,x_0+\delta)$上有界.

【函数极限的夹逼准则】若在点 x_0 的某个去心邻域内有

$$\alpha(x)\leqslant f(x)\leqslant\beta(x),$$

且

$$\lim_{x\to x_0}\alpha(x)=\lim_{x\to x_0}\beta(x)=l,$$

则

$$\lim_{x\to x_0} f(x)=l.$$

对其他五种极限过程也有类似的结果.

【函数极限的保号性】设$\lim\limits_{x\to x_0} f(x)=l$. 若$l>0$,则在$x_0$的某个去心邻域内有 $f(x)>0$.

例　如果$\lim\limits_{x\to a}\dfrac{f(x)-f(a)}{(x-a)^2}=-1$,**则函数**$f(x)$**在点 $x=a$ 处(　　).**

A. 有非零的导数　　B. 取得极大值

C. 取得极小值　　D. 导数不存在

解　利用保号性,在点 $x=a$ 附近,变量 $\dfrac{f(x)-f(a)}{(x-a)^2}$ 应当为负号,从而其分子恒为负.于是,应当选择B.

【函数极限的比较性质】若在 x_0 的某个去心邻域内成立 $f(x)\geqslant g(x)$,且下面所写极限都存在,则有

$$\lim_{x\to x_0} f(x)\geqslant \lim_{x\to x_0} g(x).$$

【两个重要的函数极限】

$$\lim_{x\to 0}\frac{\sin x}{x}=1,\quad \lim_{x\to 0}(1+x)^{\frac{1}{x}}=\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{x}=\mathrm{e}.$$

【函数极限四则运算法则】设 $\lim f(x)=A$,$\lim g(x)=B$,则

$$\lim[f(x)\pm g(x)]=\lim f(x)\pm\lim g(x)=A\pm B,$$

$$\lim[f(x)g(x)]=\lim f(x)\lim g(x)=AB,$$

$$\lim[f(x)/g(x)]=[\lim f(x)/\lim g(x)]=A/B,\quad B\neq 0.$$

【左、右极限与极限的关系】

$$\lim_{x\to x_0} f(x)=l\Leftrightarrow f(x_0^+)=f(x_0^-)=l,$$

$$\lim_{x\to\infty} f(x)=l\Leftrightarrow f(+\infty)=f(-\infty)=l.$$

【判定函数极限敛散性的方法】

(1) 初等函数在定义区间内每个点的极限都存在且等于该点的函数值;

(2) 分段函数在分段点的极限的存在性要看函数在该点的左、右极限是否相等;

(3) 使用等式变形和极限的四则运算;

(4) 使用不等式变形和夹逼准则.

【无穷小量】在极限过程中趋于零的变量称为(对应于该极限过程中的)无穷小量.

例　$n\to\infty$ 时,$\frac{1}{n}$,$\sqrt{n+1}-\sqrt{n}$ 都是无穷小量;$x\to 0$ 时,x^2,$\sin x$,e^x-1 都是无穷小量.

【基本无穷小量】当 $x\to a$ 时,称 $x-a$ 为基本无穷小量,作为比较的标准,其阶数视为 1.

【无穷小量的比较】设 u,v 是同一极限过程中的无穷小量,则

(1) 当 $\lim\frac{u}{v}=0$ 或者(等价地)$\lim\frac{v}{u}=\infty$ 时,称 u 是 v 的高阶无穷小,或 v 是 u 的低阶无穷小,记作 $u=o(v)$.

(2) 当 $\lim\frac{u}{v}=l$(l 是非零常数)时,称 u 与 v 是同阶的无穷小,记作 $u=O(v)$.

(3) 当 $\lim\frac{u}{v}=1$ 时,称 u 与 v 是等价的无穷小,记作 $u\sim v$.

〖点拨〗等价关系是同阶关系的特例,即 $u\sim v$ 时必有 $u=O(v)$;反过来,若 u 与 v 同阶,如 $\lim\frac{u}{v}=l$ $(l\neq 0)$,则也可得到等价关系式 $u\sim lv$. 等价关系式使用的较多.

【无穷小量的三种变形法】

(1) 等式变形 $u=v$;

(2) 等价变形 $u\sim v$;

(3) **不等式变形** $u\leqslant v$.

〖点拨〗**等式变形能保留变量的所有信息.通常,对可微函数采用泰勒公式变形,对三角函数、根式函数采用初等变形.不等式变形可单向或双向(放大与缩小)地化简变量,但要控制适当不太容易.等价变形则能在保留无穷小量的阶数、主部等主要特征的前提下较大程度地化简变量,且由于等价转换有公式和规则可循,因而在涉及变量极限问题中使用较多.**

【无穷小量的主部和阶】若当 $x\to a$ **时**,$u\sim A(x-a)^r$ $(r>0,A\neq 0)$,**则称** $A(x-a)^r$ **是** u **的主部**,r **是** u **的阶数(或阶).**

例　**当** $x\to 0$ **时**,$1-\cos x\sim\frac{1}{2}x^2$,**故** $1-\cos x$ **的主部是** $\frac{1}{2}x^2$,**阶数是** 2.

【无穷小量运算性质】

(1) **两个无穷小量的和与积仍是无穷小量;**

(2) **有界变量与无穷小量的积是无穷小量.**

例　$n\to\infty$ **时**,$\frac{1}{n}+\frac{1}{\sqrt{n}},\frac{1}{n\ln n},\frac{\sin n}{n}$ **是无穷小量**;$x\to 0$ **时**,$x+\sin x,x\sin x,x\sin\frac{1}{x}$ **是无穷小量**;$x\to+\infty$ **时**,$\frac{1}{x}+e^{-x},\frac{1}{x}e^{-x},\frac{1}{x}\sin x$ **是无穷小量.**

【等价代换与极限计算】设 $u\sim\bar{u}$,**则有**

$$\lim uw=\lim\bar{u}w,\quad \lim w/u=\lim w/\bar{u}.$$

【等价代换与四则运算】设 u,v **是同一极限过程中的无穷小量,它们分别与** $\bar{u},\bar{v}$ **等价,那么:**

(1) $uv\sim\bar{u}\,\bar{v}$, $u/v\sim\bar{u}/\bar{v}$;

(2) 若 $w\to a\neq 0$,则 $wu\sim au$;

(3) 若 $v=o(u)$,则 $u\pm v\sim u$;

(4) 若 $\lim(\bar{u}/\bar{v})\neq -1$,则 $u+v\sim\bar{u}+\bar{v}$.

例　当 $x\to 0$ 时,有

$$\sin x+2\ln(1+x)\sim x+2x=3x,$$

$$\sin x\ln(1+x)\sim x^2,$$

$$\sin 3x+x^2\sim 3x,$$

$$(\cos x+5)\ln(1+x)\sim 6x.$$

注意:$\sin x-\tan x\sim x-x$ 没有意义,上述(4)中的条件不可缺少.

【等价代换与变限积分】若连续函数 $u(x)$,$v(x)$是 $x\to 0$时的等价无穷小量,则有$\int_0^x u(t)\mathrm{d}t\sim\int_0^x v(t)\mathrm{d}t$.

例　当 $x\to 0$ 时,有

$$\int_0^{x^2}\tan\sqrt{t}\,\mathrm{d}t\sim\int_0^{x^2}\sqrt{t}\,\mathrm{d}t=\frac{2}{3}x^3.$$

【无穷大量】在极限过程中绝对值无限增大的变量称为无穷大量,视其最终的符号恒正或者恒负而分作正无穷大量及负无穷大量两类.

【无穷大量与无界量】无穷大量一定是无界量,但反之不然. 例如,当 $x\to\infty$时 $x\sin x$ 是无界量,但不是无穷大量.

【无穷大量运算性质】两个同号无穷大量之和还是无穷大量,两个无穷大量之积也是无穷大量,无穷大量与有界量之和是无穷大量.

【无穷大量与无穷小量】设$u\neq 0$. 若u 是无穷大量,则

$\frac{1}{u}$是无穷小量；若u是无穷小量，则$\frac{1}{u}$是无穷大量.

【洛必达法则】 $\lim\limits_{x}\frac{f(x)}{g(x)}=\lim\limits_{x}\frac{f'(x)}{g'(x)}$，

其中$f(x)$，$g(x)$在同一变化过程（用$\lim\limits_{x}$泛指）中都趋于0或都趋于∞. 并且等式右边的极限为有限值或无穷大.

〖点拨〗仅当函数极限是上述$\frac{0}{0}$型或$\frac{\infty}{\infty}$时，才可使用这一法则，通常在$f'(x)/g'(x)$较$f(x)/g(x)$简单时，洛必达法则奏效.

【如何用好洛必达法则】为了使$f'(x)/g'(x)$的极限易于计算，通常应结合以下方法化简：

(1)使用等价代换化简分子或分母中某个因式；

(2)使用等式变形化简函数形式.

常见题型·应对

【计算数列极限】通过数列极限的计算问题可以考察极限的概念、性质、计算方法的掌握程度. 在数项级数的问题中经常遇到.

〖一般解法〗由于数列的形式多样，使用的解法也很多. 计算数列极限的关键是对数列化简，通常采用等式变形法、等价变形法、放大缩小法等化简手段.

(1)对幂指形式如$x_n=a_n^{b_n}$形式，考虑等式变形$a_n^{b_n}=\mathrm{e}^{b_n\ln a_n}$转换为乘积形式，而对于常考的$1^\infty$型变量$(a_n\to1,b_n\to\infty)$的极限，则应当直接转换为以下极限：

$$\lim x_n=\mathrm{e}^{\lim b_n(a_n-1)}.$$

(2)对根式差形式，如$x_n=n(\sqrt{n^2+1}-n)$，考虑

用有理化法化简；对下面的 n 次方根数列，可以利用通用结果：

$$\lim\sqrt[n]{a^n+b^n}=\max\{a,b\}\quad(a,b>0).$$

(3) 对分式如 $x_n=\dfrac{n^2+n-1}{2n^2-n+1}$，可考虑分式上下除以高阶无穷大量 n^2 来化简.

(4) 对含有参数者，如 $x_n=\dfrac{a^n}{1+a^n}(a>0)$，须分别讨论 a 的各种取值.

(5) 对一般的 n 项和式，如

$$x_n=\frac{n}{n^2+1}+\frac{n}{n^2+2}+\cdots+\frac{n}{n^2+n},$$

可以考虑用夹逼准则.

(6) 对特殊的 n 项和式，如 $x_n=\dfrac{1}{n}\sum\limits_{k=1}^{n}f\left(\dfrac{k}{n}\right)$ 或者 $x_n=\dfrac{1}{n}\sum\limits_{k=1}^{n}f\left(\dfrac{k-1}{n}\right)$，可以化作区间 $[0,1]$ 上的定积分：

$$\lim_{n\to\infty}x_n=\int_0^1 f(x)\mathrm{d}x.$$

(7) 对于特殊的通项 $x_n=f(n)$，可以考虑将数列极限 $\lim\limits_{n\to\infty}x_n$ 变作函数极限 $\lim\limits_{x\to+\infty}f(x)$ 计算，因为后者有较多的化简方法.

(8) 对递归列 $x_{n+1}=f(x_n)$，则要求首先用单调有界收敛准则证明其收敛，然后从 $a=f(a)$ 中解得极限值 $a=\lim\limits_{n\to\infty}x_n$.

【计算函数极限】 函数极限可以作为单独的考点，也可以作为综合问题中的考点. 在使用四则运算极限

规则的基础上,可以考虑以下较典型的计算方法.

〖分侧计算法〗对于分段函数在分段点的极限,含有绝对值或指数函数的初等函数在特殊点的极限,应当分侧处理.例如:

$$\lim_{x\to 0}\left(\frac{2+e^{1/x}}{1+e^{4/x}}+\frac{\sin x}{|x|}\right)(=1),$$

$\lim\limits_{x\to 0}\left(\dfrac{1}{1+e^{1/x}}\right)$(不存在,左、右极限不相等).

〖函数值代入法〗对于初等函数在定义区间内部的点的极限,该点的函数值便是所求的极限值:$\lim\limits_{x\to a}f(x)=f(a)$.其中 $f(x)$是初等函数,a 在 $f(x)$的定义区间中.例如:

$$\lim_{x\to 1}\frac{\sin x}{x}=\sin 1,\quad \lim_{x\to e}\ln(e+x)=\ln 2+1.$$

许多函数极限的最后一步可能归结到这一情况.

〖定性分析法〗根据函数的定义和变量的结构、性质来计算极限.例如:

$\lim\limits_{x\to 0^+}\ln x=-\infty$　(结合函数的图形),

$\lim\limits_{x\to\infty}\dfrac{\sin x}{x}=0$　(有界量 $\sin x$ 与无穷小量$\dfrac{1}{x}$的积还是无穷小量).

〖利用重要极限法〗重要极限公式中的变量x 可以换作任何趋于零的变量$u(x)$.例如:

$$\lim_{x\to 1}\frac{\sin\ln x}{\ln x}=1,\text{因为 } \ln x\to 0,$$

$$\lim_{x\to -\infty}(1+e^x)^{e^{-x}}=e,\text{因为 } e^x\to 0.$$

〖不等式夹逼法〗适当地放大或缩小变量.

〖等式变形法〗通过对函数的合并、因式分解、根式有理化等手段使得函数化简，极限容易计算.

〖换元变形法〗通过新的自变量的引进，化简函数形式. 例如函数极限$l=\lim\limits_{x\to 0^-}\dfrac{\mathrm{e}^{1/x}}{x}$，采用倒代换$t=\dfrac{1}{x}$之后变作$l=\lim\limits_{t\to\infty}\dfrac{t}{\mathrm{e}^t}$，便容易调用洛必达法则计算.

〖等价代换法〗通过函数中的因式的替换能较大程度地化简函数，并且不改变函数极限的存在性和大小.

〖洛必达法则〗以下五种未定式的函数极限可以转换为$\dfrac{0}{0}$型或$\dfrac{\infty}{\infty}$型极限问题，从而调用洛必达法则计算.

(1)$\infty-\infty$型：通分化作分式.

(2)$0\cdot\infty$型：直接化作分式.

(3)1^∞型：取对数化作$0\cdot\infty$型.

(4)0^0型：取对数化作$0\cdot\infty$型.

(5)∞^0型：取对数化作$0\cdot\infty$型.

〖泰勒公式法〗属于等式变形法，信息无损失. 一旦将函数都写成多项式及其余项的和式，则极限计算便非常简单. 例如：

$$\lim_{x\to 0}\frac{\cos x-\mathrm{e}^{-x^2/2}}{x^4}$$

$$=\lim_{x\to 0}\frac{1-\dfrac{1}{2}x^2+\dfrac{1}{4!}x^4-o(x^4)-1+\dfrac{1}{2}x^2-\dfrac{1}{8}x^4+o(x^4)}{x^4}$$

$$=-\frac{1}{12}.$$

〖中值定理法〗属于等式变形法,适合于同名函数差形式,相当于泰勒公式的特殊情况. 思路如下:

$$\lim_{x\to a}\frac{f(x)-f(a)}{g(x)-g(a)}=\lim_{\xi\to a}\frac{f'(\xi)}{g'(\xi)}.$$

【常用的数列极限】当 $n\to\infty$ 时,有

(1) $\left(1+\frac{1}{n}\right)^n\to\mathrm{e}$, $\left(1+\frac{k}{n}\right)^n\to\mathrm{e}^k$;
(2) $\sqrt[n]{a}\to1\ (a>0)$, $\sqrt[n]{n}\to1$;
(3) $a^n\to0\quad(\lvert a\rvert<1)$;
(4) $\frac{a^n}{n!}\to0$, $\frac{n^a}{a^n}\to0\quad(a>1)$.

〖点拨〗将以上公式中的 n 换作随 n 增大而趋于正无穷大的 $\alpha(n)$ 后,结论依然正确. 例如:

$$\left(1+\frac{1}{3n}\right)^{3n}\to\mathrm{e},\quad \sqrt[2n]{2n}\to1.$$

【常用等价无穷小公式】当 $x\to0$ 时,有

$\sin x\sim x$, $\tan x\sim x$,
$\mathrm{e}^x-1\sim x$, $a^x-1\sim x\ln a$,
$\arcsin x\sim x$, $\arctan x\sim x$,
$\ln(1+x)\sim x$, $(1+x)^\alpha-1\sim\alpha x$,
$1-\cos x\sim\frac{1}{2}x^2$, $\tan x-\sin x\sim\frac{1}{2}x^3$.

〖点拨1〗当上述公式中的 x 换作中间变量 $\alpha(x)$ 时,只要 $\alpha(x)\to0$,公式就依旧可用. 例如:

$$\ln x=\ln(1+x-1)\sim x-1\quad(x\to1);$$

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

$$\sim\frac{1}{2}(\cos x-1)\sim-\frac{1}{4}x^2(x\to 0).$$

〖点拨 2〗结合因式分解，可以将等价的无穷小因式和极限非零的因式进行替换. 例如：

当 $x\to 0^+$ 时，有

$$\sqrt{x}-\sqrt{x/(x+1)}=\sqrt{x}\left(1-\frac{1}{\sqrt{x+1}}\right)$$

$$=\sqrt{x}\,(\sqrt{x+1}-1)/\sqrt{x+1}$$

$$\sim\sqrt{x}\cdot\frac{1}{2}x\Big/1=\frac{1}{2}x^{3/2}$$

（其中 $\sqrt{x+1}-1\sim x/2$，$\sqrt{x+1}$ 是极限非零的因式，可以用其极限值来替换）.

1.3　连　　续

大纲要求

理解函数连续性（含左连续和右连续）的概念，会判断函数间断点的类型. 了解连续函数的性质和初等函数的连续性，理解闭区间上连续函数的性质（有界性，最大值和最小值定理，介值定理），并会应用这些性质.

知识点 · 点拨

【在一点连续】设函数 $f(x)$ 在 x_0 的某个邻域有定义. 若 $\lim\limits_{x\to x_0}f(x)=f(x_0)$，则称 $f(x)$ 在 x_0 处连续.

【在一点左（右）连续】设函数 $f(x)$ 在 x_0 的某个左（右）邻域有定义，若 $\lim\limits_{x\to x_0^-}f(x)=f(x_0)$（$\lim\limits_{x\to x_0^+}f(x)=f(x_0)$），则称 $f(x)$ 在点 x_0 左（右）连续.

【在一点连续的充要条件】

函数 $f(x)$ 在点 x_0 处连续

$\Leftrightarrow$ 函数 $f(x)$ 在点 x_0 既是左连续又是右连续

$\Leftrightarrow \lim\limits_{x\to x_0}(f(x)-f(x_0))=0$

$\Leftrightarrow f(x_0^-)=f(x_0^+)=f(x_0)$.

【在开区间连续】 若 $f(x)$ 在每个点 $x_0\in(a,b)$ 都连续,称函数 $f(x)$ 在开区间 (a,b) 连续.

【在闭区间连续】 若 $f(x)$ 在点 a 右连续,在点 b 左连续,且在开区间 (a,b) 连续,称函数 $f(x)$ 在闭区间 $[a,b]$ 连续.

【间断点】 使函数 $f(x)$ 不连续的点称为间断点.

〖点拨〗 通常,被考察的点的某个单侧或双侧去心邻域上函数 $f(x)$ 应当有定义.

【间断点分类】 函数 $f(x)$ 的间断点分为第一类(可去间断点和跳跃间断点)以及第二类间断点(无穷间断点和振荡间断点).

【第一类间断点】 若 $f(x_0^+)$ 与 $f(x_0^-)$ 都存在,则当 $f(x_0^+)\neq f(x_0^-)$ 时,称 x_0 为跳跃间断点;而当 $f(x_0^+)=f(x_0^-)$,但是不与 $f(x_0)$ 相等,或 $f(x)$ 于 x_0 无定义时,称 x_0 为可去间断点.

例 $x=0$ 是 $f(x)=\dfrac{\sin x}{x}$ 的可去间断点,是 $g(x)=\dfrac{\sin x}{|x|}$ 的跳跃间断点,但却是

$$h(x)=\begin{cases}\dfrac{\sin x}{x}, & x\neq 0,\\ 1, & x=0\end{cases}$$

的连续点.

【第二类间断点】若 $f(x_0^+)$ 与 $f(x_0^-)$ 有一个是无穷大或无限振荡,则称 x_0 为函数 $f(x)$ 的第二类间断点.

例　由于 $x\to 0$ 时,函数 $f(x)=\frac{1}{x}\to\infty$,函数 $g(x)=\sin\frac{1}{x}$ 无限振荡,故 $x=0$ 是这两个函数的第二类间断点.

【连续函数的运算性质】

(1)若 $f(x)$,$g(x)$ 在点 x_0 处连续,则其四则运算 $f(x)\pm g(x)$,$f(x)g(x)$ 以及 $f(x)/g(x)$(要求 $g(x_0)\neq 0$)也在点 x_0 处连续.

(2)若函数 $u=\varphi(x)$ 在点 x_0 处连续且 $u_0=\varphi(x_0)$,而 $f(u)$ 在点 u_0 处连续,则复合函数 $f(\varphi(x))$ 在点 x_0 处连续.

(3)若函数 $y=f(x)$ 在区间 (a,b) 内严格单调,于点 $x_0\in(a,b)$ 处连续,则其反函数 $x=f^{-1}(y)$ 于 $y_0=f(x_0)$ 处连续.

【初等函数的连续性】初等函数在其定义区间内连续.

【利用连续性求极限】如果 $f(u)$ 是初等函数,于点 u_0 处有定义,则在 $\lim\limits_{x}\varphi(x)=u_0$ 时,便有

$$\lim_{x} f(\varphi(x))=f(\lim_{x}\varphi(x))$$

(记号 $\lim\limits_{x}$ 泛指变量 x 的任何一种确定的极限过程).

【闭区间上连续函数的性质】设 $f(x)$ 在 $[a,b]$ 上连续,则成立以下结果.

〖最值定理〗$f(x)$ 在 $[a,b]$ 上有界,且函数值中

存在最大值M及最小值m.

〖零点定理〗当$f(a)f(b)<0$时,存在点$x_0\in(a,b)$,使$f(x_0)=0$.

〖介值定理〗任给$A\in[m,M]$,必有$x_0\in(a,b)$,使$f(x_0)=A$.

常见题型·应对

【确定函数的间断点及其类型】首先找出可能的考察点,然后通过单侧极限判定.例如:

$$f(x)=\frac{x-1}{x^2-x^3}$$

的间断点是没有定义的点$x=0,x=1$,依次是第二类与第一类间断点.

【由分段函数的连续性求待定参数】连续性条件相当于告诉我们在连续点的函数极限存在,从而左、右极限相等,据此可以计算待定的参数.例如,若函数$f(x)$在点$x=0$处可导,分段函数

$$F(x)=\begin{cases}\dfrac{f(x)+a\sin x}{x}, & \\ 1 & \end{cases}$$

在点$x=0$连续,则可以求得

$$\begin{aligned}1&=\lim_{x\to0}F(x)=\lim_{x\to0}\left[\frac{f(x)-f(0)}{x-0}+\frac{a\sin x}{x}\right]\\&=f'(0)+a,\end{aligned}$$

从而

$$a=1-f'(0).$$

【用数列极限定义的函数的间断点】有的分段函数的形式是由数列极限结果定义的,此时要根据自变量x的取值来计算极限,得到分段函数的公式后再

判定间断点.例如:

$$f(x)=\lim_{n\to\infty}\frac{1+x}{1+x^{2n}}=\begin{cases}1+x, & |x|<1,\\ 0, & x=-1,|x|>1,\\ 1, & x=1\end{cases}$$

的间断点是(作图法判定)跳跃间断点 $x=1$.

【零点定理应用要点】若问题归结为证明方程 $f(x)=0$ 在区间 (a,b) 内有根,则只需验证 $f(x)$ 在 $[a,b]$ 的连续性,然后寻找 (a,b) 内两个点 α,β(通常是 a,b 或题目中较特殊的点),使 $f(\alpha)f(\beta)\leqslant 0$ 即可.

【介值定理应用要点】若问题归结为"存在 $\xi\in(a,b)$,使 $f(\xi)=A$",则只需证 A 在连续函数 $f(x)(a\leqslant x\leqslant b)$ 的值域中.

例 设 $f(x)$ 在 $[a,b]$ 上连续,$x_1,x_2,x_3\in[a,b]$,证明存在 $\xi\in[a,b]$,使

$$f(\xi)=\frac{1}{3}[f(x_1)+f(x_2)+f(x_3)].$$

解 注意到三个数的平均值必然在其中的最大值与最小值之间,从而必在 $f(x)$ 的值域之中,应用介值定理即得.

【最值定理应用要点】在证明函数不等式

$$A\leqslant f(x)\leqslant B\quad(a\leqslant x\leqslant b)$$

时,只需要证明 $f(x)$ 的最小值 m 与最大值 M 在 $[A,B]$ 之中便可.

【有界定理应用要点】闭区间 $[a,b]$ 上连续函数是有界函数;开区间 (a,b) 内连续函数不一定是有界函数;如果 $f(a^+),f(b^-)$ 都存在,则开区间 (a,b) 内连续函数一定是有界函数.

第 2 章　一元函数微分学

2.1　导数与微分

大纲要求

理解导数和微分的概念、相互关系，理解导数的几何意义. 会求平面曲线的切线和法线方程. 了解导数的物理意义，会用导数描述一些物理量. 理解函数的可导性与连续性的关系. 掌握导数的四则运算法则和复合函数的求导法则，掌握基本初等函数的导数公式. 了解微分的四则运算法则和一阶微分形式的不变性，会求函数的微分. 了解高阶导数的概念，会求简单函数的 n 阶导数. 会求分段函数的导数. 会求隐函数和参数方程确定的函数以及反函数的导数.

知识点 · 点拨

【改变量】设函数 $y=f(x)$ 在 (a,b) 内有定义，$x_0\in(a,b)$，$\Delta x\neq 0$ 且 $x_0+\Delta x\in(a,b)$. 称 Δx 为自变量 x 的改变量，$\Delta y=f(x_0+\Delta x)-f(x_0)$ 为函数 $y=f(x)$ 在点 x_0 处的改变量.

【点 x_0 处的导数】设函数 $y=f(x)$ 在 (a,b) 有定义，$x_0\in(a,b)$. 若 $L=\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}$ 存在，则说函数 $y=f(x)$ 在点 x_0 处可导，否则为不可导. 称 L 为函数 $y=f(x)$ 在点 x_0 处的导数，记作 $f'(x_0)$ 或 $y'(x_0)$ 或者用微分符号

表示为$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=x_0}$，$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0}$，亦即

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}.$$

【导数定义的等价形式】为了表述方便，导数定义中的极限有时写成以下等价形式：

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0};$$

$$f'(x_0)=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}.$$

〖点拨〗以下两种表达方式不是等价形式：

$$f'(x_0)=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0-h)}{2h};$$

$$f'(x_0)=\lim_{h\to +\infty}\frac{f(x_0+1/h)-f(x_0)}{1/h}.$$

因为前者没有包含 $f(x_0)$，后者是单侧极限.

【左导数】设函数 $y=f(x)$ 在包含 x_0 的某个区间 $(a,x_0]$ 上有定义. 若极限 $L=\lim\limits_{\Delta x\to 0^-}\frac{\Delta y}{\Delta x}$ 存在，则说函数 $y=f(x)$ 在点 x_0 处左可导，并称 L 为函数 $y=f(x)$ 在点 x_0 处的左导数，记作 $f'_-(x_0)$ 或 $y'_-(x_0)$，亦即

$$f'_-(x_0)=\lim_{\Delta x\to 0^-}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0}.$$

【右导数】设函数 $y=f(x)$ 在包含 x_0 的某个区间 $[x_0,b)$ 上有定义. 若极限 $L=\lim\limits_{\Delta x\to 0^+}\frac{\Delta y}{\Delta x}$ 存在，则说函数 $y=f(x)$ 在点 x_0 处右可导，并称 L 为函数 $y=f(x)$ 在

点 x_0 处的右导数,记作 $f'_+(x_0)$ 或 $y'_+(x_0)$,亦即

$$f'_+(x_0)=\lim_{\Delta x\to 0^+}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{x\to x_0^+}\frac{f(x)-f(x_0)}{x-x_0}.$$

【开区间上的导函数】如果 $f(x)$ 在 (a,b) 内每一点都可导,则说函数 $f(x)$ 在开区间 (a,b) 内可导,此时 $f'(x)\ (x\in(a,b))$ 是开区间 (a,b) 内的函数,称为导函数.

【闭区间上的导函数】如果 $f(x)$ 在 (a,b) 内可导,在端点 a 处右可导,在端点 b 处左可导,则说函数 $f(x)$ 在闭区间 $[a,b]$ 上可导. 此时称

$$f'(x)=\begin{cases}f'_+(a), & x=a,\\ f'(x), & a<x<b,\\ f'_-(b), & x=b\end{cases}$$

为 $f(x)$ 在闭区间 $[a,b]$ 上的导函数.

【可导与连续】若 $f(x)$ 在点 x_0 处可导,则 $f(x)$ 在点 x_0 处连续,反之不然.

〖点拨〗经常使用其逆否命题:若 $f(x)$ 在点 x_0 处不连续,则 $f(x)$ 在点 x_0 处不可导.

【连续但不可导的例子】函数 $f(x)=|x|$ 在点 $x=0$ 处连续但左右导数不相等,因而不可导.

【仅在一点可导的例子】函数

$$f(x)=\begin{cases}x^2, & x\text{ 为有理数},\\ 0, & x\text{ 为无理数}\end{cases}$$

处处有定义,但是仅在点 $x_0=0$ 处连续、可导.

〖点拨〗因

$$0 \leqslant f'_+(0) = \lim_{x \to 0^+} \frac{f(x)}{x} \leqslant \lim_{x \to 0^+} \frac{x^2}{x} = 0,$$

故 $f'_+(0)=0$；同理可证 $f'_-(0)=0$. 所以 $f(x)$ 在 $x_0=0$ 处可导，从而连续. 其次，当 $x_0 \neq 0$ 时，选取趋于 x_0 的无理数列 $\{a_n\}$ 及有理数列 $\{b_n\}$，则有

$$\lim_{n \to \infty} f(a_n) = 0, \quad \lim_{n \to \infty} f(b_n) = \lim_{n \to \infty} b_n^2 = x_0^2.$$

可见 $\lim\limits_{x \to x_0} f(x)$ 不存在（否则，以上两种方式的极限应当相等），故 $f(x)$ 于 x_0 处不连续，从而不可导.

〖推论〗当 $f(x)$ 在一点可导时，不可以说 $f(x)$ 在该点附近可导或连续.

【不可导点分类】不可导点有三个类型：

(1)不连续点；

(2)左、右导数都存在但不相等的连续点（称为折点，见图 2-1-1）；

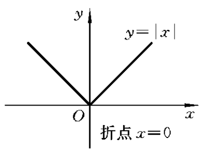


图 2-1-1

(3)左、右导数至少有一个不存在的连续点（见图 2-1-2、图 2-1-3）.

【四则运算求导规则】设 $u=u(x)$，$v=v(x)$ 是定义在

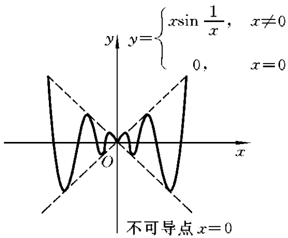


图 2-1-2

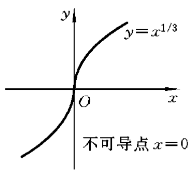


图 2-1-3

同一区间上的可导函数，α,β **是常数，则**

〖**线性规则**〗　$(\alpha u+\beta v)'=\alpha u'+\beta v'$；

〖**积规则**〗　$(uv)'=u'v+uv'$；

〖**商规则**〗　$(u/v)'=(u'v-uv')/v^2\quad(v\neq 0)$.

【复合函数求导规则】设函数 $y=f(u)$ **和** $u=g(x)$ **都是可导函数，且复合函数** $y=f(g(x))$ **有意义，则有**

$$[f(g(x))]'=f'(g(x))\cdot g'(x),$$

或者用微分表示为

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}.$$

【反函数求导法则】设二阶可导函数 $y=f(x)$ 在区间 I 上有反函数 $x=f^{-1}(y)$，$f'(x)\neq 0$，则有反函数的一阶和二阶求导公式

$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{1}{\mathrm{d}y/\mathrm{d}x}=\frac{1}{f'(x)},$$

$$\frac{\mathrm{d}^2x}{\mathrm{d}y^2}=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{f'(x)}\right)\cdot\frac{\mathrm{d}x}{\mathrm{d}y}=-\frac{f''(x)}{f'^3(x)}.$$

【奇偶性与导函数】设 $f(x)$ 是区间 $(-a,a)$ 内的可导函数，则

(1) 若 $f(x)$ 是奇函数，则 $f'(x)$ 是偶函数；

(2) 若 $f(x)$ 是偶函数，则 $f'(x)$ 是奇函数.

【周期性与导函数】设 $f(x)$ 是以 T 为周期的可导函数，则其导函数也是以 T 为周期的函数.

【单调性与导函数】单调函数的导函数不一定单调. 例如：$f(x)=x^3(-\infty<x<+\infty)$.

【有界性与导函数】有界函数的导函数不一定是有界函数. 例如：函数 $f(x)=\sqrt{x}$ 在区间 $(0,1)$ 内有界，但是 $f'(x)=\dfrac{1}{2\sqrt{x}}$ 在 $(0,1)$ 内无界.

【函数极限与导函数极限】由 $f(+\infty)=$ 常数，不能推出 $f'(+\infty)=0$，如函数 $f(x)=\dfrac{\sin x^2}{x}$；另一方面，由 $f'(+\infty)=0$ 也不能推出 $f(+\infty)=$ 常数，如函数 $f(x)=\ln x$.

【二阶导数】若函数 $y=f(x)$ 在 (a,b) 内处处可导，则

导函数$f'(x)$在点$x_0\in(a,b)$处的导数称为$f(x)$在点x_0处的二阶导数,记作$f''(x_0)$或$y''(x_0)$,其极限形式为

$$f''(x_0)=\lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{x-x_0}.$$

【一阶导数的几何意义】一阶导数$f'(x)$表示曲线$y=f(x)$在点$(x,f(x))$处的切线斜率,其符号表示倾斜方向,而绝对值表示倾斜程度.

【二阶导数的几何意义】二阶导数$f''(x)$表示曲线$y=f(x)$在点$(x,f(x))$处的凹凸性质,其符号表示凹凸方向,$f''(x)>0$时表示曲线下凸,$f''(x)<0$时表示曲线上凸.

【导数的物理意义】一阶导数$s'(t)$表示直线运动$s(t)$的速度,而二阶导数$s''(t)$表示速度$s'(t)$的变化率,即加速度.

【对数求导法】公式$y'=y(\ln y)'$适合于幂指函数或连乘连除函数的导数计算.例如:

(1)$y=u^v,y'=u^v(v\ln u)'$;

(2)$y=uv/w,y'=y(\ln u+\ln v-\ln w)'$.

【n阶导数】若函数$y=f(x)$在(a,b)内有$n-1$阶导数$(n\geqslant 1)$,则$n-1$阶导函数$f^{(n-1)}(x)$在点$x_0\in(a,b)$处的导数称为$f(x)$在点x_0处的n阶导数,记作$f^{(n)}(x_0)$或$y^{(n)}(x_0)$.

【n阶导数计算规则】可以使用以下规则来简化n阶导数的计算.

(1)线性规则　$(\alpha u+\beta v)^{(n)}=\alpha u^{(n)}+\beta v^{(n)}$;

(2)**乘积规则**　$(uv)^{(n)}=\sum_{k=0}^{n}C_n^k u^{(n-k)}v^{(k)}$;

(3)**链导规则**　$[u(ax+b)]^{(n)}=a^n u^{(n)}(ax+b)$.

其中,α,β **是常数**,u,v **是** n **阶可导的函数**.

【微分】设函数 $y=f(x)$**在点** x **的某个邻域有定义.若存在与** Δx **无关的常数** A,**使得**

$$\Delta y=f(x+\Delta x)-f(x)=A\Delta x+o(\Delta x),$$

则说函数 $f(x)$**在点** x **处可微,称** $A\Delta x$ **为函数** $y=f(x)$**在点** x **处的微分,记作** $\mathrm{d}y$ **或** $\mathrm{d}f(x)$,**即** $\mathrm{d}y=A\Delta x$.

〖**点拨 1**〗**称** Δx **为自变量的微分,通常用** $\mathrm{d}x$ **来记** Δx. Δx **的大小与** x **没有关系**.

〖**点拨 2**〗**计算微分时,务必记得写上因式** $\mathrm{d}x$.

【可微与可导】函数 $f(x)$**在点** x **处:可微**$\Leftrightarrow$**可导**.

【微分与导数的比较】导数 $f'(x)$**是两个改变量** Δx,Δy **的比的极限** $\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}$,**微分** $\mathrm{d}y$ **是改变量** Δy **的线性主部** $A\Delta x$. **从计算公式** $\mathrm{d}f(x)=f'(x)\mathrm{d}x$ **来看,微分的大小取决于** x **及** Δx **这两个变量,而** $f'(x)$**只取决于** x.

例　$f(x)=x^2$ **在** $x=1$ **处的导数与微分分别为** $f'(1)=2$,$\mathrm{d}f(1)=2\mathrm{d}x$,**只有在给出** $\mathrm{d}x$ **之值后,微分值** $\mathrm{d}f(1)$**才是常量**.

【微分计算规则】设 α,β **为常数**,f,u,v **为可微函数,则有:**

(1)**线性规则**　$\mathrm{d}(\alpha u+\beta v)=\alpha\mathrm{d}u+\beta\mathrm{d}v$;

(2)**积规则**　$\mathrm{d}(uv)=u\mathrm{d}v+v\mathrm{d}u$;

(3)**商规则**　$\mathrm{d}(u/v)=(v\mathrm{d}u-u\mathrm{d}v)/v^2\quad(v\neq 0)$;

(4)**链规则**　$\mathrm{d}f(u)=f'(u)\mathrm{d}u$.

【导数看作微分的商】可以通过分别计算微分 $\mathrm{d}y$ **与** $\mathrm{d}x$(**称作微商法**)**来得到导数** $y'(x)=\dfrac{\mathrm{d}y}{\mathrm{d}x}$.

例　**设** $x=\ln(1+t^2)$, $y=t-\arctan t$ **确定了函数** $y=y(x)$. **求** $y'(x)$, $y''(x)$.

解　**因** $\mathrm{d}x=\dfrac{2t}{1+t^2}\mathrm{d}t$, $\mathrm{d}y=\left(1-\dfrac{1}{1+t^2}\right)\mathrm{d}t$, **故**

$$y'(x)=\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{t^2}{2t}=\frac{t}{2},\quad \mathrm{d}y'(x)=\frac{1}{2}\mathrm{d}t,$$

$$y''(x)=\frac{\mathrm{d}y'(x)}{\mathrm{d}x}=\left(\frac{1}{2}\mathrm{d}t\right)\Big/\left(\frac{2t}{1+t^2}\mathrm{d}t\right)=\frac{1+t^2}{4t}.$$

〖**点拨**〗**与计算参变量函数的二阶导数的复合法与公式法相比较,以上的微商法较简单**.

常见题型·应对

【函数可导性的讨论问题】目的在于考察导数的概念. 较多的题目是研究分段函数在指定的分段点处的连续性和可导性,或者直接计算导数. 应对方法有两种:主观题使用导数定义求,客观题使用导函数的极限来求.

例　**设** $f(x)=\begin{cases}x^3, & x<0,\\ x^2, & x\geqslant 0,\end{cases}$ **讨论** $f'(0)$.

方法1　**分侧依左、右导数定义计算:**

$$f'_-(0)=\lim_{x\to 0^-}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0^-}\frac{x^3}{x}=0,$$

$$f'_+(0)=\lim_{x\to 0^+}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0^+}\frac{x^2}{x}=0.$$

可见, $f'_-(0)=f'_+(0)$, $f'(0)$**存在且等于** 0.

方法 2　首先确定 $f(x)$ 在 $x=0$ 处连续，然后应用求导公式求出 $f(x)$ 在 $x=0$ 两侧的导函数：

$$f'(x)=\begin{cases}3x^2, & x<0,\\ 2x, & x\geqslant 0.\end{cases}$$

由于极限 $f'(0^+)$ 及 $f'(0^-)$ 存在，故 $f'_+(0)=f'(0^+)=0$，$f'_-(0)=f'(0^-)=0$，从而 $f'(0)=0$ 存在.

〖点拨〗方法 2 的原理：若函数 $y=f(x)$ 在点 $x=0$ 处连续，且 $f'(0^+)=L$ 存在，则由导数定义及洛必达法则即知右导数

$$f'_+(0)=\lim_{x\to 0^+}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0^+}f'(x)=L.$$

类似地说明左导数 $f'_-(0)$ 的存在性. 其中连续性是必要条件，而导函数极限存在是充分条件. 方法 2 多用在无需表示过程的题型中.

【由分段函数可导性求参数问题】若已知分段函数 $f(x)$ 在分段点 x_0 处可导，则意味着

$$f(x_0^-)=f(x_0^+),\quad f'_-(x_0)=f'_+(x_0).$$

从中可求出函数中的待定参数.

例　设函数 $f(x)=\begin{cases}x^2, & x\leqslant 1,\\ ax+b, & x>1\end{cases}$ 在点 $x=1$ 处可导，则 $a=$________，$b=$________.

解　由于可导则连续，故由 $f(1^-)=f(1^+)$，得 $1=a+b$，而由 $f'_-(1)=f'_+(1)$ 或用加强的条件 $f'(1^-)=f'(1^+)$（导函数的左右极限相等），得 $2=a$，从而解出 $b=-1$.

〖点拨〗对需要写出计算过程的试题，只可以用导数定义来判定 $f'_-(1)=f'_+(1)$；而对客观题则可以

用容易求得的导函数的左右极限来代替左右导数 $f'(1^-)=f'_-(1^+)$，$f'(1^+)=f'_+(1)$. 这一方法要求事先明确函数在分段点为连续，并且导函数的极限是收敛的.

【计算参变量函数的导数】设方程组 $\begin{cases} x=x(t), \\ y=y(t) \end{cases}$ 确定了一元函数 $y=y(x)$，则可以采用如下两种方法：

〖复合法〗 $$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y'(t)}{x'(t)},$$
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{y'(t)}{x'(t)}\right)\cdot\frac{\mathrm{d}t}{\mathrm{d}x}.$$

〖公式法〗 $$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{x'(t)y''(t)-x''(t)y'(t)}{x'(t)^3}.$$

【计算隐函数的导数】设方程 $F(x,y)=0$ 确定了一元函数 $y=y(x)$，则先在方程式两边同时对 x 求导，再从结果中解出导数即可.

例 计算由方程 $\arctan\frac{y}{x}=\ln\sqrt{x^2+y^2}$ 所确定的函数 y 的一阶导数.

解 $$\frac{1}{1+\left(\frac{y}{x}\right)^2}\left(\frac{y}{x}\right)'=\frac{1}{2}\frac{1}{x^2+y^2}(x^2+y^2)',$$

从中解得 $$y'=\frac{x+y}{x-y}.$$

【计算 n 阶导数问题】有以下两种题型：

(1)计算函数在任意一点的 n 阶导数. 考虑使用相应的求导规则和公式.

(2)求函数在特定点 x_0 的 n 阶导数. 除了可以先计算任意一点的 n 阶导数之后再代点 x_0 的值之外，

还可以使用函数在点 x_0 的泰勒公式计算：从 $a_n=\frac{f^{(n)}(x_0)}{n!}$ 中解出 n 阶导数 $f^{(n)}(x_0)$.

【相关变化率问题】若在某个系统中，两个一元函数 $x(t)$ 及 $y(t)$ 之间满足关系式

$$F(x(t),y(t))=0,$$

则于该关系式两边同时对 t 求导，便可得到联系着两个变化率 $x'(t),y'(t)$ 的关系式

$$G(x(t),y(t),x'(t),y'(t))=0,$$

联立上述两个关系式便可以求解所给问题.

例　设以 $50\ \mathrm{cm}^3/\mathrm{s}$ 的速率给一气球充气. 若在充气过程中气球内气压保持为常数，气球保持为球形，问气球半径为 $5\ \mathrm{cm}$ 时半径增加的变化率是多大？

解　分别以 V,r 表示气球的体积与半径，则 V,r 都是时间 t 的函数，它们满足关系式

$$V(t)=\frac{4}{3}\pi r^3(t),$$

对 t 求导后得 $V'(t)=4\pi r^2(t)r'(t)$. 于是所求的变化率为 $r'(t)=V'(t)/4\pi r^2(t)$，代入 $V'(t)=50\ \mathrm{cm}^3/\mathrm{s}$，$r(t)=5\ \mathrm{cm}$，便可以解得 $r'(t)=\frac{1}{2\pi}\ \mathrm{cm/s}$.

【导数计算公式表】

(1) $(C)'=0$；
(2) $(x^\alpha)'=\alpha x^{\alpha-1}\quad(x>0)$；
(3) $(a^x)'=a^x\ln a\quad(1\neq a>0)$；

(4) $(e^x)'=e^x$;

(5) $(\ln x)'=\dfrac{1}{x}$;

(6) $(\log_a x)'=\dfrac{1}{x\ln a}\quad(1\neq a>0)$;

(7) $(\sin x)'=\cos x,\quad(\mathrm{sh}x)'=\mathrm{ch}x$;

(8) $(\cos x)'=-\sin x,\quad(\mathrm{ch}x)'=\mathrm{sh}x$;

(9) $(\tan x)'=\dfrac{1}{\cos^2 x}=\sec^2 x$,

$(\mathrm{th}x)'=\dfrac{1}{\mathrm{ch}^2 x}$;

(10) $(\cot x)'=-\dfrac{1}{\sin^2 x}=-\csc^2 x$,

$(\mathrm{coth}x)'=-\dfrac{1}{\mathrm{sh}^2 x}$;

(11) $(\arcsin x)'=\dfrac{1}{\sqrt{1-x^2}}$,

$(\arccos x)'=-\dfrac{1}{\sqrt{1-x^2}}$;

(12) $(\arctan x)'=\dfrac{1}{1+x^2}$,

$(\mathrm{arccot}x)'=-\dfrac{1}{1+x^2}$.

表中 C,α,a 表示常数.

【微分计算公式表】

(1) $\mathrm{d}C=0$;

(2) $\mathrm{d}x^\alpha=\alpha x^{\alpha-1}\mathrm{d}x\quad(x>0)$;

(3) $\mathrm{d}a^x=a^x\ln a\mathrm{d}x\quad(1\neq a>0)$;

(4) $\mathrm{d}e^x=e^x\mathrm{d}x$;

(5) $\mathrm{d}\ln x=\frac{1}{x}\mathrm{d}x$；

(6) $\mathrm{d}\log_a x=\frac{1}{x\ln a}\mathrm{d}x \quad (1\neq a>0)$；

(7) $\mathrm{d}\sin x=\cos x\mathrm{d}x$，　$\mathrm{d}\,\mathrm{sh}x=\mathrm{ch}x\mathrm{d}x$；

(8) $\mathrm{d}\cos x=-\sin x\mathrm{d}x$，　$\mathrm{d}\,\mathrm{ch}x=\mathrm{sh}x\mathrm{d}x$；

(9) $\mathrm{d}\tan x=\frac{1}{\cos^2 x}\mathrm{d}x=\sec^2 x\mathrm{d}x$；

(10) $\mathrm{d}\cot x=-\frac{1}{\sin^2 x}\mathrm{d}x=-\csc^2 x\mathrm{d}x$；

(11) $\mathrm{d}\arcsin x=\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$，

$\mathrm{d}\arccos x=-\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$；

(12) $\mathrm{d}\arctan x=\frac{1}{1+x^2}\mathrm{d}x$，

$\mathrm{d}\,\mathrm{arccot}x=-\frac{1}{1+x^2}\mathrm{d}x$.

表中 C,α,a 表示常数，u 为可微函数.

【常见函数的 n 阶导数公式】

$$\left(\frac{1}{ax+b}\right)^{(n)}=(-1)^n\frac{a^n n!}{(ax+b)^{n+1}};$$

$$\left(\frac{1}{x}\right)^{(n)}=(-1)^n\frac{n!}{x^{n+1}};$$

$$(\sin x)^{(n)}=\sin\left(x+\frac{n\pi}{2}\right);$$

$$(\cos x)^{(n)}=\cos\left(x+\frac{n\pi}{2}\right),$$

$$(a^x)^{(n)}=a^x(\ln a)^n,$$

$$(\mathrm{e}^x)^{(n)}=\mathrm{e}^x.$$

2.2 导数的应用

大纲要求

理解并会用罗尔定理、拉格朗日中值定理和泰勒定理,了解并会用柯西中值定理.掌握用洛必达法则求未定式极限的方法.理解函数的极值概念,掌握用导数判断函数的单调性和求函数极值的方法,掌握函数最大值和最小值的求法及其简单应用.会用导数判断函数图形的凹凸性,会求函数图形的拐点以及水平、铅直和斜渐近线,会描绘函数的图形.了解曲率和曲率半径的概念,会计算曲率和曲率半径.

知识点·点拨

【平面曲线的切线与法线】曲线$y=y(x)$**在其上一点**(x_0,y_0)**处的切线方程为**

$$y-y_0=y'(x_0)(x-x_0),$$

与切线垂直的过点(x_0,y_0)**的直线称为曲线的法线,其方程为**

$$y-y_0=-\frac{1}{y'(x_0)}(x-x_0)\quad (y'(x_0)\neq 0).$$

【渐近线】如果曲线$y=f(x)$**上的点**$P(x,y)$**与某条直线**L**的距离**$d(P,L)$**当点**$P(x,y)$**趋于无穷(即**$x^2+y^2\to\infty$**)时趋于零,则称直线**L**是曲线**$y=f(x)$**的一条渐近线.渐近线分为水平、铅直、斜渐近线三种(单侧极限成立便可以).**

〖**铅直渐近线**$x=a$〗**满足**$\lim\limits_{x\to a^{\pm}}f(x)=\infty$.

〖**水平渐近线**$y=A$〗**满足**$\lim\limits_{x\to\pm\infty}f(x)=A$.

〖斜渐近线 $y=kx+b$〗满足

$$\lim_{x\to\pm\infty}\frac{f(x)}{x}=k,\quad \text{且}\quad \lim_{x\to\pm\infty}(f(x)-kx)=b.$$

【曲率】曲率是曲线在一点的弯曲程度的衡量. 若函数 $y=f(x)$ 二阶可导,则在点 x 处的曲率公式为

$$\kappa=|y''|/[1+(y')^2]^{3/2}.$$

【曲率半径】曲率的倒数称为曲率半径:$R=1/\kappa$.

【单调性的判定】(1)函数 $f(x)$ 在区间 (a,b) 内单调增(减)$\Leftrightarrow f'(x)\geqslant 0(\leqslant 0)(a<x<b)$;

(2)函数 $f(x)$ 在 (a,b) 内严格单调增(减)$\Leftrightarrow f'(x)\geqslant 0(\leqslant 0)(a<x<b)$ 且使 $f(x)=0$ 的点不充满任何子区间 $(\alpha,\beta)\subseteq(a,b)$.

【凸性的定义】若对任何 $x_1,x_2\in(a,b)$,有

$$f\left(\frac{x_1+x_2}{2}\right)\leqslant(\geqslant)\frac{f(x_1)+f(x_2)}{2},$$

称函数 $f(x)(a<x<b)$ 为下凸(上凸)函数.

〖点拨〗下凸函数的曲线称为下凸弧,表示从曲线的下面来看,曲线是凸起的. 也可称之为上凹,指从曲线的上面来看,曲线是凹进的. 类似地理解上凸弧.

【凸性的判别】若 $f''(x)\geqslant 0(f''(x)\leqslant 0)(a<x<b)$,则 $f(x)$ 是下凸(上凸)函数.

【拐点的定义】设函数 $f(x)$ 在连续点 x_0 的两侧的凸性相反,则称点 x_0 是函数 $f(x)$ 的拐点,等价地称点 $(x_0,f(x_0))$ 是曲线 $y=f(x)$ 的拐点.

【拐点的判定】首先确定函数的凸性区间分布,两侧凸性发生改变的连续点便是拐点.

〖点拨 1〗当函数 $y=f(x)$ 二次可微时，$f''(x)=0$ 是拐点的必要条件，但非充分条件. 如 $y=x^4$ 在原点的二阶导数为零，但原点不是其拐点，因为 $f''(x)$ 在原点两侧没有变号，两侧的凸性相同.

〖点拨 2〗拐点是连续点的条件不可少. 例如曲线 $y=\dfrac{1}{x}$ 在原点(间断点)两侧凸性相反，但不可以讲原点是该函数的拐点.

【函数的极值】若有 x_0 的一个邻域 $(x_0-\delta, x_0+\delta)$，在其上 $f(x)\leqslant f(x_0)$(或 $f(x)\geqslant f(x_0)$)，则称函数值 $f(x_0)$ 是函数 $f(x)$ 的一个极大值(或极小值). 此邻域可以很大，也可以很小.

【函数的最值】若对所有的 $x\in I$，$f(x)\leqslant f(x_0)$(或者 $f(x)\geqslant f(x_0)$)，则称函数值 $f(x_0)$ 是函数 $f(x)$ 在区间 I 上的最大值(或最小值).

〖点拨〗谈及最值时要指明其控制范围，极值则不需要. 依据定义，极值不可以在定义区间的端点取得，因为它必须与两侧的函数值进行比较.

【极值点与最值点】当 $f(x_0)$ 为函数的极值时，称 x_0 为函数 $f(x)$ 的极值点，称 $(x_0, f(x_0))$ 为曲线 $y=f(x)$ 的极值点. 最值点类似理解.

【驻点】使函数 $f(x)$ 的导数为零的点 $x=x_0$.

【极值必要条件】函数 $f(x)$ 的极值点要么是驻点，要么是不可导点.

【极值充分条件 1】对连续点 x_0，当 $f'(x)$ 在其两侧左正右负时，$f(x_0)$ 为极大值；当 $f'(x)$ 在其两侧左负右正时，$f(x_0)$ 为极小值.

【极值充分条件 2】对驻点 x_0，当 $f''(x_0)>0$ 时，$f(x_0)$ 为极小值；当 $f''(x_0)<0$ 时，$f(x_0)$ 为极大值.

【最值的存在性】只有明确了所论最值的存在性后方可进行搜寻. 以下是相关结果：

(1)闭区间 $[a,b]$ 上连续函数一定有最大值和最小值；

(2)开区间 (a,b) 内连续函数可能没有、也可能有最值. 例如 $f(x)=x^2$ 在 $(-1,1)$ 内只有最小值 $f(0)$，$f(x)=x^3$ 在 $(-1,1)$ 内既没有最大值也没有最小值.

【闭区间 $[a,b]$ 上连续函数的最值求法】首先求出 (a,b) 内的全部驻点以及不可导点，然后计算这些点的函数值，再和两个端点的函数值比较，其中最大者即最大值，最小者即最小值.

【开区间 (a,b) 内连续函数的最值求法】通过一阶导数的符号确定 (a,b) 内的单调区间分布，结合几何图形确定最值是否存在以及大小.

【不连续函数的最值求法】对于分段连续函数（指有有限个间断点的函数），应当在每个连续区间上进行最值分析然后综合.

【罗尔定理】设 $f(x)$ 在 $[a,b]$ 上连续，在 (a,b) 内可导，则当 $f(a)=f(b)$ 时存在 $\xi\in(a,b)$ 使得 $f'(\xi)=0$.

【拉格朗日中值定理】设 $f(x)$ 在 $[a,b]$ 上连续，在 (a,b) 内可导，则有 $\xi\in(a,b)$，使得

$$f(b)-f(a)=f'(\xi)(b-a),$$

或

$$f'(\xi)=\frac{f(b)-f(a)}{b-a}.$$

【柯西中值定理】设 $f(x)$,$g(x)$在$[a,b]$上连续,在(a,b)内可导,且 $g'(x)\neq 0$,则有 $\xi\in(a,b)$,使

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}.$$

【泰勒多项式】设 $f(x)$在 x_0 有 n 阶导数,则称

$$\begin{aligned}T_n(x)=f(x_0)&+f'(x_0)(x-x_0)\\&+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots\\&+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\end{aligned}$$

为函数 $f(x)$在 x_0 的 n 阶泰勒多项式.

【泰勒定理】设 $f(x)$在(a,b)上有 $n+1$ 阶导数,x_0,$x\in(a,b)$,则有介于 x_0 与 x 之间的 ξ,使得

$$f(x)=T_n(x)+R_n(x),$$

其中 $$R_n(x)=\frac{1}{(n+1)!}f^{(n+1)}(\xi)(x-x_0)^{n+1}$$

称为拉格朗日余项.

〖点拨〗若可导条件减弱为 $f(x)$在(a,b)内有 n 阶导数,则上述公式中的余项必须换作

$$R_n(x)=o((x-x_0)^n),$$

称之为皮亚诺余项.其意义在于说明,当 $x\to x_0$ 时,泰勒多项式与 $f(x)$的误差至少是 $x-x_0$ 的 n 阶无穷小量.

【直接法求函数的泰勒多项式】首先求出直到 n 阶的导数值 $f^{(k)}(x_0)$ $(1\leqslant k\leqslant n)$,然后写出 $T_n(x_0)$及余项.

【间接法求函数的泰勒多项式】将所给函数 $f(x)$用已知其泰勒多项式的函数 $u(x)$,$v(x)$表示,例如:

$$f(x)=u(x)+v(x),$$

$$f(x)=u'(x),\quad f(x)=\int_a^x u(x)\mathrm{d}x,$$

再代入$u(x)$,$v(x)$的泰勒多项式即可.

【展开点的调整】如果要将函数$f(x)$展开为$x-a$的泰勒多项式,则应当遵循以下步骤:

(1)作代换$x-a=t$,使

$$f(x)=f(x-a+a)=f(t+a);$$

(2)将$g(t)=f(t+a)$展开为t的泰勒多项式;

(3)代回$t=x-a$.

常见题型·应对

【相关变化率问题】设在某个考察过程中,变量u与v都是x的函数,并且相互之间满足某种关系$F(u,v)=0$,则变化率$u'(x)$,$v'(x)$之间也应当有相应的制约:$[F(u,v)]'_x=0$,从中可以解出所需的未知量.

〖点拨〗通常是已知$u(x)$,$v(x)$,$u'(x)$,$v'(x)$中的任意两个,求另外两个,结合方程式$F(u,v)=0$,及$[F(u,v)]'_x=0$,这是可以成功的.

【函数中值问题概述】

(1)函数方程$f(x)=0$在区间(a,b)内有根;

(2)函数$f(x)$在(a,b)内存在零点;

(3)证明存在$\xi\in(a,b)$,使$g(\xi)=h(a,b)$.

以上三种形式的问题可以互相转换,例如问题(3)相当于证明函数$f(x)=g(x)-h(a,b)$存在零点,因此归结为同一种题型.使用的工具(根据题给条件选择)是几个相关定理:连续函数的介值定理、微分中值定理、积分中值定理.

【如何说明零点的惟一性】证明零点的惟一性可以

采用两个方法：

(1)函数连续且严格单调；

(2)反证法，假设不惟一来导出矛盾.

【一个中值的问题 · 解法思路】待证关系为：存在 $\xi\in(a,b)$，使 $\varphi(a,b,\xi)=0$.

(1)若条件中的函数 $f(x)$ 仅假定为连续，则考虑用介值定理；若假定为在 (a,b) 内可导，则考虑用微分中值定理；若假定为二阶可导，则应当考虑使用泰勒定理.

(2)使用罗尔定理的程序. 适当地变形 $\varphi(a,b,\xi)$(等式变形或乘以非零因子)成为 $f(\xi)$，以 x 代替 ξ，求出 $f(x)$ 的原函数 $F(x)$，验证 $F(x)$ 在 $[a,b]$ 上某两点 x_1,x_2 成立 $F(x_1)=F(x_2)$ 后，即由罗尔定理得出要证的结果.

(3)使用拉格朗日或柯西中值定理的程序. 适当地变形 $\varphi(a,b,\xi)=0$，使之成为 $h(\xi)=\varphi(a,b)$，探求 $f(x)$ 使得 $f'(\xi)=h(\xi)$，同时 $\dfrac{f(b)-f(a)}{b-a}=\varphi(a,b)$，再应用拉格朗日中值定理；或者探求 $f(x)$ 及 $g(x)$ 使得 $\dfrac{f'(\xi)}{g'(\xi)}=h(\xi)$，同时 $\dfrac{f(b)-f(a)}{g(b)-g(a)}=\varphi(a,b)$，再应用柯西中值定理论证.

例 1　设 $f(x)$ 在 $[a,b]$ 上连续，在 (a,b) 内可导，$f(a)=f(b)=0$，λ 是常数，证明存在 $\xi\in(a,b)$，使 $f'(\xi)+\lambda f(\xi)=0$.

解　以 $e^{\lambda\xi}$ 乘等式两边，得

$$e^{\lambda\xi}f'(\xi)+\lambda e^{\lambda\xi}f(\xi)=0,$$

以 x 代 ξ，则问题变为

$$(e^{\lambda x}f(x))'=e^{\lambda x}f'(x)+\lambda e^{\lambda x}f(x)$$

的零点问题. 恰好题设条件已告知$e^{\lambda x}f(x)$在$x=a$及$x=b$相等,从而应用罗尔定理完成证明.

例 2　设$f(x),g(x)$在$[a,b]$上连续,在(a,b)内可导,$g'(x)\neq 0$,证明存在$\xi\in(a,b)$,使得

$$\frac{f(\xi)-f(a)}{g(b)-g(\xi)}=\frac{f'(\xi)}{g'(\xi)}.$$

解　消去分母,将方程变形为

$$f'(\xi)g(\xi)+f(\xi)g'(\xi)-f'(\xi)g(b)-f(a)g'(\xi)=0,$$

以x代ξ,则问题变为

$$F(x)=f(x)g(x)-f(x)f(b)-f(a)g(x)$$

的导函数$F'(x)$的零点问题,恰好可验证$F(a)=F(b)$. 从而应用罗尔定理完成证明.

例 3　设$f(x)$在$[a,b]$上连续,在(a,b)内可导,$0<a<b$,证明存在$\xi\in(a,b)$,使得

$$a\ln b-b\ln a=(ab^2-ba^2)(1-\ln\xi)/\xi^2.$$

解　分离a,b与ξ,得$\dfrac{\dfrac{\ln b}{b}-\dfrac{\ln a}{a}}{b-a}=\dfrac{1-\ln\xi}{\xi^2}$,则可看出,对$f(x)=\dfrac{\ln x}{x}$,应用拉格朗日中值定理即可得证.

例 4　设$0<a<b$,$f(x)$在$[a,b]$连续,在(a,b)可导,证明存在$\xi\in(a,b)$,使

$$f(b)-f(a)=\xi f'(\xi)\ln(b/a).$$

解　分离a,b与ξ,得$\dfrac{f(b)-f(a)}{\ln b-\ln a}=\xi f'(\xi)$,记$g(x)=\ln x$,则左边$=\dfrac{f(b)-f(a)}{g(b)-g(a)}$,右边恰好是$\dfrac{f'(\xi)}{g'(\xi)}$,故应用柯西中值定理完成证明.

【两个中值的问题 · 解法思路】待证关系为:存在ξ,$\eta\in(a,b)$,使$\varphi(a,b,\xi,\eta)=0$.一般的分析思路为:分离a,b与ξ,η,对含ξ部分及含η部分分别恰当地使用微分中值定理表示为a,b的函数,代回原方程式即可.

例1　设$0<a<b$,$f(x)$在$[a,b]$上连续,在(a,b)内可导,证明存在$\xi,\eta\in(a,b)$,使得

$$2\eta f'(\xi)=(a+b)f'(\eta).$$

解　分离ξ,η得

$$f'(\xi)=(a+b)\frac{f'(\eta)}{2\eta},$$

调用拉格朗日中值定理得

$$f'(\xi)=\frac{f(b)-f(a)}{b-a},$$

调用柯西中值定理得

$$\frac{f'(\eta)}{2\eta}=\frac{f(b)-f(a)}{b^2-a^2},$$

回代后即得证.

例2　$f(x)$在$[a,b]$上连续,在(a,b)内可导,$f(a)=0$,$f(b)=1$,证明存在不同的$\xi,\eta\in(a,b)$,使得

$$\frac{1}{f'(\xi)}+\frac{1}{f'(\eta)}=2(b-a).$$

解　由介值定理,选取$c\in(a,b)$,使$f(c)=\dfrac{1}{2}$.在$[a,c]$,$[c,b]$上分别调用拉格朗日中值定理,得

$$f'(\xi)=\frac{f(c)-f(a)}{c-a}=\frac{1}{2(c-a)},$$

$$f'(\eta)=\frac{f(b)-f(c)}{b-c}=\frac{1}{2(b-c)}.$$

回代后即得证.

【结合积分条件的中值问题】如果出现积分,则可调用积分中值定理或N-L公式进行分析.

例　设$f(x)$在$[0,1]$上连续,在$(0,1)$内可导,且满足$f(1)=a\int_0^{1/a}xe^{1-x}f(x)dx\quad(a>1)$,证明存在$\xi\in(0,1)$,使得$f'(\xi)=(1-\xi^{-1})f(\xi)$.

解　使用积分中值定理知,存在$c\in\left(0,\frac{1}{a}\right)$,使$f(1)=ce^{1-c}f(c)$,记$g(x)=xe^{1-x}f(x)$,则$g(1)=g(c)$.在区间$[c,1]$上对函数$g(x)$应用罗尔定理即得.

【罗尔定理推广】若$f(x)$在(a,b)上可导,$f(a^+)=f(b^-)$,则存在$\xi\in(a,b)$使$f'(\xi)=0$.

【拉格朗日中值定理推广】设以下函数在$[a,b]$上连续,在(a,b)内所写导数存在,则有

(1) 若曲线$y=f(x)$与曲线$y=g(x)$在$[a,b]$上有两个交点,则有$\xi\in(a,b)$,使$f'(\xi)=g'(\xi)$.取$g(x)$为通过$P(a,f(a)),Q(b,f(b))$的直线,便是拉格朗日中值定理.

(2) 若曲线$y=f(x)$与曲线$y=g(x)$在$[a,b]$上有三个交点,则有$\xi\in(a,b)$,使$f''(\xi)=g''(\xi)$.

例　若在区间$[0,2]$上可导曲线$y=f(x)$经过点$O(0,0)$,$A(1,1)$及$B(2,4)$,则存在$\xi\in(0,2)$,使$f''(\xi)=2$.

解　取$g(x)=x^2$,它也通过O,A,B三点,引用推广结果(2)即得.

【不等式问题概述】常见的几种不等式:

$$f(x)\leqslant g(x),\quad f(x)\leqslant M,\quad P\leqslant f(x)\leqslant Q$$

均可通过移项以及适当的化简归结为

$$F(x)\geqslant 0 \quad (a\leqslant x\leqslant b).$$

证明的方式主要有以下三种：

(1) 单调法(主要方法)　若 $F(x)$ 在 $[a,b]$ 上连续，在 (a,b) 内可导. 则当

$$F(a)\geqslant 0,\quad F'(x)\geqslant 0 \quad (a<x<b)$$

时便有　$F(x)\geqslant 0 \quad (a\leqslant x\leqslant b)$.

(2) 最值法　若能求得 $F(x)$ 在区间 $[a,b]$ 上的最小值为 0，则有

$$F(x)\geqslant 0,\quad a\leqslant x\leqslant b.$$

(3) 泰勒公式法　使用 0 阶或 1 阶泰勒公式：

$$f(x)=f(a)+f'(\xi)(x-a),$$

$$f(x)=f(a)+f'(a)(x-a)+\frac{1}{2}f''(\xi)(x-a)^2,$$

适当地放大或缩小余项也能得到函数不等式(例如凸函数问题).

例 1　证明 $1-\dfrac{x^2}{2}<\cos x<1-\dfrac{x^2}{\pi}\ \left(0<x<\dfrac{\pi}{2}\right)$.

解　将所证不等式变形为

$$-\frac{1}{2}<\frac{\cos x-\cos 0}{x^2-0}<-\frac{1}{\pi},$$

调用柯西中值定理作等式变形，得

$$\frac{1}{\pi}<\frac{\sin\xi}{2\xi}<\frac{1}{2}\quad\left(0<\xi<\frac{\pi}{2}\right).$$

记 $f(x)=\dfrac{\sin x}{2x}$，易知 $f'(x)<0\quad\left(0<x<\dfrac{\pi}{2}\right)$，从而由单调法知

$$\frac{1}{\pi}=f\left(\frac{\pi}{2}\right)<f(x)<f(0^+)=\frac{1}{2}\quad\left(0<x<\frac{\pi}{2}\right).$$

例 2　证明 $e^x-1>(1+x)\ln(1+x)\quad(x>0)$.

解　以下做法是标准方法.令

$$f(x)=\mathrm{e}^x-1-(1+x)\ln(1+x),$$

则

$$f'(x)=\mathrm{e}^x-\ln(1+x)-1.$$

由于$f'(x)$符号难定,因而考虑$f''(x)=\mathrm{e}^x-\dfrac{1}{1+x}$,由于$f''(x)>0(x>0)$及$f'(0)=0$,故推出

$$f'(x)>0(x>0),$$

再由$f(0)=0$推出$f(x)>0\ (x>0)$.

【常见函数的泰勒公式】在$x=0$点的展开公式,ξ在0与x之间.

$\mathrm{e}^x=1+x+\dfrac{1}{2!}x^2+\cdots+\dfrac{1}{n!}x^n+\dfrac{\mathrm{e}^\xi}{(n+1)!}x^{n+1},\quad \lvert x\rvert<+\infty.$
$\sin x=x-\dfrac{1}{3!}x^3+\dfrac{1}{5!}x^5-\cdots+(-1)^{n-1}\cdot\dfrac{1}{(2n-1)!}x^{2n-1}+(-1)^{n+1}\dfrac{1}{(2n+1)!}\cos\xi x^{2n+1},\lvert x\rvert<+\infty.$
$\cos x=1-\dfrac{1}{2!}x^2+\dfrac{1}{4!}x^4-\cdots+(-1)^n\cdot\dfrac{1}{(2n)!}x^{2n}+(-1)^{n+1}\dfrac{\cos\xi}{(2n+2)!}x^{2n+2},\lvert x\rvert<+\infty.$
$\dfrac{1}{1-x}=1+x+x^2+\cdots+x^n+o(x^n),\lvert x\rvert<1.$
$\ln(1+x)=x-\dfrac{1}{2}x^2+\dfrac{1}{3}x^3-\cdots+(-1)^{n-1}\dfrac{1}{n}x^n,+o(x^n),\lvert x\rvert<1.$

第3章　一元函数积分学

3.1　不定积分

大纲要求

理解原函数和不定积分的概念. 掌握不定积分的基本公式,掌握不定积分的性质,掌握不定积分的换元积分法与分部积分法. 会求有理函数、三角函数有理式和简单无理函数的积分.

知识点·点拨

【原函数定义】若在区间 I 上,$F'(x)=f(x)$,则称 $F(x)$为 $f(x)$的原函数. 等价说法是:$f(x)$是 $F(x)$的导函数.

【原函数的存在性】如果 $f(x)$是区间$[a,b]$上的连续函数,则 $F(x)=\int_a^x f(t)\mathrm{d}t$ 便是它的一个原函数.

〖点拨〗“有原函数”并不意味着可以使用初等函数将原函数表示出来. 例如连续函数 e^{x^2} 的原函数存在但不是初等函数.

【原函数的结构】若 $F(x)$是 $f(x)$的一个原函数,则 $F(x)+C$ 都是 $f(x)$的原函数,C 为任意常数;并且 $F(x)+C$ 包含了 $f(x)$的全部原函数.

【不定积分定义】若 $F(x)$是 $f(x)$的一个原函数,则原函数全体 $F(x)+C$ 称为 $f(x)$的不定积分,记作

$$\int f(x)\mathrm{d}x=F(x)+C.$$

〖点拨 1〗利用导数理解积分公式：

$$\int f(x)\mathrm{d}x=F(x)+C \Leftrightarrow \frac{\mathrm{d}F(x)}{\mathrm{d}x}=f(x).$$

〖点拨 2〗以下三个关系式等价：

(1) $\int f(u(x))\mathrm{d}v(x)=F(x)+C$；

(2) $\dfrac{\mathrm{d}F(x)}{\mathrm{d}v(x)}=f(u(x))$；

(3) $F'(x)=f(u(x))v'(x)$.

【不定积分互逆性质】求导与求积的关系：

$$\left(\int f(x)\mathrm{d}x\right)'=f(x),\quad \mathrm{d}\left(\int f(x)\mathrm{d}x\right)=f(x)\mathrm{d}x;$$

$$\int F'(x)\mathrm{d}x=F(x)+C,\quad \int \mathrm{d}F(x)=F(x)+C.$$

【不定积分线性性质】设 α,β 是常数，则有

$$\int(\alpha f(x)+\beta g(x))\mathrm{d}x=\alpha\int f(x)\mathrm{d}x+\beta\int g(x)\mathrm{d}x.$$

【凑微分积分法】若有 $\int f(u)\mathrm{d}u=F(u)+C$，则取 $u=\varphi(x)$，便产生新的积分公式：

$$\int f(\varphi(x))\mathrm{d}\varphi(x)=F(\varphi(x))+C.$$

【如何凑微分】当积分式中出现因式 $f(\varphi(x))$ 时，可考虑凑出 $u=\varphi(x)$ 的微分而将积分化作 $\int f(u)\mathrm{d}u$ 后进行积分. 例如：

$\int f(x^n)x^{n-1}\mathrm{d}x$	$u=x^n$
$\int f(\sin x)\cos x\mathrm{d}x$	$u=\sin x$

$\int f(\ln x)\frac{1}{x}\mathrm{d}x$	$u=\ln x$
$\int f(\arctan x)\frac{1}{1+x^2}\mathrm{d}x$	$u=\arctan x$
$\int f(\mathrm{e}^x)\mathrm{e}^x\mathrm{d}x$	$u=\mathrm{e}^x$

【换元积分法】 令 $x=\varphi(t)$($\varphi'(t)$连续且非零),便有

$$\int f(x)\mathrm{d}x=\int f(\varphi(t))\varphi'(t)\mathrm{d}t.$$

【常用换元形式】 引入新的积分变量的目的在于化简积分式,以下是几种消除根式的常见方法:

所含根式	作代换($a>0$)
$\sqrt{a^2-x^2}$	$x=a\sin t$
$\sqrt{x^2-a^2}$	$x=a\tan t$
$\sqrt{a^2+x^2}$	$x=a\sec t$
$\sqrt[n]{a+bx}$	$t=\sqrt[n]{a+bx}$

例 求 $I=\int \mathrm{e}^{\sqrt{x}}\mathrm{d}x$.

解 令 $\sqrt{x}=t$,即 $x=t^2$,有

$$I=\int \mathrm{e}^t\cdot 2t\mathrm{d}t=2\int t\mathrm{d}\mathrm{e}^t=2\left(t\mathrm{e}^t-\int \mathrm{e}^t\mathrm{d}t\right)$$

$$=2(t-1)\mathrm{e}^t+C=2(\sqrt{x}-1)\mathrm{e}^{\sqrt{x}}+C.$$

【分部积分法】 若 $u'(x)$较 $u(x)$简单,$v'(x)$不比 $v(x)$复杂,则以下分部积分公式能够简化积分计算:

$$\int u(x)\mathrm{d}v(x)=u(x)v(x)-\int v(x)u'(x)\mathrm{d}x.$$

〖点拨〗通常取 $u=\arctan x, \ln x, x^n$.

例 1 $\int x\mathrm{e}^x\mathrm{d}x=\int x\mathrm{d}\mathrm{e}^x=x\mathrm{e}^x-\int \mathrm{e}^x\mathrm{d}x$

$=\mathrm{e}^x(x-1)+C$.

例 2 $\int x\arctan x\mathrm{d}x$

$$
\begin{aligned}
&=\int \arctan x\mathrm{d}\,\frac{x^2}{2}\\
&=\frac{1}{2}x^2\arctan x-\frac{1}{2}\int\frac{x^2}{1+x^2}\mathrm{d}x\\
&=\frac{1}{2}x^2\arctan x-\frac{1}{2}\int\frac{1+x^2-1}{1+x^2}\mathrm{d}x\\
&=\frac{1}{2}x^2\arctan x-\frac{1}{2}(x-\arctan x)+C.
\end{aligned}
$$

【奇偶函数的原函数】设 $F(x)$ 是 $f(x)$ 在对称区间 $[-a,a]$ 上的原函数,则有

(1) $f(x)$ **为奇函数** $\Leftrightarrow F(x)$ **为偶函数**;

(2) $f(x)$ **为偶函数** $\Leftarrow F(x)$ **为奇函数**.

〖点拨〗偶函数的原函数并不全是奇函数. 如 $f(x)=x^2$ 有原函数 $F(x)=\frac{1}{3}x^3+C$,当 $C\neq 0$ 时,$F(x)$ 便不是奇函数 (参见 1.1 节).

【周期函数的原函数】设 $f(x)$ 是区间 $(-\infty,+\infty)$ 上的周期函数,周期为 T. 则 $f(x)$ 的原函数是周期为 T 的周期函数 $\Leftrightarrow$ 定积分 $\int_0^T f(t)\mathrm{d}t=0$(参见 1.1 节).

【有理函数】形为 $\frac{P(x)}{Q(x)}$(P,Q 为多项式)的函数称为有理函数,它包括多项式及分式函数.

【有理函数的积分】理论上可以证明,有理函数的原函数是初等函数,因而其积分总可以求出. 实际计算

时,采用以下步骤积分.

(1)**分项**　使用多项式除法和部分分式法将有理函数表示为多项式及以下的最简分式之和:

$$a(x)=\frac{1}{(x-a)^n},$$

$$b(x)=\frac{Ax+B}{(x^2+px+q)^n}\quad (p^2<4q).$$

(2)**逐项积分**

$$\int a(x)\mathrm{d}x=\frac{1}{1-n}\frac{1}{(x-a)^{n-1}}+C,$$

$$\begin{aligned}\int b(x)\mathrm{d}x &=\frac{A}{2}\int\frac{\mathrm{d}(x^2+px+q)}{(x^2+px+q)^n}\\ &\quad+\left(B-\frac{Ap}{2}\right)\int\frac{\mathrm{d}x}{(x^2+px+q)^n}\\ &=\frac{A}{2(1-n)}\frac{1}{(x^2+px+q)^{n-1}}\\ &\quad+\left(B-\frac{Ap}{2}\right)\int\frac{\mathrm{d}\left(x+\frac{p}{2}\right)}{\left[\left(x^2+\frac{p}{2}\right)^2+q-\frac{p}{4}\right]^n},\end{aligned}$$

最后一个积分可以用分部积分法关于 n 递推计算.

【部分分式法】将一个有理函数分解为多项式及最简分式的过程叫部分分式法.根据有理函数特点,写出相应的包含待定系数的分解式,然后在消除分母后,比较同次项系数后求出待定系数即成.例如(用 $p_k(x)$ 表示 k 次多项式):

$$\frac{p_1(x)}{(x-1)^2}=\frac{A}{x-1}+\frac{B}{(x-1)^2},$$

$$\frac{p_3(x)}{(x^2+x+1)^2}=\frac{Ax+B}{x^2+x+1}+\frac{Cx+D}{(x^2+x+1)^2},$$

$$\frac{p_5(x)}{(x-1)^2(x^2+x+1)^2}$$
$$=\frac{A}{x-1}+\frac{B}{(x-1)^2}+\frac{Cx+D}{x^2+x+1}+\frac{Ex+F}{(x^2+x+1)^2}.$$

【常见分式的分项式】使用部分分式法时，计算量较大. 对一些典型情形，往往直接将分子变形以得到分项公式. 例如：

$$\frac{1}{1-x^2}=\frac{1}{2}\frac{1-x+1+x}{(1-x)(1+x)}=\frac{1}{2}\left(\frac{1}{1+x}+\frac{1}{1-x}\right),$$

$$\frac{x^2}{1+x^2}=\frac{1+x^2-1}{1+x^2}=1-\frac{1}{1+x^2}.$$

常见题型·应对

较少出现单独考察不定积分法的试题，因此掌握基本的积分法便可以. 在定积分、重积分、线面积分、微分方程中都会用到不定积分法.

【常用不定积分表】

$$\int k\mathrm{d}x=kx+C;$$

$$\int x^\alpha\mathrm{d}x=\frac{x^{\alpha+1}}{\alpha+1}+C\quad(\alpha\neq-1);$$

$$\int\frac{1}{x}\mathrm{d}x=\ln|x|+C;$$

$$\int a^x\mathrm{d}x=\frac{a^x}{\ln a}+C\quad(0<a\neq1);$$

$$\int \mathrm{e}^x\mathrm{d}x=\mathrm{e}^x+C;$$

$$\int \ln x\mathrm{d}x=x\ln x-x+C;$$

$$\int \cos x\mathrm{d}x=\sin x+C;$$

$$\int \sin x \mathrm{d}x = -\cos x + C;$$

$$\int \frac{1}{\cos^2 x} \mathrm{d}x = \tan x + C;$$

$$\int \frac{1}{\sin^2 x} \mathrm{d}x = -\cot x + C;$$

$$\int \sec x \tan x \mathrm{d}x = \sec x + C;$$

$$\int \csc x \cot x \mathrm{d}x = -\csc x + C;$$

$$\int \frac{\mathrm{d}x}{\sqrt{1+x^2}} = \ln(x + \sqrt{1+x^2}) + C;$$

$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x + C;$$

$$\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + C.$$

表中C为任意常数.

3.2　定积分·广义积分

大纲要求

理解定积分的概念.掌握定积分的性质及定积分中值定理,掌握换元积分法与分部积分法.理解积分上限的函数,会求它的导数,掌握牛顿-莱布尼兹公式,了解广义积分的概念并会计算广义积分.

知识点·点拨

【定积分定义】有界函数$f(x)$在闭区间$[a,b]$上的定积分是以计算曲边梯形面积或变速运动的路程为背景问题的一个数学概念.其定义式为

$$\int_a^b f(x)\mathrm{d}x=\lim_{\lambda\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i,$$

其中
$$a=x_0<x_1<\cdots<x_n=b,$$
$$x_i\leqslant\xi_i\leqslant x_{i+1},\Delta x_i=x_{i+1}-x_i$$
$$(i=0,1,2,\cdots,n),\quad \lambda=\max\{\Delta x_i\,|\,0\leqslant i\leqslant n\}.$$

当上述极限存在时，称 $f(x)$ 在区间 $[a,b]$ 上可积，积分值即该极限值.

【可积函数】狄利克雷函数是不可积函数，但是以下函数都是可积函数：

(1) $[a,b]$ 上的分段连续函数（即只有有限个第一类间断点的函数）；

(2) $[a,b]$ 上分段单调的函数（指函数曲线由有限个单调段构成）.

【定积分的整体性】定积分的可积性和积分值的大小不会因为个别点的函数值的改变而受影响.

【定积分的基本性质】

〖分项积分公式〗（和的积分等于积分的和）

$$\int_a^b (f(x)+g(x))\mathrm{d}x=\int_a^b f(x)\mathrm{d}x+\int_a^b g(x)\mathrm{d}x.$$

〖常数因子〗设 α 是常数，则

$$\int_a^b \alpha f(x)\mathrm{d}x=\alpha\int_a^b f(x)\mathrm{d}x.$$

〖分段积分公式〗设 $a<c<b$，则

$$\int_a^b f(x)\mathrm{d}x=\int_a^c f(x)\mathrm{d}x+\int_c^b f(x)\mathrm{d}x,$$

当 $a<b<c$ 或 $c<a<b$ 时分段积分公式还是成立.

〖比较定理〗设 $a<b$，则

$$f(x)\leqslant g(x)\Rightarrow\int_a^b f(x)\mathrm{d}x\leqslant\int_a^b g(x)\mathrm{d}x.$$

〖估界定理〗设 $m\leqslant f(x)\leqslant M(a\leqslant x\leqslant b)$,则

$$m(b-a)\leqslant\int_a^b f(x)\mathrm{d}x\leqslant M(b-a).$$

〖积分的绝对值〗设 $a<b$,则

$$\left|\int_a^b f(x)\mathrm{d}x\right|\leqslant\int_a^b |f(x)|\mathrm{d}x.$$

【牛顿-莱布尼兹公式】设 $F(x)$ 是连续函数 $f(x)$ $(a\leqslant x\leqslant b)$ 的一个原函数,则

$$\int_a^b f(x)\mathrm{d}x=F(b)-F(a).$$

〖点拨〗若 $f(x)$ 是分段连续函数,则可以首先使用分段积分公式,再使用以上公式计算积分值,被积函数在端点的第一类间断性可以忽略.

【定积分的换元积分法】令 $x=\varphi(t)$($\varphi'(t)$ 连续且非零),则

$$\int_a^b f(x)\mathrm{d}x=\int_\alpha^\beta f(\varphi(t))\varphi'(t)\mathrm{d}t,$$

其中 $t=\alpha$ 时,$x=\alpha$;$t=\beta$ 时,$x=b$.

【定积分的分部积分公式】若 $u'(x)$ 较 $u(x)$ 简单,$v'(x)$ 不比 $v(x)$ 复杂,则

$$\int_a^b u(x)\mathrm{d}v(x)=u(x)v(x)\Big|_a^b-\int_a^b v(x)u'(x)\mathrm{d}x.$$

【变上限积分】设 $f(x)$ 在区间 $[a,b]$ 上连续,对任意 $x\in[a,b]$,称 $[a,x]$ 上的定积分

$$F(x)=\int_a^x f(t)\mathrm{d}t\quad(a\leqslant x\leqslant b)$$

为变上限积分. 它是$[a,b]$上的一个函数.

【变上限积分的导数】

$$F'(x)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\mathrm{d}t\right)=f(x),\quad a\leqslant x\leqslant b.$$

【复合变限积分的导数】

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{a(x)}^{b(x)} f(t)\mathrm{d}t\right)=f(b(x))b'(x)-f(a(x))a'(x).$$

〖点拨〗变限积分是考试重点,求导公式务必熟练掌握. 当被积式中含有 x 时便不能直接应用以上求导公式,而要先将 x 从被积式中提出来或代换掉后再用上述求导公式.

例　求 $F(x)=\int_0^x (x+t)f(t)\mathrm{d}t$ 的导数.

解　$F(x)=x\int_0^x f(t)\mathrm{d}t+\int_0^x tf(t)\mathrm{d}t$,

$$\begin{aligned}F'(x)&=\int_0^x f(t)\mathrm{d}t+xf(x)+xf(x)\\&=\int_0^x f(t)\mathrm{d}t+2xf(x).\end{aligned}$$

【奇偶函数的定积分】若函数 $f(x)$是区间$[-a,a]$上连续的奇函数,则

$$\int_{-a}^{a} f(x)\mathrm{d}x=0;$$

若函数 $f(x)$是区间$[-a,a]$上连续的偶函数,则

$$\int_{-a}^{a} f(x)\mathrm{d}x=2\int_0^a f(x)\mathrm{d}x.$$

【周期函数的定积分】若函数 $f(x)$是周期为 T 的区间$(-\infty,+\infty)$上连续周期函数,则对任意实数 a,有

$$\int_0^T f(x)\mathrm{d}x=\int_a^{a+T} f(x)\mathrm{d}x.$$

【积分中值定理】设在区间$[a,b]$上函数$f(x)$连续，$g(x)$可积且不变号，则存在$\xi\in[a,b]$，使得

$$\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x.$$

【积分中值定理· 特殊情形】设$f(x)$在$[a,b]$上连续，则存在$\xi\in(a,b)$，使得

$$\int_a^b f(x)\mathrm{d}x=f(\xi)(b-a).$$

〖点拨〗使用微分中值定理证明此公式，可以将一般公式中的$\xi\in[a,b]$加强为$\xi\in(a,b)$.

【柯西积分不等式】设函数$f(x)$，$g(x)$在区间$[a,b]$上连续，则

$$\left[\int_a^b f(x)g(x)\mathrm{d}x\right]^2\leqslant\int_a^b f^2(x)\mathrm{d}x\int_a^b g^2(x)\mathrm{d}x.$$

【$[a,+\infty)$上的广义积分】设$f(x)$在$[a,+\infty)$上有定义. 若对任意$b>a$，$f(x)$在$[a,b]$上可积，则极限

$$l=\lim_{b\to+\infty}\int_a^b f(x)\mathrm{d}x\quad\left(\text{记作}\int_a^{+\infty}f(x)\mathrm{d}x\right)$$

称为$f(x)$在无穷区间$[a,+\infty)$上的广义积分，l为有限时说广义积分收敛，否则说广义积分发散.

类似地定义$(-\infty,b]$上的广义积分，记作

$$\int_{-\infty}^b f(x)\mathrm{d}x=\lim_{a\to-\infty}\int_a^b f(x)\mathrm{d}x.$$

【$(-\infty,+\infty)$上的广义积分】设$f(x)$在$(-\infty,+\infty)$上有定义. 若对某个c，$\int_{-\infty}^c f(x)\mathrm{d}x$及$\int_c^{+\infty}f(x)\mathrm{d}x$都收敛，则说$f(x)$在$(-\infty,+\infty)$上的广义积分收敛，记作

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{-\infty}^{c} f(x)\mathrm{d}x+\int_{c}^{+\infty} f(x)\mathrm{d}x.$$

〖点拨〗若 $\int_{-\infty}^{c} f(x)\mathrm{d}x$ 及 $\int_{c}^{+\infty} f(x)\mathrm{d}x$ 中有一个发散，则称广义积分 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$ 发散. 例如，$\int_{-\infty}^{+\infty} \frac{1}{x}\mathrm{d}x$ 是发散的，因为 $\int_{1}^{+\infty} \frac{1}{x}\mathrm{d}x=+\infty$.

【函数的奇点】 设 $c\in[a,b]$. 若从区间内的单侧极限 $f(c^{+})=\infty$ 或 $f(c^{-})=\infty$，则称 c 是函数 $f(x)$ 在 $[a,b]$ 上的奇点.

【上限为奇点的广义积分】 设 $f(x)$ 在 $[a,b)$ 上有定义，$f(b^{-})=\infty$. 若对任意 $\beta\in(a,b)$，$f(x)$ 在 $[\alpha,\beta]$ 上可积，则极限

$$l=\lim_{\beta\to b^{-}}\int_{\alpha}^{\beta} f(x)\mathrm{d}x \quad \left(\text{记作}\int_{a}^{b} f(x)\mathrm{d}x\right)$$

称为无界函数在 $[a,b]$ 上的广义积分，l 为有限时，说广义积分收敛，否则说广义积分发散.

类似地定义当 $f(a^{+})=\infty$ 时的广义积分，记作

$$\int_{a}^{b} f(x)=\lim_{\alpha\to a^{+}}\int_{\alpha}^{b} f(x)\mathrm{d}x.$$

【内部有奇点的广义积分】 设 $c\in(a,b)$ 是 $f(x)$ 的奇点，则当广义积分 $\int_{a}^{c} f(x)\mathrm{d}x$ 及 $\int_{c}^{b} f(x)\mathrm{d}x$ 都收敛时，说 $f(x)$ 在 $[a,b]$ 上的广义积分收敛，记作

$$\int_{a}^{b} f(x)\mathrm{d}x=\int_{a}^{c} f(x)\mathrm{d}x+\int_{c}^{b} f(x)\mathrm{d}x.$$

〖点拨〗若 $\int_{a}^{c} f(x)\mathrm{d}x$ 及 $\int_{c}^{b} f(x)\mathrm{d}x$ 中有一个发散，则称广义积分 $\int_{a}^{b} f(x)\mathrm{d}x$ 发散. 例如，$\int_{-1}^{1} \frac{1}{x}\mathrm{d}x$ 是

发散的,因为$\int_0^1 \frac{1}{x}\mathrm{d}x=-\infty$.

【广义积分的计算公式】 设$F'(x)=f(x)$,则

$$\int_a^{+\infty} f(x)\mathrm{d}x=F(+\infty)-F(a),$$

$$\int_a^b f(x)\mathrm{d}x=F(b^-)-F(a)\quad (b\text{ 为奇点}).$$

例 $\int_1^{+\infty}\mathrm{e}^{-2x}\mathrm{d}x=\left(-\frac{1}{2}\mathrm{e}^{-2x}\right)\Big|_1^{+\infty}=\frac{1}{2}$.

【广义积分的换元积分法】 换元的目的是化简函数形式或调整积分区间,因此遵循与定积分换元法相同的原则:

$$\int_a^b f(x)\mathrm{d}x=\int_\alpha^\beta f(\varphi(t))\varphi'(t)\mathrm{d}t,$$

其中,$x=\varphi(t)$,$t\to\alpha^+$时,$x\to a$;$t\to\beta^-$时,$x\to b$.

【广义积分的分部积分公式】

$$\int_a^b u(x)\mathrm{d}v(x)=u(x)v(x)\Big|_a^b-\int_a^b v(x)u'(x)\mathrm{d}x,$$

其中 $u(x)v(x)\Big|_a^b=u(b^-)v(b^-)-u(a^+)v(a^+)$.

〖点拨〗当极限式$u(x)v(x)\Big|_a^b$不存在时,不可以推断$\int_a^b u(x)\mathrm{d}v(x)$发散,此时应当将积分的原函数求出后再代入上下限判定.

例 计算广义积分$I=\int_0^1 \frac{\ln x}{(1+x)^2}\mathrm{d}x$.

解 奇点是0,使用分部积分法,得

$$I=-\int_0^1 \ln x\mathrm{d}\,\frac{1}{1+x}=-\frac{\ln x}{1+x}\Big|_{0^+}^1+\int\frac{1}{1+x}\mathrm{d}\ln x,$$

虽然极限 $-\left.\dfrac{\ln x}{1+x}\right|_{0^+}^{1}$ 不收敛，也不能断定积分发散，而应当在原函数完全求出来之后再考虑积分限：

$$
\begin{aligned}
F(x)&=-\int\ln x\mathrm{d}\,\frac{1}{1+x}=-\frac{\ln x}{1+x}+\int\frac{\mathrm{d}x}{(1+x)x}\\
&=-\frac{\ln x}{1+x}+\ln x-\ln(1+x)\\
&=\frac{x\ln x}{1+x}-\ln(1+x),
\end{aligned}
$$

由于 $\lim\limits_{x\to0^+}x\ln x=0$，故 $I=F(x)\Big|_{0^+}^{1}=-\ln 2$.

常见题型·应对

定积分计算作为基本运算主要出现在多元函数的积分、微分方程等问题中. 但是也有一些专门的题型和考点.

【变限积分的问题概述】涉及变限积分的试题较多. 将变限积分视为一个普通的函数，考察其极限、连续性、导数、几何性质、不等式等问题.

应对要点主要是掌握求导公式，然后按照标准的方法求解.

例1　求正常数 a,b，使得

$$\lim_{x\to0}\frac{1}{bx-\sin x}\int_0^x\frac{t^2}{\sqrt{a+t^2}}\mathrm{d}t=1.$$

解　使用洛必达法则，得 $\lim\limits_{x\to0}\dfrac{\dfrac{2x}{\sqrt{a+x^2}}}{b-\cos x}=1$，于是推出 $b=1$，进而使用等价代换求得 $a=4$.

例2　函数 $F(x)=\displaystyle\int_1^x\left(2-\frac{1}{\sqrt{t}}\right)\mathrm{d}t\,(x>0)$ 的单

调减少区间为________.

解　求得导数$F'(x)=2-\dfrac{1}{\sqrt{x}}(x>0)$,便可求得单调减少区间为$\left(0,\dfrac{1}{4}\right)$.

【分段函数的定积分计算】考察分段积分公式和牛顿-莱布尼兹公式,在题给的分段点分段积分即可.

例　设$f(x)=\begin{cases}1+x^2, & x\leqslant 0,\\ \mathrm{e}^{-x}, & x>0,\end{cases}$求定积分

$$I=\int_1^3 f(x-2)\mathrm{d}x.$$

解　代换$x-2=t$,得$I=\int_{-1}^1 f(t)\mathrm{d}t$,分段代入被积函数计算便可,即

$$I=\int_{-1}^0(1+t^2)\mathrm{d}t+\int_0^1\mathrm{e}^{-t}\mathrm{d}t=\frac{7}{3}-\frac{1}{\mathrm{e}}.$$

【证明积分恒等式】解法思路有两种:

(1)通过分部积分或变量代换改变定积分的被积表示式或积分限.

(2)视为变限积分函数的问题,转化为导数的应用(参见2.2节).

例1　设$f(x)$处处连续,证明

$$\int_0^a x^3 f(x^2)\mathrm{d}x=\frac{1}{2}\int_0^{a^2}xf(x)\mathrm{d}x.$$

证1　以积分上限a为变量,设置辅助函数

$$h(t)=\int_0^t x^3 f(x^2)\mathrm{d}x-\frac{1}{2}\int_0^{t^2}xf(x)\mathrm{d}x,$$

则
$$h'(t)=t^3 f(t^2)-\frac{1}{2}t^2 f(t^2)\cdot 2t=0,$$

故 $h(t)=h(0)=0$, **从而** $h(a)=0$.

证2 使用换元法,令 $x^2=t$,得

$$\text{左}=\frac{1}{2}\int_0^a x^2 f(x^2)\mathrm{d}x^2=\text{右}.$$

例2 设 $f'(x)$ 处处连续,证明

$$\int_0^a f(x)\mathrm{d}x=af(0)+\int_0^a (a-x)f'(x)\mathrm{d}x.$$

证 对出现导数的积分,使用分部积分法,

$$\begin{aligned}&\int_0^a (a-x)f'(x)\mathrm{d}x\\&=\int_0^a (a-x)\mathrm{d}f(x)\\&=(a-x)f(x)\Big|_0^a+\int_0^a f(x)\mathrm{d}x\\&=-af(0)+\int_0^a f(x)\mathrm{d}x.\end{aligned}$$

【证明积分不等式】解法思路有三种:

(1)通过定积分估计性质来分析大小.

(2)若不等式 $\int_a^b f(x)\mathrm{d}x\geqslant 0$ 与积分限的大小无关,则等价于证明函数不等式

$$F(t)=\int_a^t f(x)\mathrm{d}x\geqslant 0,$$

借用 $F'(t)$ 的符号与 $F(a)$ 或 $F(b)$ 之值可以证明.

(3)借助重要不等式,例如柯西不等式

$$\left[\int_a^b f(x)g(x)\mathrm{d}x\right]^2\leqslant\int_a^b f^2(x)\mathrm{d}x\int_a^b g^2(x)\mathrm{d}x.$$

例1 设 $f'(x)$ 在 $(0,a)$ 内连续, $f(0)=0$, $|f'(x)|\leqslant M$,证明 $\left|\int_0^a f(x)\mathrm{d}x\right|\leqslant\frac{M}{2}a^2$.

证　因 $f(x)=f(0)+f'(\xi)x\leqslant Mx$,故

$$\left|\int_0^a f(x)\mathrm{d}x\right|\leqslant\int_0^a|f(x)|\mathrm{d}x\leqslant M\int_0^a x\mathrm{d}x$$
$$=\frac{1}{2}Ma^2.$$

例 2　设 $f(x)$ 在 $[a,b]$ 上连续,证明

$$\left[\int_a^b f(x)\mathrm{d}x\right]^2\leqslant(b-a)\int_a^b f^2(x)\mathrm{d}x.$$

证 1　利用柯西不等式(视 $g=1$)得

$$\left[\int_a^b f(x)\mathrm{d}x\right]^2\leqslant\int_a^b f^2\mathrm{d}x\int_a^b 1\mathrm{d}x=(b-a)\int_a^b f^2(x)\mathrm{d}x.$$

证 2　移项之后,以积分上限 b 为变量,设

$$F(t)=(t-a)\int_a^t f^2(x)\mathrm{d}x-\left[\int_a^t f(x)\mathrm{d}x\right]^2,$$

则
$$F'(t)=\int_a^t f^2(x)\mathrm{d}x+(t-a)f^2(t)$$
$$-2\int_a^t f(x)\mathrm{d}x\cdot f(t)$$
$$=\int_a^t[f^2(x)+f^2(t)-2f(t)f(x)]\mathrm{d}x$$
$$=\int_a^t[f(x)-f(t)]^2\mathrm{d}x\geqslant 0,$$

又 $F(a)=0$,故从 $F(t)$ 在 $[a,b]$ 上连续知 $F(t)\geqslant 0$,于是 $F(b)\geqslant 0$ 即为所证.

【积分中值问题】通常是积分中值定理、介值定理和罗尔定理的联合使用.

例 1　设 $f(x)$ 在 $[a,b]$ 上连续,且满足

$$\int_a^b f(x)\mathrm{d}x=0,\quad \int_a^b xf(x)\mathrm{d}x=0,$$

证明:存在 $\xi,\eta\in(a,b)$,使得 $f(\xi)=f(\eta)=0$.

证　反证法.如果 $f(x)$ 只有一个根 ξ,则 $f(x)$ 必

然在ξ两侧反号,于是构建函数$g(x)=x-\xi$,利用条件得

$$\int_a^b g(x)f(x)\mathrm{d}x=0,$$

但另一方面,除$x=\xi$外,$g(x)f(x)$在$[a,b]$上恒正或恒负,从而

$$\int_a^b g(x)f(x)\mathrm{d}x\neq 0,\quad \text{矛盾}.$$

例 2　设$f(x)$在闭区间$[a,b]$上连续,且满足$f(a)\leqslant f(x)\leqslant f(b)$,证明:存在$c\in[a,b]$,使得

$$\int_a^b f(x)\mathrm{d}x=f(a)(c-a)+f(b)(b-c).$$

证　记$F(x)=f(a)(x-a)+f(b)(b-x)$,则有$F(b)\leqslant\int_a^b f(x)\mathrm{d}x\leqslant F(a)$,对$F(x)$使用介值定理即知,有$c\in[a,b]$,使$F(c)=\int_a^b f(x)\mathrm{d}x$.

【使用定积分求和式极限】(参见1.2节)

【瓦里斯公式】

$$\int_0^{\pi/2}\sin^n x\mathrm{d}x=\int_0^{\pi/2}\cos^n x\mathrm{d}x=\begin{cases}\dfrac{(n-1)!!}{n!!}, & n\text{为奇数},\\[2mm] \dfrac{(n-1)!!}{n!!}\dfrac{\pi}{2}, & n\text{为偶数},\end{cases}$$

其中,$5!!=5\cdot 3\cdot 1$, $6!!=6\cdot 4\cdot 2$,等等.

【常用定积分代换公式】

$$\int_0^{\pi}xf(\sin x)\mathrm{d}x=\frac{\pi}{2}\int_0^{\pi}f(\sin x)\mathrm{d}x;$$

$$\int_0^a f(x)\mathrm{d}x=\int_0^a f(a-x)\mathrm{d}x;$$

$\int_0^a f(x)\mathrm{d}x=\frac{1}{2}\int_0^a (f(x)+f(a-x))\mathrm{d}x$;

$\int_{-a}^a f(x)\mathrm{d}x=\int_0^a (f(x)+f(-x))\mathrm{d}x$.

【广义积分公式】

$$\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x=\frac{\sqrt{\pi}}{2},\quad \int_0^{+\infty} \frac{\sin x}{x}\mathrm{d}x=\frac{\pi}{2}.$$

3.3 定积分应用

大纲要求

掌握用定积分表达和计算一些几何量与物理量(平面图形的面积、平面曲线的弧长、旋转体的体积及侧面积、平行截面面积为已知的立体体积、功、引力、压力、质心等)及函数的平均值.

知识点·点拨

【定积分的几何应用】弧长和侧面积计算由于积分中多出现根式,不易计算,故试题较少涉及;面积计算与体积计算的问题在考研题中出现频率较高.结合待定参数,优化问题的综合题常常出现.

【平面曲线的弧长 s】

(1)**设曲线 L 的方程为** $y=y(x),a\leqslant x\leqslant b$,**则**

$$s=\int_a^b \sqrt{1+y'^2(x)}\mathrm{d}x.$$

(2)**设曲线 L 的方程为** $r=r(\theta),a\leqslant\theta\leqslant b$,**则**

$$s=\int_a^b \sqrt{r^2(\theta)+r'^2(\theta)}\mathrm{d}\theta.$$

(3)**设曲线 L 的方程为** $\begin{cases}x=x(t),\\ y=y(t),\end{cases} a\leqslant t\leqslant b$,**则**

$$s=\int_a^b \sqrt{x'^2(t)+y'^2(t)}\mathrm{d}t.$$

例 1　**求曲线** $x^{\frac{2}{3}}+y^{\frac{2}{3}}=a^{\frac{2}{3}}$ **的弧长** s.

解　**由对称性,只需考虑曲线在第一象限内的一段,尽管曲线方程中** y **不易解出,但利用隐函数求导法和等量替换可以求得**

$$(x^{\frac{2}{3}}+y^{\frac{2}{3}})'=\frac{2}{3}x^{-\frac{1}{3}}+\frac{2}{3}y^{-\frac{1}{3}}y'=0,$$

从而

$$y'=-\left(\frac{y}{x}\right)^{\frac{1}{3}},$$

$$1+y'^2(x)=\frac{x^{\frac{2}{3}}+y^{\frac{2}{3}}}{x^{\frac{2}{3}}}=x^{-\frac{2}{3}}a^{\frac{2}{3}},$$

故　$s=4\int_0^a \sqrt{1+y'^2(x)}\mathrm{d}x=4\int_0^a a^{\frac{1}{3}}x^{-\frac{1}{3}}\mathrm{d}x=6a.$

例 2　**求曲线** $x=a(t-\sin t)\ y=a(1-\cos t)$ **的弧长** s $(0\leqslant t\leqslant 2\pi)$.

解　$x'(t)=a(1-\cos t)$,　$y'(t)=a\sin t$,**于是**

$$\begin{aligned}\sqrt{x'^2(t)+y'^2(t)}&=a\sqrt{(1-\cos t)^2+\sin^2 t}\\&=a\sqrt{2(1-\cos t)}\\&=2a|\sin(t/2)|,\end{aligned}$$

$$s=\int_0^{2\pi}2a|\sin(t/2)|\mathrm{d}t=\int_0^{2\pi}\sin(t/2)\mathrm{d}t=8a.$$

注意　**凡是从平方根式中开出的结果均应非负,绝对值记号的解除要依据函数在积分区间的符号而定.**

例 3　**求心脏线** $r=a(1+\cos\theta)$ **的弧长** s.

解　$\sqrt{r^2(\theta)+r'^2(\theta)}=2a|\cos(\theta/2)|$,**利用对**

称性,可以只考虑曲线在上半平面的一段,故有

$$s=2\int_0^{\pi}2a\cos(\theta/2)\mathrm{d}\theta=8a.$$

【平面区域的面积S】如图3-3-1所示,设D为x-型区域:$\begin{cases}y_1(x)\leqslant y\leqslant y_2(x),\\a\leqslant x\leqslant b,\end{cases}$则$D$的面积为

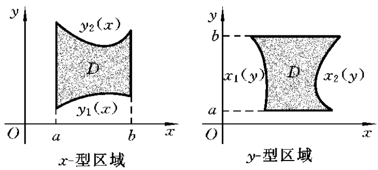


图 3-3-1

$$S=\int_a^b(y_2(x)-y_1(x))\mathrm{d}x,$$

设D为y-型区域:$\begin{cases}x_1(y)\leqslant x\leqslant x_2(y),\\a\leqslant y\leqslant b,\end{cases}$则

$$S=\int_a^b(x_2(y)-x_1(y))\mathrm{d}y.$$

【平行截面体的体积V】设沿着x轴方向,截面体的截面积是$S(x)$,$a\leqslant x\leqslant b$,则

$$V=\int_a^b S(x)\mathrm{d}x.$$

【绕x轴的旋转体的体积V与侧面积S】如图3-3-2所示,设旋转体由D:$\begin{cases}0\leqslant y\leqslant y(x),\\a\leqslant x\leqslant b\end{cases}$绕$x$轴旋转而得,则

$$V=\int_a^b \pi y^2(x)\mathrm{d}x \quad \text{（截面是圆）},$$

$$S=\int_a^b 2\pi y(x)\sqrt{1+y'^2(x)}\mathrm{d}x.$$

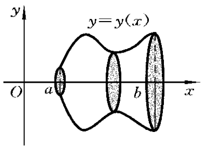


图 3-3-2

【绕 y 轴的旋转体的体积 V】如图 3-3-3 所示，设旋转体由 $D:\begin{cases}0\leqslant y\leqslant y(x),\\0\leqslant a\leqslant x\leqslant b\end{cases}$ 绕 y 轴旋转而得，则

$$V=\int_a^b 2\pi xy(x)\mathrm{d}x \quad \text{（截面是圆柱面）}.$$

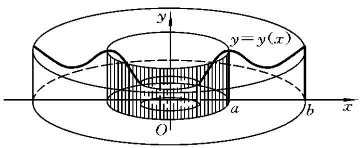


图 3-3-3

【古鲁津定律】如图 3-3-4 所示，平面图形 D 绕同一平面内 D 外的直线 L 旋转，所得旋转体的体积 V 及表面积 S 分别为

$$V=2\pi R\cdot\sigma,\quad S=2\pi R\cdot s.$$

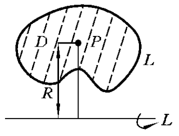


图 3-3-4

其中 R 是 D 的质心 P 到 L 的距离，σ 与 s 分别是 D 的面积与边界曲线的周长.

【定积分的物理应用概述】积分在物理学中的应用主要有以下两种：

(1)计算沿直线移动变力做功，计算沿直线移动变速运动路程，计算静压力；

(2)计算质量、重心、转动惯量、沿曲线移动做功、通过曲面一侧的流量.

其中(1)适合于用定积分表现，(2)则适合于多元函数的积分表现. 在定积分的物理应用中，变力做功是常考问题.

【变力做功】设物体 Ω 移动的路程区间是 $[a,b]$，若能求出在 $[x,x+\mathrm{d}x]$ 段上移动 Ω 所做的功 $\mathrm{d}W$，则

$$W=\int_a^b \mathrm{d}W.$$

例 1　弹簧从原长压缩 1 m 需做功 60 J，在弹性限度内，再压缩 1 m 需做功多少？

解　由虎克定律，弹性力 f 与压缩长度 x 成比例，即 $f(x)=kx$. 于是，在路程段 $[x,x+\mathrm{d}x]$ 中，系统做功为 $\mathrm{d}W=f(x)\mathrm{d}x$，故所求为

$$W=\int_1^2 f(x)\mathrm{d}x=\int_1^2 kx\mathrm{d}x=\frac{3}{2}k,$$

由条件知 $60=\int_0^1 kx\mathrm{d}x$，即 $k=120$，从而 $W=180$ (J).

例2　使用抓斗从井底提升污泥到井口. 已知井深30 m，抓斗自重400 N，缆绳每米重50 N，抓斗提升速度为 3 m/s，抓起的污泥重 2000 N，提升过程中污泥以 20 N/s 的速率从抓斗中漏出，求系统克服重力做功的大小.

解　以提升距离 x 为积分变量，则在路程段$[x,x+\mathrm{d}x]$$(0\leqslant x\leqslant 30)$中，系统做功为

$$\mathrm{d}W=\left[400+(30-x)\times 50+\left(2000-\frac{x}{3}\times 20\right)\right]\mathrm{d}x,$$

$$W=\int_0^{30}\mathrm{d}W=91500\ (\mathrm{J}).$$

【微元法计算做功】设物体 Ω 分布的位置区间是$[a,b]$. 若能求出区间$[x,x+\mathrm{d}x]$上对应的$\mathrm{d}\Omega$ 被移动时的做功 $\mathrm{d}W$，则 $W=\int_a^b \mathrm{d}W$.

例　如图3-3-5 所示，从平地堆出一个底半径为 r，高为 h 的圆锥形沙堆，求做功大小 W(设沙子密度为 ρ).

解　使用微元法，以沙堆的高度 x 为积分变量，有 $0\leqslant x\leqslant h$，于是堆积高度范围为$[x,x+\mathrm{d}x]$的沙层的做功微元为

$$\mathrm{d}W=gx\mathrm{d}m=gx\pi\rho\left(\frac{h-x}{h}r\right)^2\mathrm{d}x,$$

故
$$W=\int_0^h \mathrm{d}W=\frac{g\rho\pi r^2}{h^2}\int_0^h x(h-x)^2\mathrm{d}x$$

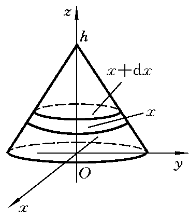


图 3-3-5

$$=\frac{1}{12}g\pi\rho h^2r^2.$$

【变速运动路程计算】设物体以速度 $v(t)$ 沿直线移动,则其在时段 $a\leqslant t\leqslant b$ 内的移动路程为

$$s=\int_a^b v(t)\mathrm{d}t.$$

【变压强下压力计算】设物体 Ω 承受的压强大小是 $f(h)$,$a\leqslant h\leqslant b$. 且在 $[h,h+\mathrm{d}h]$ 段所对应的部分受压力为 $\mathrm{d}F=f(h)\mathrm{d}\sigma$($\mathrm{d}\sigma$ 为受压面积微元),则物体 Ω 所受压力为

$$F=\int_a^b \mathrm{d}F.$$

【函数的平均值】函数 $f(x)$($a\leqslant x\leqslant b$)在区间 $[a,b]$ 上的平均值是:$\overline{f}=\dfrac{1}{b-a}\displaystyle\int_a^b f(x)\mathrm{d}x$.

常见题型 · 应对

将定积分计算与微分学的极值问题相结合产生的综合题较多. 单独考察定积分的几何应用(面积、

体积),物理应用(变力做功、质心、质量等问题)也有一些.以下举例说明.

【面积计算结合极值问题举例】

例　求 $a(0<a<1)$,使曲线 $y=x^2$, $y=ax$ 与 $x=1$ 所围的两块面积之和最小.

解　如图 3-3-6 所示,所求总面积为

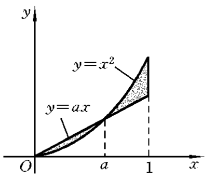


图 3-3-6

$$
\begin{aligned}
S &= \int_0^a (ax-x^2)\mathrm{d}x + \int_a^1 (x^2-ax)\mathrm{d}x \\
&= \frac{1}{2}a^3 - \frac{1}{3}a^3 + \frac{1}{3}(1-a^3) - \frac{a}{2}(1-a^2) \\
&= \frac{1}{3} - \frac{1}{2}a + \frac{1}{3}a^3.
\end{aligned}
$$

由 $S'(a)=-\frac{1}{2}+a^2$, $S''(a)=2a>0$ 知,驻点 $a=\frac{\sqrt{2}}{2}$ 即为所求.

【体积计算结合极值问题举例】

例　设平面区域 D_1, D_2 如图 3-3-7 所示, $0<a<2$,(1)求 D_1 绕 x 轴旋转所成体积 V_1; D_2 绕 y 轴旋转所成体积 V_2.(2)a 为何值时, $V=V_1+V_2$ 最大? 求

$V_{\max}$.

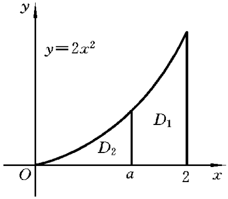


图 3-3-7

解 (1)直接套用旋转体体积公式,得

$$V_1=\int_a^2 \pi y^2(x)\mathrm{d}x=\frac{4\pi}{5}(32-a^5),$$

$$V_2=\int_0^a 2\pi xy(x)\mathrm{d}x=\int_0^a 2\pi x\cdot 2x^2\mathrm{d}x=\pi a^4.$$

(2)$V'(a)=4\pi a^3-4\pi a^4=4\pi a^3(1-a)$在 $a=1$ 两侧左正右负,故 $a=1$ 时 V 取最大值,$V_{\max}=\dfrac{129}{5}\pi$.

第4章　向量代数和空间解析几何

大纲要求

理解空间直角坐标系,理解向量的概念及其表示.掌握向量的运算(线性运算、数量积、向量积、混合积),了解两个向量垂直、平行的条件.理解单位向量、方向数与方向余弦、向量的坐标表达式,掌握用坐标表达式进行向量运算的方法.掌握平面方程和直线方程及其求法.会求平面与平面、平面与直线、直线与直线之间的夹角,并会利用平面、直线的相互关系(平行、垂直、相交等)解决有关问题.会求点到直线以及点到平面的距离.了解曲面方程和空间曲线方程的概念.了解常用二次曲面的方程及其图形,会求以坐标轴为旋转轴的旋转曲面及母线平行于坐标轴的柱面方程.了解空间曲线的参数方程和一般方程.了解空间曲线在坐标平面上的投影,并会求其方程.

注　数学二不要求本章内容.

4.1　向 量 代 数

知识点·点拨

【空间直角坐标系】三根交于一点且两两垂直的数轴构成一个所在空间的直角坐标系.按照右手法则依次给这三个数轴命名为x轴、y轴、z轴.每两个数

轴构成的平面称为坐标平面,命名为 Oxy 面、Oyz 面、Ozx 面. 这三个平面将三维空间划分为八个卦限. 通常用 $Oxyz$ 表示空间直角坐标系(见图 4-1-1).

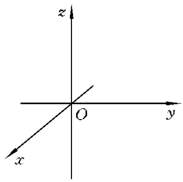


图 4-1-1

【两个点的距离】点 $P(x_1,y_1,z_1)$ 与点 $Q(x_2,y_2,z_2)$ 之间的距离为

$$\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$$

【向量】向量(也称为矢量)是从一类有大小、有方向并满足一定运算规则的物理量和几何量中抽象出来的量.

【向量的几何表示】向量可以用空间中从点 A 到点 B 的有向线段来表示,记作 $\overrightarrow{AB}$,也可用黑体字母表示向量,如 $\boldsymbol{a}$,$\boldsymbol{b}$ 等.

【向量的两个要素:模和方向】有向线段 $\overrightarrow{AB}$ 的长度称为向量的模,记作 $|\overrightarrow{AB}|$,方向由 A 指向 B.

【向量的相等 · 可移动性】两个向量相等当且仅当它们的两个要素一样,即方向一致,并且模相等. 经

过平移后完全重合的两个向量是相同的向量.

【特殊的向量】

〖零向量〗其模为 0,规定它有任意方向,记作 $\mathbf{0}$.

〖单位向量〗其模为 1,这样的向量有无限个. 与 $\boldsymbol{a}$ 同向的单位向量记作 $\boldsymbol{a}^0$ 或 $\boldsymbol{e}_a$,$\boldsymbol{a}^0=\boldsymbol{a}/|\boldsymbol{a}|$.

〖负向量〗称 $\overrightarrow{BA}$ 为 $\overrightarrow{AB}$ 的负向量,记作 $\overrightarrow{BA}=-\overrightarrow{AB}$.

〖基向量〗三个坐标轴 Ox,Oy,Oz 上与坐标轴同向的单位向量,依次记作 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$.

【向量的和】两个向量 $\boldsymbol{a},\boldsymbol{b}$ 的"和"还是向量,记作 $\boldsymbol{a}+\boldsymbol{b}$,其定义如图 4-1-2 所示.

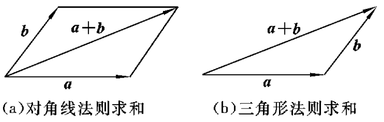


图 4-1-2

【向量的差】两个向量 $\boldsymbol{a},\boldsymbol{b}$ 的差定义为如下向量:

$$\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b}).$$

【向量的数乘】向量 $\boldsymbol{a}$ 与数量 λ 的乘积是一个向量,记作 $\lambda\boldsymbol{a}$,其模定义为 $|\lambda||\boldsymbol{a}|$,方向与向量 $\boldsymbol{a}$ 相同($\lambda>0$ 时)或相反($\lambda<0$ 时).

【向量的夹角】两个向量的起点重合之后两线段形成的较小的角度,如图 4-1-3 所示.

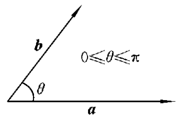


图 4-1-3

【向量的方向角·方向余弦】向量$\boldsymbol{a}$与三个基向量的夹角称为方向角,记作α,β,γ;方向角的余弦称为向量的方向余弦.

〖点拨〗任给三个角α,β,γ不一定能够成为某个向量的三个方向角.能够成为某个向量的三个方向角的充分必要条件是它们的余弦的平方和为1:

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1.$$

【向量的数量积·点乘】两个向量$\boldsymbol{a},\boldsymbol{b}$的数量积(也称为点乘)记作$\boldsymbol{a}\cdot\boldsymbol{b}$,是一个数量:

$$\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta,$$

其中θ是向量$\boldsymbol{a},\boldsymbol{b}$的夹角.$\boldsymbol{a}\cdot\boldsymbol{a}$可以记作$\boldsymbol{a}^2$.

【向量的向量积·叉乘】两个向量$\boldsymbol{a},\boldsymbol{b}$的向量积(也称为叉乘)记作$\boldsymbol{a}\times\boldsymbol{b}$,是一个向量:其模为$|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\theta$,方向垂直于$\boldsymbol{a},\boldsymbol{b}$所在的平面,依右手法则确定,如图4-1-4所示.

【三个向量的混合积】三个向量$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$的混合积是指既做点乘,又做叉乘的运算.例如$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$,它是一个数量,依次按照两种乘法的定义计算.

【向量运算的基本算律】

〖交换律〗$\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$,　$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$,

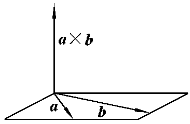


图 4-1-4

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}=(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}.$$

〖反交换律〗$\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}$.

〖分配律〗$k(\boldsymbol{a}+\boldsymbol{b})=k\boldsymbol{a}+k\boldsymbol{b}$,

$$\boldsymbol{c}\cdot(\boldsymbol{a}+\boldsymbol{b})=\boldsymbol{c}\cdot\boldsymbol{a}+\boldsymbol{c}\cdot\boldsymbol{b},$$

$$\boldsymbol{c}\times(\boldsymbol{a}+\boldsymbol{b})=\boldsymbol{c}\times\boldsymbol{a}+\boldsymbol{c}\times\boldsymbol{b}.$$

〖结合律〗$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$,

$$k(\boldsymbol{a}\cdot\boldsymbol{b})=(k\boldsymbol{a})\cdot\boldsymbol{b}=\boldsymbol{a}\cdot(k\boldsymbol{b}),$$

$$k(\boldsymbol{a}\times\boldsymbol{b})=(k\boldsymbol{a})\times\boldsymbol{b}=\boldsymbol{a}\times(k\boldsymbol{b}).$$

【使用向量运算判断非零向量的位置关系】

$$\boldsymbol{a}\parallel\boldsymbol{b}\Leftrightarrow\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0},$$

$$\boldsymbol{a}\parallel\boldsymbol{b}\Leftrightarrow\boldsymbol{a}=k\boldsymbol{b}(\text{共线}),$$

$$\boldsymbol{a}\perp\boldsymbol{b}\Leftrightarrow\boldsymbol{a}\cdot\boldsymbol{b}=0,$$

$$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\text{ 共面}\Leftrightarrow(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=0.$$

〖点拨〗由于向量是可以平行移动的,故向量的平行与共线是等价的;任何两个向量都是共面的.这一点与直线不相同.

【坐标系中的径向量】由原点 $O(0,0,0)$ 到点 $P(x,y,z)$的向量可以表示为三个基向量$\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 的线性组合$\overrightarrow{OP}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$,其坐标表示为 $\boldsymbol{r}=\{x,y,z\}$.

x,y,z称为径向量的坐标.

【坐标系中的任意向量】起点为$P(x_1,y_1,z_1)$,终点为$Q(x_2,y_2,z_2)$的向量的坐标表示为

$$\overrightarrow{PQ}=\{x_2-x_1,y_2-y_1,z_2-z_1\},$$

其中x_2-x_1,y_2-y_1,z_2-z_1称为该向量的坐标.

【使用坐标计算向量的模】设向量$\boldsymbol{a}=\{x,y,z\}$,则其模为$\sqrt{x^2+y^2+z^2}$.

【使用坐标计算向量的方向余弦】设$\boldsymbol{a}=\{x,y,z\}$,则将其单位化之后的单位向量$\boldsymbol{a}/|\boldsymbol{a}|$的三个分量$x/|\boldsymbol{a}|,y/|\boldsymbol{a}|,z/|\boldsymbol{a}|$便是它的三个方向余弦.

例 给定$A(0,1,-1),B(2,1,3)$两点,求向量$\overrightarrow{AB}$的坐标表示、模、单位向量、方向余弦.

解 $\overrightarrow{AB}=\{2-0,1-1,3-(-1)\}=\{2,0,4\}$,

$$|\overrightarrow{AB}|=\sqrt{2^2+0^2+4^2}=2\sqrt{5},$$

$$\overrightarrow{AB}^0=\{2,0,4\}/\sqrt{20},\quad \cos\alpha=\sqrt{5}/5,$$

$$\cos\beta=0/\sqrt{20}=0,\quad \cos\gamma=2\sqrt{5}/5.$$

【使用坐标进行向量运算】设

$$\boldsymbol{a}=\{a_1,a_2,a_3\},\boldsymbol{b}=\{b_1,b_2,b_3\},\boldsymbol{c}=\{c_1,c_2,c_3\},$$

则有

$$m\boldsymbol{a}+n\boldsymbol{b}=\{ma_1+nb_1,ma_2+nb_2,ma_3+nb_3\},$$

$$\boldsymbol{a}\cdot\boldsymbol{b}=a_1b_1+a_2b_2+a_3b_3,$$

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix},$$

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}=(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}$$
$$=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

【使用向量运算计算几何量】

(1)**线段** AB **的长度为** $|\overrightarrow{AB}|$ ；

(2)**平行四边形** $ABCD$ **的面积是** $|\overrightarrow{AB}\times\overrightarrow{AD}|$；

(3)$\triangle ABC$ **的面积是** $|\overrightarrow{AB}\times\overrightarrow{AC}|/2$；

(4)**以三个向量** $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ **为棱的平行六面体的体积为三个向量的混合积的绝对值** $|(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}|$；

(5)**线段** AB **与** AC **的夹角为**

$$\arccos(\overrightarrow{AB}^0\cdot\overrightarrow{AC}^0).$$

4.2　空间解析几何

【平面的法向量】与平面垂直的非零向量.

【平面的方程】主要有以下几种.

〖**点法式方程**〗**如图**4-2-1 **所示,若平面**π **经过点** $P_0(x_0,y_0,z_0)$,**法向量为** $\boldsymbol{n}=\{A,B,C\}$,**则其方程为**

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0.$$

〖**三点式方程**〗**若平面** π **经过点** $P(x_1,y_1,z_1)$, $Q(x_2,y_2,z_2)$,$R(x_3,y_3,z_3)$,**则其方程为**

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0.$$

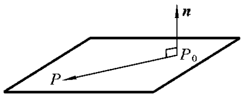


图 4-2-1

〖截距式方程〗如图4-2-2 所示,若平面π 在三个坐标轴上的截距依次为 $a,b,c,abc\neq 0$,则其截距式方程为

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1.$$

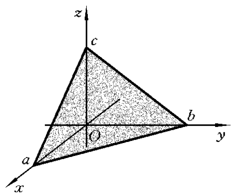


图 4-2-2

〖一般方程〗平面方程都可化作三元一次方程

$$Ax+By+Cz+D=0,$$

称之为一般方程. 其中变量的系数构成平面的法向量 $\boldsymbol{n}=\{A,B,C\}$.

〖点拨〗一般方程的形式简单,便于比较;截距式方程便于构图;寻找平面的方程时,较多的是使用点法式方程. 为了确定法向量 $\boldsymbol{n}$,常常寻找与 $\boldsymbol{n}$ 垂直

的两个向量，将这两个向量做叉乘运算来得到 $\boldsymbol{n}$.

【直线的方向向量】与直线平行的非零向量 $\boldsymbol{\tau}$.

【直线的方程】主要有以下几种.

〖点向式方程〗如图 4-2-3 所示，若直线 L 经过点 $P(x_0,y_0,z_0)$，方向向量为 $\boldsymbol{\tau}=\{l,m,n\}$，则其方程为

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}\quad(\text{也称对称式}).$$

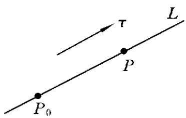


图 4-2-3

〖两点式方程〗若直线 L 经过两点 $P(x_1,y_1,z_1)$，$Q(x_2,y_2,z_2)$，则其两点式方程为

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1}.$$

〖参数方程〗若直线 L 经过点 $P(x_0,y_0,z_0)$，方向向量为 $\boldsymbol{\tau}=\{l,m,n\}$，则其参数式方程为

$$\begin{cases}x=x_0+lt,\\ y=y_0+mt,\quad(-\infty<t<+\infty).\\ z=z_0+nt\end{cases}$$

〖一般方程〗将直线看做两个不平行的平面的交线，联立这两个平面方程便得到直线的代数方程，称为一般方程：

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases}$$

〖点拨 1〗确定直线的方程时,关键是寻找方向向量 τ,如果不能直接得到其信息,通常应该考虑寻找与 τ 垂直的两个向量,将这两个向量做叉乘运算便得到 τ.

〖点拨 2〗使用较多的是点向式方程. 但是若知道直线所在的两个平面,则使用一般方程.

【平面束方程】如图 4-2-4 所示,通过一条直线的所有平面的集合叫平面束. 设这条直线的方程为

$$L:\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0, \end{cases}$$

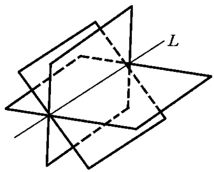


图 4-2-4

则含参数 λ 的三元一次方程

$$A_1x + B_1y + C_1z + D_1 + \lambda(A_2x + B_2y + C_2z + D_2) = 0$$

包含了通过直线 L 的除 $A_2x+B_2y+C_2z+D_2=0$ 之外的所有平面,称之为平面束方程.

【距离公式】设 M,L,π 分别表示点、直线、平面,M_0,

M_1, M_2 表示直线或平面上的点，则有如下距离公式.

〖两点距〗　$d(M_1M_2)=|\overrightarrow{M_1M_2}|$.

〖点线距〗如图 4-2-5 所示，

$$d(M,L)=|\overrightarrow{M_0M}\times\boldsymbol{\tau}^0|.$$

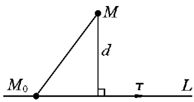


图 4-2-5

〖点面距〗如图 4-2-6 所示，

$$d(M,\pi)=|\overrightarrow{M_0M}\cdot\boldsymbol{n}^0|.$$

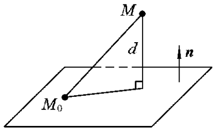


图 4-2-6

〖点拨 1〗线线距、线面距、面面距均可归于以上情形.

〖点拨 2〗原点 O 到平面 $Ax+By+Cz+D=0$ 的距离为

$$d(O,\pi)=|D|/\sqrt{A^2+B^2+C^2}.$$

〖异面直线之距离〗如图 4-2-7 所示,

$$d(L_1,L_2)=|\overrightarrow{M_1M_2}\cdot(\boldsymbol{\tau}_1\times\boldsymbol{\tau}_2)^0|.$$

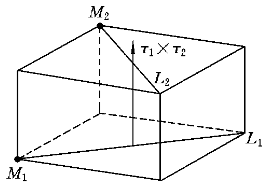


图 4-2-7

【夹角公式】直线、平面之间的夹角可以用它们的方向向量和法向量的夹角来表示.

〖线线夹角〗两直线 L_1, L_2 的夹角,记作 $\langle L_1,L_2\rangle$,定义为两个方向向量 τ_1,τ_2 的夹角 $\langle\tau_1,\tau_2\rangle$(起点重合后的小角度值).

〖面面夹角〗两平面 π_1,π_2 的夹角,记作 $\langle\pi_1,\pi_2\rangle$,定义为两个法向量 $\boldsymbol{n}_1,\boldsymbol{n}_2$ 的夹角 $\langle\boldsymbol{n}_1,\boldsymbol{n}_2\rangle$.

〖线面夹角〗如图 4-2-8 所示,直线 L 和平面 π 的夹角,记作 $\langle L,\pi\rangle$,定义为直线 L 与它在平面 π 上的投影 L' 之间的夹角:$\langle L,\pi\rangle=\left|\dfrac{\pi}{2}-\langle\boldsymbol{n},\tau\rangle\right|$.

【判定两条直线是否共面】两条直线共面的充要条件为方向向量 τ_1,τ_2 和 $\overrightarrow{P_1P_2}$ 为共面向量. 在共面时,若方向向量 τ_1,τ_2 不平行,则两直线相交.

【曲面的一般方程】形为 $F(x,y,z)=0$.

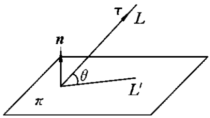


图 4-2-8

【柱面】动直线沿着定曲线平移形成的曲面. 其中动直线称为母线,定曲线称为准线.

【标准柱面的方程】母线平行于 z 轴的柱面方程为 $F(x,y)=0$,母线平行于 y 轴的柱面方程为 $F(x,z)=0$,母线平行于 x 轴的柱面方程为 $F(z,y)=0$.

【旋转面】平面曲线 C 绕同一平面内的直线 L 旋转一周所成的曲面称为旋转面. 称 L 为旋转轴.

【标准旋转面方程】平面曲线 $\begin{cases} f(y,z)=0, \\ x=0 \end{cases}$ 或 $\begin{cases} f(x,z)=0, \\ y=0 \end{cases}$ 绕 z 轴旋转产生的旋转面方程为

$$f(\pm\sqrt{x^2+y^2},z)=0.$$

例　曲面 $z=x^2+y^2$ 可以看作是由曲线 $\begin{cases} z=x^2, \\ x=0 \end{cases}$ 绕 z 轴旋转产生的旋转面.

【空间曲线的方程】

〖一般方程〗看作两个曲面的交线. 方程为

$$\begin{cases} F(x,y,z)=0, \\ G(x,y,z)=0. \end{cases}$$

〖参变量方程〗 $\begin{cases} x=x(t), \\ y=y(t), \\ z=z(t) \end{cases}$ $(\alpha\leqslant t\leqslant\beta)$.

【常用空间曲线】

〖**直线**〗$\dfrac{x-x_0}{l}=\dfrac{y-y_0}{m}=\dfrac{z-z_0}{n}$.

〖**平面曲线**〗落在某个平面上的曲线.

〖**柱面曲线**〗落在某个柱面上的曲线.

〖**球面曲线**〗落在某个球面上的曲线.

例　如图 4-2-9 所示,球面 $x^2+y^2+z^2=R^2(z>0)$和柱面$\left(x-\dfrac{R}{2}\right)^2+y^2=\left(\dfrac{R}{2}\right)^2$的交线,既是柱面曲线,又是球面曲线,称之为维维安妮曲线.

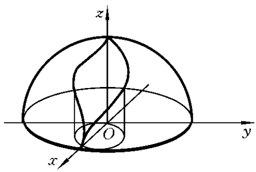


图 4-2-9

【空间曲线的投影柱面·投影曲线】从空间曲线方程$\begin{cases} F(x,y,z)=0, \\ G(x,y,z)=0 \end{cases}$中消去变量 z,得出的方程 $H(x,y)=0$便是空间曲线在Oxy面上的投影柱面方程.投影柱面与Oxy面的交线便是空间曲线在Oxy面上的投影曲线,其方程为$\begin{cases} H(x,y)=0, \\ z=0. \end{cases}$

〖点拨〗空间曲线$\begin{cases}x=x(t),\\y=y(t),(\alpha\leqslant t\leqslant\beta),\\z=z(t)\end{cases}$在 Oxy 面上的投影柱面方程为$\begin{cases}x=x(t),\\y=y(t)\end{cases}(\alpha\leqslant t\leqslant\beta)$，投影曲线方程为$\begin{cases}x=x(t),\\y=y(t),(\alpha\leqslant t\leqslant\beta).\\z=0\end{cases}$

常见题型·应对

单独考察这一部分的试题较少，确定直线、平面或旋转曲面的方程是多元微积分学的基础，在积分区域绘制，极值问题中会用到.

【确定平面的方程】

例　求满足以下条件的平面方程：

(1)平行于平面$3x-7y+5z-12=0$，过点$(4,1,-2)$.

(2)过两点 $A(8,-3,1)$，$B(4,7,2)$且垂直于平面 $3x+5y-7z-21=0$.

(3)过三点 $A(7,6,7)$，$B(5,10,5)$及 $C(-1,8,9)$.

(4)与坐标轴的截距都相等且过点 $P(6,2,-4)$.

解　(1)由条件知$\boldsymbol{n}=\{3,-7,5\}$，故由点法式方程得所求平面为

$$3(x-4)-7(y-1)+5(z+2)=0.$$

(2) 所求平面的法向量$\boldsymbol{n}$与

$\overrightarrow{AB}=\{4-8,7-(-3),2-1\}=\{-4,10,1\}$
以及$\{3,5,-7\}$垂直. 因为

$$\{-4,10,1\}\times\{3,5,-7\}$$
$$=\{-75,-25,-50\}\ /\!/\ \{3,1,2\},$$

故可取$\boldsymbol{n}=\{3,1,2\}$,结合点A,由点法式方程得所求平面为$3x+y+2z-23=0$.

(3) 法向量 $\boldsymbol{n}$ 与$\overrightarrow{AB}=\{-2,4,-2\}$及$\overrightarrow{AC}=\{-8,2,2\}$垂直,且

$$\overrightarrow{AB}\times\overrightarrow{AC}=4\{3,5,7\}\ /\!/\ \{3,5,7\},$$

故可取$\boldsymbol{n}=\{3,5,7\}$,结合点A得到所求方程为

$$3x+5y+7z+100=0.$$

(4) 套用截距式方程,因$a=b=c$,故有$\frac{1}{a}(x+y+z)=1$,将点P的坐标代入,得$a=4$,故所求方程为

$$\frac{x}{4}+\frac{y}{4}+\frac{z}{4}=1.$$

【确定直线的方程】点向式方程是基本思路.

例　求满足下列条件的直线的方程:

(1)过点$(2,3,4)$且垂直于平面$3x-5y+7z+6=0$;

(2)过点$(2,-3,8)$且平行于z轴;

(3)过原点和点(a,b,c);

(4)过点$A(1,2,3)$且与z轴相交,垂直于直线$x=y=z$.

解　(1)所求直线的方向向量便是题给平面的法向量$\{3,-5,7\}$,故由点向式方程得所求方程为

$$\frac{x-2}{3}=\frac{y-3}{-5}=\frac{z-4}{7}.$$

(2)依题意(与z轴平行),可取方向向量为$\boldsymbol{k}=\{0,0,1\}$,故由点向式方程得所求方程为(分母可以写作0)

$$\frac{x-2}{0}=\frac{y+3}{0}=\frac{z-8}{1}.$$

(3)方向向量平行于所给两点的连线,取原点及使用方向向量$\tau=\{a,b,c\}$,得所求直线方程为

$$\frac{x}{a}=\frac{y}{b}=\frac{z}{c}.$$

(4)设所求直线与z轴交点为$B(0,0,c)$,则$\overrightarrow{AB}$与直线$x=y=z$的方向向量垂直,即

$$\{1,2,3-c\}\perp\{1,1,1\},\quad 1+2+3-c=0,$$

故$c=6$.从而$\tau=\{1,2,-3\}$,所求方程为

$$\frac{x-1}{1}=\frac{y-2}{2}=\frac{z-3}{-3}.$$

【常用简单曲面】必须掌握图形的曲面有如下几种.

〖**球面**〗　$x^2+y^2+z^2=R^2$.

〖**圆柱面**〗　$x^2+y^2=R^2$.

〖**圆锥面**〗　$z=\sqrt{x^2+y^2}$.

〖**旋转抛物面**〗　$z=x^2+y^2$.

【二次曲面】只需了解的二次曲面有如下几种.

〖**椭球面**〗　$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$.

〖**单叶双曲面**〗　$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$.

〖**双叶双曲面**〗 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=-1.$

〖**椭圆抛物面**〗 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=z.$

〖**双曲抛物面**〗 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=z.$

〖**二次锥面**〗 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=0.$

第5章　多元函数微分学

5.1　多元函数偏导数与全微分

大纲要求

理解多元函数的概念,理解二元函数的几何意义.了解二元函数的极限与连续的概念以及有界闭区域上连续函数的性质.理解多元函数偏导数和全微分的概念,会求全微分,了解全微分存在的必要条件和充分条件,了解全微分形式的不变性.理解方向导数与梯度的概念并掌握其计算方法.掌握多元复合函数一阶、二阶偏导数的求法.了解隐函数存在定理,会求多元隐函数的偏导数.

注　数学二不要求方向导数、梯度.

知识点·点拨

【平面区域基本概念】

〖邻域〗Oxy 面上的以点 P_0 为圆心、以 δ 为半径的不含圆周的圆盘称为点 P_0 的一个邻域,记作 $N(P_0,\delta)$.

〖内点〗设 D 是平面点集.若点 P_0 有一个邻域包含在 D 中,则称点 P_0 是该点的内点.

〖边界点〗设 D 是平面点集.若点 P_0 的任何邻域中都包含 D 内和 D 外的点,则称点 P_0 是 D 的边界点.

〖开区域〗若 D 中每一个点都是 D 的内点,且任何两点都可以用全在 D 中的折线连接(连通性),则

称 D 为开区域.

〖闭区域〗开区域连同其边界点全体构成的集合称为闭区域.它一定是连通的.

〖有界点集〗能够被某个圆所包含的平面点集称为有界点集.

【常见二维区域】由一条或多条平面曲线围成.

【常见三维区域】由一片或多片曲面围成.例如具备以下特征的点集.

(1)开球体:$x^2+y^2+z^2<1$.

(2)闭球体:$x^2+y^2+z^2\leqslant 1$.

(3)圆锥体:由平面 $z=1$、锥面 $z=\sqrt{x^2+y^2}$ 围成.

【二元函数】二元函数是由平面区域 D 到实数集的一个对应规则,记作 $z=f(x,y),(x,y)\in D$.

【二元函数的几何表示】

〖空间曲面〗$z=f(x,y),(x,y)\in D$ 的图形指以下点集:

$$S=\{(x,y,z)\mid z=f(x,y),(x,y)\in D\},$$

如图 5-1-1 所示.

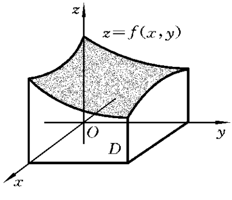


图 5-1-1

〖等量线〗二元函数 $f(x,y)$的定义域 D 中函数值等于 c 的点形成一条曲线，称为等量线

$$L_C = \{(x,y) \mid f(x,y) = c\}.$$

等量线的分布表现了函数 $z=f(x,y)$的变化特征. 其中，c 在函数的值域中变动. 如图 5-1-2 所示.

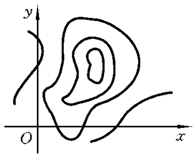


图 5-1-2

【二元函数在一点的极限】若任给$\varepsilon>0$，存在$\delta>0$，使得不等式

$$|f(P) - l| < \varepsilon$$

在 $0<|P-P_0|<\delta$ 时成立，则称 l 是 $f(P)$在点 P_0 的极限，记作

$$l = \lim_{P\to P_0} f(P) = \lim_{\substack{x\to x_0\\ y\to y_0}} f(x,y).$$

【路径极限】动点 P 沿着一条给定的曲线 C 趋于 P_0，所得的极限称做路径极限，记作 $\lim\limits_{\substack{P\to P_0\\ P\in C}} f(P)$.

【路径极限与极限的关系】$A=\lim\limits_{P\to P_0} f(P) \Leftrightarrow$沿任何过点 P 的曲线 C，路径极限都为 A.

〖点拨 1〗此定理常常用来说明一个极限不存在.

〖点拨2〗沿任何直线方向的路径极限存在且相等时,函数的极限还是可能不存在.

例1　说明极限$\lim\limits_{\substack{x\to 0\\ y\to 0}}\frac{xy}{x^2+y^2}$不存在.

证　沿着曲线$x=y$,所论极限为

$$\lim_{x\to 0}\frac{x^2}{x^2+x^2}=\frac{1}{2},$$

而沿着曲线$y=-x$,所论极限为

$$\lim_{x\to 0}\frac{-x^2}{x^2+x^2}=-\frac{1}{2}.$$

沿不同的路径得出的路径极限不相等说明该极限不存在.

例2　函数$f(x,y)=\begin{cases}1, & y=x^2,\\ 0, & y\neq x^2\end{cases}$在原点沿任何直线的路径极限是0,但是沿着曲线$y=x^2$的路径极限为1,故该函数在原点的极限不存在.

【二元函数在一点连续】称函数$f(P)$在点P_0处连续,若$f(P)$在点P_0的某个邻域有定义,并且

$$\lim_{P\to P_0}f(P)=f(P_0).$$

【二元初等函数的连续性】由自变量x,y的基本初等函数经有限次复合运算、四则运算构成的二元初等函数在其定义域上连续.

【连续函数的运算性质】四则运算、复合运算均保持函数的连续性.

【最值定理】设D是有界闭区域,则其上的连续函数具有最大值与最小值,从而是有界函数.

【介值定理】设D是有界闭区域,f是D上的连续函

数. 若实数 c 在函数值 $f(A)$ 和 $f(B)$ 之间,则必有点 $P \in D$,使得 $f(P)=c$.

【偏导数】二元函数 $z=f(x,y)$ 在点 (x_0,y_0) 关于 x 的偏导数定义为 x 的一元函数 $z=f(x,y_0)$ 在点 $x=x_0$ 关于 x 的导数:

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x},$$

记作 $z_x(x_0,y_0)$,或 $\dfrac{\partial z}{\partial x}(x_0,y_0)$,或 $f_x(x_0,y_0)$. 类似地,可以定义和标记关于 y 的偏导数.

【偏导数存在与函数的连续性关系】与一元函数不同的是,多元函数 $f(P)$ 在点 P 处连续与在点 P 处偏导数存在之间没有逻辑关系.

例 1　函数 $z=\sqrt{x^2+y^2}$ 在原点连续但是偏导不存在.

例 2　函数 $z=\begin{cases}\dfrac{xy}{x^2+y^2}, & x^2+y^2\neq 0,\\ 0, & x^2+y^2=0\end{cases}$ 在原点偏导数存在但是不连续.

【全增量】$z=f(x,y)$ 在点 (x_0,y_0) 处与附近的点的函数值的差,称为全增量,记作

$$\Delta z=f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0).$$

【全微分】若 $z=f(x,y)$ 在点 (x_0,y_0) 处的全增量 Δz 能写为以下形式:

$$\Delta z=A\Delta x+B\Delta y+o\left(\sqrt{\Delta x^2+\Delta y^2}\right),$$

则称 $z=f(x,y)$ 在点 (x_0,y_0) 处可微,此时,若记 $\Delta x=\mathrm{d}x,\Delta y=\mathrm{d}y$,称

$$A\mathrm{d}x + B\mathrm{d}y$$

为 $z=f(x,y)$ 在点 (x_0,y_0) 处的全微分，记作 $\mathrm{d}f(x_0,y_0)$.

【全微分计算公式】当 $f(x,y)$ 可微时，便有

$$\mathrm{d}f(x,y) = f_x(x,y)\mathrm{d}x + f_y(x,y)\mathrm{d}y.$$

【方向导数】以下极限称为函数 $f(P)$ 在点 P_0 处沿方向向量 $\boldsymbol{l}$ 的方向导数(其中 PP_0 平行于向量 $\boldsymbol{l}$)：

$$\frac{\partial f}{\partial l}(P_0) = \lim_{P\to P_0}\frac{f(P)-f(P_0)}{|P_0P|}.$$

【方向导数计算公式】当二元函数 $f(P)$ 于 P_0 可微时，可推出方向导数计算公式：

$$\frac{\partial f}{\partial l}(P_0) = \{f_x(P_0), f_y(P_0)\}\cdot \boldsymbol{l}^0.$$

【梯度】由函数 $f(x,y)$ 在点 P 的偏导数构成的向量 $\{f_x, f_y\}$ 称为函数 f 在该点的梯度，记作 $\mathbf{grad}f(P)$. 函数在该点的所有方向导数中最大的方向导数即是此梯度向量的模.

〖点拨〗三元函数 $f(x,y,z)$ 在点 P 的梯度 $\mathbf{grad}f(P)$ 为矢量 $\{f_x, f_y, f_z\}$.

例　求向径 $\boldsymbol{r}=\{x,y,z\}$ 的模的梯度.

解　向径的模为 $r=\sqrt{x^2+y^2+z^2}$,

$$\mathbf{grad}r = \{r_x, r_y, r_z\} = \left\{\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right\} = \boldsymbol{r}^\circ.$$

【可微的必要条件】由函数 $f(P)$ 在点 P 处可微推出以下结果(推理均不可逆)：

(1)该函数在点 P 处连续；

(2)该函数在点 P 处的偏导都存在；

(3)该函数在点P处的任何方向的方向导数都存在.

【可微的充要条件】函数$z=f(x,y)$在点(x_0,y_0)处可微等价于

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\Delta z-f_x(x_0,y_0)\Delta x+f_y(x_0,y_0)\Delta y}{\sqrt{\Delta x^2+\Delta y^2}}=0.$$

其中　$$\Delta z=f(x,y)-f(x_0,y_0).$$

【可微的充分条件】若函数$f(x,y)$的偏导函数$f_x(x,y)$,$f_y(x,y)$在点(x_0,y_0)处是连续的,则函数$f(x,y)$在点(x_0,y_0)处可微.

【多元复合函数的微分法】对复合函数$z=f(u(x,y),v(x,y))$,有链导法则:

$$z_x=f_u(u,v)\cdot u_x+f_v(u,v)\cdot v_x,$$

$$z_y=f_u(u,v)\cdot u_y+f_v(u,v)\cdot u_y.$$

例　设$z=f(x-y,x^2y)$,求z_x,z_y.

解　$z_x=f_1+2xyf_2$,　$z_y=f_1+x^2f_2$.

注　数字下标用来表示关于自变量的偏导数:

f_1表示f_u,　f_2表示f_v,　f_{12}表示f_{uv}.

【二阶偏导数】对一阶偏导函数z_x,z_y再求偏导,便成为二阶偏导数,简记为z_{xx},z_{xy},z_{yy},z_{yx}.

〖点拨〗当z_{xy},z_{yx}连续时,它们是相等的.

例　设$z=x^2e^y$,求z的所有二阶偏导数.

解　$z_x=2xe^y$,$z_{xx}=2e^y$,$z_{xy}=2xe^y$,

$z_y=x^2e^y$,$z_{yy}=x^2e^y$,$z_{yx}=2xe^y$.

【复合函数的二阶偏导数】对抽象函数的复合函数,例如$z=f(x-y,x^2y)$,计算二阶偏导数的关键是要

将一阶偏导函数看作是与函数 $z=f(x-y,x^2y)$ 具有相同复合结构的函数继续求偏导.

常见题型 · 应对

【多元复合函数的一阶或二阶偏导数计算】 此类试题为基本题,按照链导法则计算便是.

例 设 $z=f(xy,x^2-y^2)$,求 z_{xx}.

解 $z_x=f_1\cdot y+f_2\cdot 2x$,

$$z_{xx}=y(f_{11}\cdot y+f_{12}\cdot 2x)+2f_2+2x(f_{21}\cdot y+f_{22}\cdot 2x),$$

其中,导函数 $f_1(xy,x^2-y^2)$ 对 x 再求偏导时遵循与函数 $f(xy,x^2-y^2)$ 对 x 求偏导时相同的链导法则,即

$$(f_1)_x=f_{11}\cdot(xy)'_x+f_{12}\cdot(x^2-y^2)'_x=yf_{11}+2xf_{12}.$$

【一个方程确定的隐函数的微分法】 设 $z=z(x,y)$ 由方程 $F(x,y,z)=0$ 所确定. 两边对 x 求导数,得

$$F_1(x,y,z)+F_3(x,y,z)\frac{\partial z}{\partial x}=0,$$

故有
$$\frac{\partial z}{\partial x}=-F_1/F_3.$$

类似地,推得 $\dfrac{\partial z}{\partial y}=-F_2/F_3$.

例 设方程 $x+y+z=\mathrm{e}^z$ 确定了二元函数 $z=z(x,y)$,求 z_x.

解1 视 y 为常数,z 为 x 的函数,于方程两边同时对 x 求导,得 $1+z_x=\mathrm{e}^z\cdot z_x$,故有

$$z_x=1/(\mathrm{e}^z-1).$$

解2　使用公式$\frac{\partial z}{\partial x}=-F_1/F_3$,记函数

$$F(x,y,z)=x+y+z-\mathrm{e}^z,$$

由于$F_x=1,F_z=1-\mathrm{e}^z$,故有$z_x=1/(\mathrm{e}^z-1)$.

【多个方程所确定的函数的偏导数计算】由几个方程联立的函数方程组也可确定一个或多个函数关系,其中自变量个数通常是变量个数与方程个数之差,自变量确定之后剩余的变量便是因变量. 这类隐函数求导方法仍然与上面的例子一致:分清自变量与因变量,等式两边同时求导便可.

例1　设方程组$\begin{cases}x^2+y^2+z^2=R^2,\\ x+y+z=1\end{cases}$确定了两个一元函数$y=y(x),z=z(x)$,求$y'(x),z'(x)$.

解　以x为自变量,y,z为x的一元隐函数. 两边对x求导,得

$$\begin{cases}2x+2yy'+2zz'=0,\\ 1+y'+z'=0.\end{cases}$$

解得$y'=(z-x)/(z-y),z'=(x-y)/(z-y)$.

例2　设$y=f(x,t),F(x,y,t)=0$,求$\frac{\mathrm{d}y}{\mathrm{d}x}$.

解　视为方程组,两边对x求导(t,y均看作x的一元函数),得

$$y'=f_1+f_2t',F_1+F_2y'+F_3t'=0,$$

解出
$$\frac{\mathrm{d}y}{\mathrm{d}x}=y'=\frac{f_1F_2-F_1F_3}{F_3+f_2F_2}.$$

【方向导数计算·最大方向导数计算】通常是在可微的条件下计算沿指定方向的方向导数. 依据以下公式即可:

$$\frac{\partial f}{\partial l}(P_0) = \{f_x(P_0), f_y(P_0)\} \cdot l^0.$$

例1　**求**$u_x = xe^{2y}$**在点**$P(1,1)$**沿**$\boldsymbol{l}=\{1,-1\}$**的方向导数**$\frac{\partial u}{\partial l}$.

解　$u_x = e^{2y}$，$u_y = 2xe^{2y}$，$\boldsymbol{l}^0 = \{1,-1\}/\sqrt{2}$，

故
$$\left.\frac{\partial u}{\partial l}\right|_P = \{e, 2e\} \cdot \frac{\{1,-1\}}{\sqrt{2}} = -\frac{e^2}{\sqrt{2}}.$$

例2　**求** $u = x^2 - xy + y^2$ **在点** $P(1,1)$**的最大方向导数**.

解　由梯度的涵义知,本题所求即梯度的模:

$$|\mathbf{grad}u(P)| = |\{2x-y, 2y-x\}|_P = \sqrt{2}.$$

【数量函数的梯度计算】套用基本公式:

$$\mathbf{grad}u(P) = \{u_x, u_y, u_z\}_P.$$

【向量函数的散度和旋度计算】设有向量函数 $\boldsymbol{F} = \{P(x,y,z), Q(x,y,z), R(x,y,z)\}$,则其散度 $\mathrm{div}P$ 是一个数量,旋度 $\mathbf{rot}\boldsymbol{F}$ 是一个向量:

散度　$\mathrm{div}F = P_x + Q_y + R_z$,

旋度
$$\mathbf{rot}\boldsymbol{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{vmatrix}.$$

【分段函数的分析性质讨论】这类试题要求讨论多元函数在一点的连续性、偏导存在性和可微性.直接依据定义写出要讨论的问题的极限关系式后进行讨论便是.注意利用上述三种性质的相互关系.

5.2　多元微分学的应用

大纲要求

了解空间曲线的切线和法平面及曲面的切平面和法线的概念，会求它们的方程. 了解二元函数的二阶泰勒公式. 理解多元函数极值和条件极值的概念，掌握多元函数极值存在的必要条件，了解二元函数极值存在的充分条件，会求二元函数的极值，会用拉格朗日乘数法求条件极值，会求简单多元函数的最大值和最小值，并会解决一些简单的应用问题.

注　数学二仅要求极值内容.

知识点 · 点拨

【曲线的切线和法平面】设曲线 C 的参数方程为 $\begin{cases} x=x(t), \\ y=y(t), \\ z=z(t), \end{cases}$ 则曲线 C 在点 $P(x(t),y(t),z(t))\in C$ 的切向量为 $\{x'(t),y'(t),z'(t)\}$，切线方程为

$$\frac{x-x(t)}{x'(t)}=\frac{y-y(t)}{y'(t)}=\frac{z-z(t)}{z'(t)},$$

法平面方程为

$$x'(t)(x-x(t))+y'(t)(y-y(t))+z'(t)(z-z(t))=0.$$

【曲面的切平面和法线】曲面 $F(x,y,z)=0$ 在点 P_0 的切平面的法向量为 $\mathbf{grad}F(P_0)$，切平面的方程为

$$F_x(P_0)(x-x_0)+F_y(P_0)(y-y_0)+F_z(P_0)(z-z_0)=0,$$

与切平面垂直，过点 P_0 的直线称做该曲面的法线，

其方程为

$$\frac{x-x_0}{F_x(P_0)}=\frac{y-y_0}{F_y(P_0)}=\frac{z-z_0}{F_z(P_0)}.$$

【极值】设函数 $z=f(x,y)$ 在区域 D 有定义. 若在点 P_0 的某个邻域中,函数值 $f(P_0)$ 是最大(小)值,则称 $f(P_0)$ 是函数在区域 D 上的一个极大(小)值. 极值也称为局部最值.

【条件极值】条件极值是指某个特定范围的极值:

(1)二元函数 $z=f(x,y)$ 在平面区域 D 内某条曲线 $h(x,y)=0$ 上的极值;

(2)三元函数 $u=f(x,y,z)$ 在空间区域 V 内某条曲线 $\begin{cases}h(x,y,z)=0,\\ g(x,y,z)=0\end{cases}$ 上的极值;

(3)三元函数 $u=f(x,y,z)$ 在空间区域 V 内某块曲面 $F(x,y,z)=0$ 上的极值.

【驻点】函数 f 的驻点是使得 $\mathbf{grad} f(P)=0$ 的点 P.

【极值点的判定】

〖必要条件〗若可微函数在定义域的内点取得极值,则该点是函数的驻点.

〖充分条件〗设点 P_0 是函数 $z=f(x,y)$ 在定义域内的驻点,二阶偏导数在点 P_0 的某个邻域上连续. 记

$$A=f_{xx}(P_0),\quad B=f_{xy}(P_0),C=f_{yy}(P_0),$$
$$\Delta=AC-B^2,$$

则　(1)$\Delta<0$ 时,P_0 不是函数的极值点;

(2)$\Delta>0$ 且 $A>0$ 或 $C>0$ 时,P_0 是函数的极小值点;

(3)$\Delta>0$ 且 $A<0$ 或 $C<0$ 时,P_0 是函数的极大值点.

〖点拨〗可以使用二阶矩阵的正定、负定、不定判别法来记忆以上条件.

【条件极值问题的驻点】

(1)对于条件为 $h(x,y)=0$,目标函数为 $z=f(x,y)$ 的条件极值问题,可以按照如下方法寻找可能的极值点(称为驻点).

〖拉格朗日乘数法〗构造拉格朗日函数

$$L(x,y,\lambda)=f(x,y)+\lambda h(x,y),$$

求方程组 $\mathbf{grad}L=\{L_x,L_y,L_\lambda\}=\mathbf{0}$ 的解即可.

〖消元法〗从曲线方程 $h(x,y)=0$ 中解出 $y=y(x)$,代入目标函数 $z=f(x,y)$ 中,使之成为一元函数,然后求其驻点.

(2)对于条件为 $h(x,y,z)=0$,目标函数为 $u=f(x,y,z)$ 的条件极值问题,可以按照如下方法寻找可能的极值点.

〖拉格朗日乘数法〗构造拉格朗日函数

$$L(x,y,z,\lambda)=f(x,y,z)+\lambda h(x,y,z),$$

则方程组 $\mathbf{grad}L=\{L_x,L_y,L_z,l_\lambda\}=\mathbf{0}$ 的解即是驻点.

〖消元法〗从曲线方程 $h(x,y,z)=0$ 中解出一个变量代入到目标函数,化作二元函数后再求其驻点.

(3)对于条件为 $\begin{cases}h(x,y,z)=0,\\ g(x,y,z)=0,\end{cases}$ 目标函数为 $u=f(x,y,z)$ 的条件极值问题,可以按照如下方法寻找可能的极值点.

〖拉格朗日乘数法〗构造拉格朗日函数

$$L(x,y,z,\lambda,\mu)=$$
$$f(x,y,z)+\lambda h(x,y,z)+\mu g(x,y,z),$$

则方程组 $\mathbf{grad}L=\{L_x,L_y,L_z,L_\lambda,L_\mu\}=\mathbf{0}$ 的解即是可能的极值点.

〖消元法〗从条件方程 $\begin{cases}h(x,y,z)=0,\\ g(x,y,z)=0\end{cases}$ 中解出两个变量,代入到目标函数中,化作一元函数后再求其驻点.

〖点拨〗以上求得的点是按照条件极值点的必要条件计算的,是否为极值点,还需用其他方法确认.

如果要计算的是函数的最值,且已知最值存在于驻点之中,则直接对驻点的函数值进行比较即得.

常见题型·应对

【计算空间曲线的切线及法平面】

例 在曲线 $x=t,y=t^2,z=t^3$ 上求一点,使曲线在此点的切线平行于平面 $x+2y+z=4$.

解 切线的方向向量为 $\tau=\{1,2t,3t^2\}$,它应与所给平面的法向量 $\boldsymbol{n}=\{1,2,1\}$ 垂直,即

$$\tau\cdot\boldsymbol{n}=1+4t+3t^2=0,$$

解出 $t=-1,-\dfrac{1}{3}$,故得到两个所求切点

$$(-1,1,-1)\quad 与\quad \left(-\frac{1}{3},\frac{1}{9},-\frac{1}{27}\right).$$

【计算空间曲面的切平面及法线】

例 求曲面 $z=y+\ln(x/z)$ 在点 $P_0(1,1,1)$ 的切

平面方程.

解　记　　$F(x,y,z)=y+\ln(x/z)-z$,

则　　　$\mathbf{grad}F(P_0)=\{1,1,-2\}$,

故切平面方程为

$$(x-1)+(y-1)-2(z-1)=0.$$

【求解条件极值问题】首先明确问题中的两个函数:目标函数和条件约束函数,再通过拉格朗日乘数法求可能的极值点.

例　求圆周 $x^2+y^2=1$ 上的一个点使 x^2y 最大.

分析　这属于条件极值问题,目标函数 $z=x^2y$,约束条件是 $x^2+y^2-1=0$,以下用两种方法分别求之.

解1(代入消元法)　注意到$x^2=1-y^2$,代入z 中后化作一元函数

$$z=(1-y^2)y,\ -1\leqslant y\leqslant 1,$$

的最大值问题.由 $z_y=1-3y^2=0$ 得驻点 $y=\pm\frac{\sqrt{3}}{3}$,经比较知,当 $y=\frac{\sqrt{3}}{3}$(此时 $x=\pm\frac{\sqrt{6}}{3}$)时目标函数值最大,故所求点为

$$\left(\frac{\sqrt{6}}{3},\frac{\sqrt{3}}{3}\right)\text{及}\left(-\frac{\sqrt{6}}{3},\frac{\sqrt{3}}{3}\right).$$

解2(拉格朗日乘数法)　引入拉格朗日函数

$$L=x^2y+\lambda(x^2+y^2-1),$$

从方程组$\begin{cases}L_x=2xy+2x\lambda=0,\\L_y=x^2+2y\lambda=0,\\L_\lambda=x^2+y^2-1=0\end{cases}$ 可求出 $1-3y^2=0$,得

出与解法1相同的结果.

【条件极值快捷计算法】有时可以套用以下的和积规则直接得到所求的最大值或者最小值.

〖定积条件〗若 n 个正数的积为定数,则当这 n 个正数彼此相等时,其和最小.

〖定和条件〗若 n 个正数的和为定数,则当这 n 个正数彼此相等时,其积最大.

例1　设计一个无盖长方体水池,容积为定值 V,要求用料(即表面积)最省,问:长、宽、高之比应为多少?

解　设长、宽、高依次是 x,y,z. 目标函数是和函数

$$S = xy + 2xz + 2yz$$

注意到上式中被加数 $xy,2xz,2yz$ 的乘积

$$xy \cdot 2xz \cdot 2yz = 4(xyz)^2 = 4V^2$$

是定数,故当这些因数彼此相等时,其和最小. 由此得

$$xy = 2xz = 2yz,$$

故 $x:y:z=2:2:1$ 时用料最省.

例2　设某企业的投入为三类:人力 x,固定资产 y 和技术 z,产出 u 与投入的关系为 $u=xy^2z^3$. 若三类投入的总和为600(万元),问:如何分配三者的比例,可使产出 u 取得最大?

解　定和条件 $x+y+z=600$ 可写为

$$x + \frac{y}{2} + \frac{y}{2} + \frac{z}{3} + \frac{z}{3} + \frac{z}{3} = 600,$$

以便使得被加数的乘积与目标函数 $u=xy^2z^3$ 成正

比，于是，当被加数相等，即当 $x=y/2=z/3=100$ 时，产出 u 取得最大.

【计算最大值或最小值问题】以二元函数为例. 设函数 $z=f(x,y)$ 在平面区域 D 上有定义.

〖步骤〗首先根据实际问题的意义明确所求的最值是否存在，然后求函数在区域内部的驻点和边界曲线上的驻点（依条件极值驻点求法），再比较这些点的函数值的大小即得所求的最值.

第6章　多元函数积分学

6.1　重　积　分

大纲要求

理解二重积分、三重积分的概念，了解重积分的性质，了解二重积分的中值定理. 掌握二重积分的计算方法（直角坐标、极坐标），会计算三重积分（直角坐标、柱面坐标、球面坐标）.

注　数学二只要求二重积分.

知识点·点拨

【二重积分定义】设 $f(x,y)$ 是有界闭区域 D 上的分片连续函数，则 $f(x,y)$ 在 D 上的二重积分定义为以下极限

$$\iint\limits_D f(x,y)\mathrm{d}\sigma = \lim_{\lambda\to 0}\sum_{i=1}^{n} f(x_i,y_i)\Delta\sigma_i,$$

其中，和式 $\sum\limits_{i=1}^{n} f(x_i,y_i)\Delta\sigma_i$ 是将 D 任意地分割成 n 个小片 $\Delta S_1,\Delta S_2,\cdots,\Delta S_n$，且从每个小片中任选一点 (x_i,y_i)，再将 $f(x_i,y_i)$ 与 ΔS_i 的面积 $\Delta\sigma_i$ 相乘后关于 i 求和所得，极限过程 $\lambda\to 0$ 是指分割方式无限加细.

【三重积分定义】设 $f(x,y,z)$ 是有界闭区域 V 上的分片连续函数，则 $f(x,y,z)$ 在 V 上的三重积分定义为以下极限（记号的意义类似于二重积分）

$$\iiint\limits_V f(x,y,z)\mathrm{d}V=\lim_{\lambda\to 0}\sum_{i=1}^{n}f(x_i,y_i,z_i)\Delta V_i.$$

【重积分的性质】以二重积分为例.

〖**分项积分法**〗**对任何常数** α,β，**有**

$$\iint\limits_D(\alpha f(x,y)+\beta g(x,y))\mathrm{d}\sigma$$

$$=\alpha\iint\limits_D f(x,y)\mathrm{d}\sigma+\beta\iint\limits_D g(x,y)\mathrm{d}\sigma.$$

〖**分片积分法**〗**若** D **被分划为子区域** D_1 **及** D_2，**则**

$$\iint\limits_D f(x,y)\mathrm{d}\sigma=\iint\limits_{D_1}f(x,y)\mathrm{d}\sigma+\iint\limits_{D_2}f(x,y)\mathrm{d}\sigma.$$

〖**比较性质**〗**若在** D **上成立** $f(x,y)\leqslant g(x,y)$，**则**

$$\iint\limits_D f(x,y)\mathrm{d}\sigma\leqslant\iint\limits_D g(x,y)\mathrm{d}\sigma.$$

〖**积分中值定理**〗**若** $f(x,y)$ **在** D **上连续，则**

$$\iint\limits_D f(x,y)\mathrm{d}\sigma=f(x_0,y_0)\,|D|$$

$(x_0,y_0)\in D$（$|D|$**表示** D **的面积**）.

【常用二维区域】当积分区域 D **是以下三种类型时，可以将二重积分化作逐次积分来计算.**

直角坐标系：x-**型区域**，y-**型区域**；

极坐标系：θ-**型区域**.

【x-型区域】若与 x **轴垂直的直线与** D **相交时，交线是连续的线段或一个点，则称** D **是** x-**型区域**.

〖**点拨**〗**可以用不等式表示图** 6-1-1 **中的** x-**型区

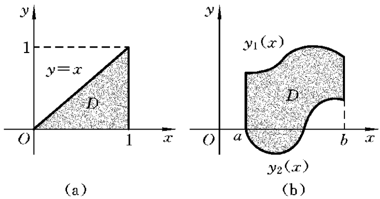


图 6-1-1

域：

(a) $\begin{cases} 0\leqslant y\leqslant x, \\ 0\leqslant x\leqslant 1; \end{cases}$

(b) $\begin{cases} y_1(x)\leqslant y\leqslant y_2(x), \\ a\leqslant x\leqslant b. \end{cases}$

【y-型区域】若与 y 轴垂直的直线与 D 相交时，交线是连续的线段或一个点，则称 D 是 y-型区域.

〖**点拨**〗**可以用不等式表示图** 6-1-2 **中的** y-**型区域：**

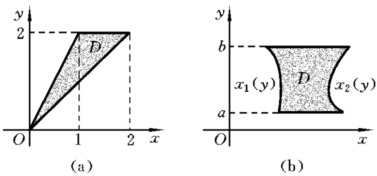


图 6-1-2

$$(\mathrm{a})\begin{cases}\dfrac{1}{2}y\leqslant x\leqslant y,\\ 0\leqslant y\leqslant 2;\end{cases}$$

$$(\mathrm{b})\begin{cases}x_1(y)\leqslant x\leqslant x_2(y),\\ a\leqslant y\leqslant b.\end{cases}$$

【θ-型区域】若从原点出发的射线与 D 相交时,交线是连续的线段或一个点,则称 D 是 θ-型区域.

〖**点拨**〗**可以用不等式表示图 6-1-3 中的 θ-型区域:**

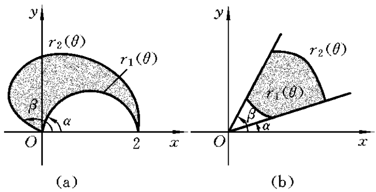


图 6-1-3

$$(\mathrm{a})\begin{cases}2\cos\theta\leqslant r\leqslant 1+\cos\theta,\\ 0\leqslant\theta\leqslant\dfrac{\pi}{2},\end{cases}$$

以及
$$\begin{cases}0\leqslant r\leqslant 1+\cos\theta,\\ \dfrac{\pi}{2}\leqslant\theta\leqslant\pi;\end{cases}$$

$$(\mathrm{b})\begin{cases}r_1(\theta)\leqslant r\leqslant r_2(\theta),\\ \alpha\leqslant\theta\leqslant\beta.\end{cases}$$

【二重积分计算公式 · 直角坐标系】根据积分区域的不同特点选择相应的计算公式,可以将二重积分转换为逐次积分算出.

〖x-型〗 $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_a^b \mathrm{d}x\int_{y_1(x)}^{y_2(x)} f(x,y)\mathrm{d}y.$

〖y-型〗 $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_a^b \mathrm{d}y\int_{x_1(y)}^{x_2(y)} f(x,y)\mathrm{d}x.$

〖θ-型〗 $\iint\limits_D f(x,y)\mathrm{d}\sigma$

$$=\int_\alpha^\beta \mathrm{d}\theta\int_{r_1(\theta)}^{r_2(\theta)} rf(r\cos\theta,r\sin\theta)\mathrm{d}r.$$

注　**当** D **不属于以上三种区域时,可以将** D **适当分割为几块,在每块上求出积分值后相加即成.**

【三重积分计算公式·直角坐标系】

〖xy-型〗**当积分区域**V(**见图**6-1-4)**为**$z_1(x,y)\leqslant z\leqslant z_2(x,y)$,$(x,y)\in D$ **时,有**

$$\iiint\limits_V f(x,y,z)\mathrm{d}V = \iint\limits_D \mathrm{d}x\mathrm{d}y\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\mathrm{d}z.$$

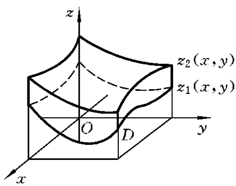


图 6-1-4

〖z-型〗**当积分区域** V(**见图** 6-1-5)**为**$(x,y,z)\in D_z$,$a\leqslant z\leqslant b$ **时,有**

$$\iiint\limits_V f(x,y,z)\mathrm{d}V = \int_a^b \mathrm{d}z\iint\limits_{D_z} f(x,y,z)\mathrm{d}x\mathrm{d}y.$$

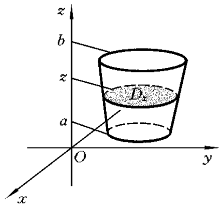


图6-1-5

〖点拨〗当V是z-型区域，且$f=h(z)$仅与z相关时，若用$|D_z|$表示D_z的面积，则z-型区域V上三重积分可直接化作定积分

$$\iiint\limits_V h(z)\mathrm{d}V=\int_a^b h(z)|D_z|\mathrm{d}z.$$

【三重积分计算公式·柱坐标】在以上xy-型及z-型区域中，若其中二维区域D或D_z适合于用极坐标表示，则可以按照θ-型区域要领来计算该二重积分(称(θ,r,z)为柱坐标).例如，对于图6-1-6中积分区域V

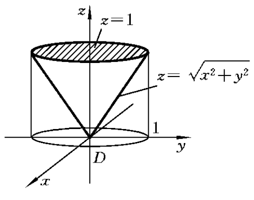


图6-1-6

有

$$\iiint\limits_V f(x,y,z)\mathrm{d}V=\int_0^{2\pi}\mathrm{d}\theta\int_0^1 r\mathrm{d}r\int_r^1 f(r\cos\theta,r\sin\theta,z)\mathrm{d}z$$

(xy-型),

或 $$\iiint\limits_V f(x,y,z)\mathrm{d}V$$

$$=\int_0^1\mathrm{d}z\int_0^{2\pi}\mathrm{d}\theta\int_0^z rf(r\cos\theta,r\sin\theta,z)\mathrm{d}r\quad (z\text{-型}).$$

【三重积分计算公式·球坐标】若V由球面、圆锥面所围,则可以用球面坐标(θ,φ,ρ)来表示区域V. 例如,对于图6-1-7中的区域V,有

$$\iiint\limits_V f(x,y,z)\mathrm{d}V$$

$$=\int_0^{2\pi}\mathrm{d}\theta\int_0^{\pi/4}\mathrm{d}\varphi\int_0^R\rho^2\sin\varphi f(\rho\sin\varphi\cos\theta,\rho\sin\varphi\sin\theta,\rho\cos\varphi)\mathrm{d}\rho.$$

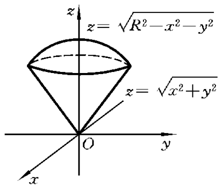


图6-1-7

【对称区域上的重积分化简】积分区域的对称性要与被积函数的奇偶性配套使用.

〖区域的对称性〗平面区域D关于y轴对称时,称为左右对称;关于x轴对称时,称为上下对称.

空间区域V关于Oyz面对称时,称为前后对称;

关于Ozx面对称时,称为左右对称;关于Oxy面对称时,称为上下对称.

〖奇函数积分〗在以下情况下,二重积分$\iint\limits_D f\mathrm{d}\sigma$为零:

(1)$f(x,y)$作为x的函数是奇函数,区域D左右对称;

(2)$f(x,y)$作为y的函数是奇函数,区域D上下对称.

在以下情况下,三重积分$\iiint\limits_V f\mathrm{d}V$为零:

(3)$f(x,y,z)$关于x是奇函数,区域V前后对称;

(4)$f(x,y,z)$关于y为奇函数,区域V左右对称;

(5)$f(x,y,z)$关于z为奇函数,区域V上下对称.

例　$\iint\limits_{x^2+y^2\leqslant 1} xy^2\mathrm{d}\sigma=0$,　$\iiint\limits_{x^2+y^2+z^2\leqslant 1} x\cos yz\mathrm{d}v=0$.

〖偶函数积分〗在以下情形,都有

$$\iint\limits_D f(x,y)\mathrm{d}\sigma = 2\iint\limits_A f(x,y)\mathrm{d}\sigma.$$

(1)区域D左右对称,右半区域记作A,且$f(x,y)$关于x是偶函数;

(2)区域D上下对称,上半区域记作A,且$f(x,y)$关于y是偶函数;

(3)类似地可以写出三重积分的对应公式(略).

注　有时候,积分区域D不是对称区域,但若可以分割出一块对称区域,则可以使用分片积分公式和对称方法化简;同样,如果被积函数不是奇偶函

数,但若它能分拆出含有奇偶函数的加项,则也可以结合分项积分公式应用此方法.

【轮称区域上的重积分化简】

〖区域的轮称性〗

(1)平面区域 D 关于直线 $y=x$ 对称时,称为轮称区域.代数特征:积分域的边界方程不因 x,y 互换而改变.

(2)空间区域 V 关于直线 $x=y=z$ 对称时,称为轮称区域.代数特征:积分域的边界方程不因 x,y,z 互换而改变.

〖轮称区域上的积分〗对二重积分来说,若 D 是轮称区域,则有(相当于作换元 $x=y,y=x$)公式

$$\iint_D f(x,y)\mathrm{d}\sigma=\iint_D f(y,x)\mathrm{d}\sigma.$$

常用方式为

$$\iint_D f(x)\mathrm{d}\sigma=\iint_D f(y)\mathrm{d}\sigma=\frac{1}{2}\iint_D (f(x)+f(y))\mathrm{d}\sigma.$$

例如,$\displaystyle\iint_{x^2+y^2\leqslant 1} x^2\mathrm{d}\sigma=\frac{1}{2}\iint_{x^2+y^2\leqslant 1}(x^2+y^2)\mathrm{d}\sigma.$

三重积分也有类似的变形.

【利用重心坐标计算重积分】将重心坐标计算公式反转即得到重积分计算公式,以二重积分为例:

$$\iint_D x\mathrm{d}\sigma=\overline{x}\,|D|,\quad \iint_D y\mathrm{d}\sigma=\overline{y}\,|D|,$$

其中,$\overline{x},\overline{y}$ 表示积分区域 D 的重心坐标,$|D|$ 表示 D 的面积.

例 $\displaystyle\iint_{(x-a)^2+y^2\leqslant r^2} x\mathrm{d}\sigma=a\cdot\pi r^2,$

$$\iiint\limits_{x^2+y^2+(z-c)^2\leqslant R^2} z\mathrm{d}V = c \cdot \frac{4}{3}\pi R^3.$$

常见题型 · 应对

【换次序计算二次积分】考察二重积分的定限方法，属于基本题. 例如，

$$\int_0^2 \mathrm{d}x\int_x^2 \mathrm{e}^{-y^2}\mathrm{d}y = \int_0^2 \mathrm{d}y\int_0^y \mathrm{e}^{-y^2}\mathrm{d}x = \frac{1}{2}(1-\mathrm{e}^{-4}).$$

【化简和计算二重积分】仅仅考察使用直角坐标或极坐标定限法的计算题较少。通常要结合奇偶函数对称法、轮称区域替换法、分片积分法、分项积分法、积分区域扩展或转换法等方法.

【使用直角坐标计算三重积分】根据区域的特征选择投影法还是截面法.

【使用柱面坐标计算三重积分】依据标准方法计算.

【使用球面坐标计算三重积分】依据标准方法计算.

【常用区域的重心】设以下形体的物体是匀质的，则有如下重心$(\overline{x},\overline{y})$位置：

半圆周 L	$x^2+y^2=R^2, y\geqslant 0$, $(\overline{x},\overline{y})=\left(0,\frac{2}{\pi}R\right)$.
半圆盘 D	$x^2+y^2\leqslant R^2, y\geqslant 0$, $(\overline{x},\overline{y})=\left(0,\frac{4}{3\pi}R\right)$.
半球面 S	$x^2+y^2+z^2=R^2, z\geqslant 0$, $(\overline{x},\overline{y},\overline{z})=(0,0,R/2)$.

续表

半球体 V	$x^2+y^2+z^2\leqslant R^2, z\geqslant 0$ $(\bar{x},\bar{y},\bar{z})=\left(0,0,\frac{3}{8}R\right)$.
圆锥体 V	$\sqrt{x^2+y^2}\leqslant z\leqslant H$ $(\bar{x},\bar{y},\bar{z})=\left(0,0,\frac{1}{4}H\right)$.

6.2　线面积分

大纲要求

理解两类曲线积分的概念. 了解两类曲线积分的性质及两类曲线积分的关系. 掌握计算两类曲线积分的方法. 掌握格林公式并会运用平面曲线积分与路径无关的条件,会求二元函数全微分的原函数.

了解两类曲面积分的概念、性质及两类曲面积分的关系,掌握计算两类曲面积分的方法,会用高斯公式、斯托克斯公式计算曲面、曲线积分.

了解散度与旋度的概念,并会计算.

注　数学二不要求本节内容.

知识点·点拨

【第一型线积分定义】从密度为 $f(x,y,z)$ 的曲线 L 的质量计算问题中导出

$$\int_L f(x,y,z)\mathrm{d}s=\lim_{\lambda\to 0}\sum_{i=1}^{n}f(x_i,y_i,z_i)\Delta S_i.$$

【第二型线积分定义】从沿有向曲线 L 移动物体,变力 $\boldsymbol{F}=\{P,Q,R\}$ 做功问题中导出

$$\int_L \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \lim_{\lambda \to 0} \sum_{i=1}^{n} \boldsymbol{F}(x_i, y_i, z_i) \cdot \Delta \boldsymbol{r}_i,$$

也记作
$$\int_L P\mathrm{d}x + Q\mathrm{d}y + Q\mathrm{d}z.$$

【第一型面积分定义】从密度为 $f(x,y,z)$**的曲面** S **的质量计算问题中导出**

$$\iint_S f(x,y,z)\mathrm{d}S = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i)\Delta S_i.$$

【第二型面积分定义】从通过曲面 S **一侧,流速为** $\boldsymbol{V}=\{P,Q,P\}$**的物质的流量计算问题中导出**

$$\iint_S \boldsymbol{V} \cdot \mathrm{d}\boldsymbol{S} = \lim_{\lambda \to 0} \sum_{1}^{n} \boldsymbol{V}(x_i, y_i, z_i) \cdot \boldsymbol{n}_i \Delta S_i,$$

也记作
$$\iint_S P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y.$$

【两类线积分关系】若 τ **是曲线** L **的与其定向一致的单位切向量,则沿曲线** L **的两种线积分的关系为**

$$\int_L \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \int_L (\boldsymbol{F} \cdot \boldsymbol{\tau})\mathrm{d}s.$$

〖点拨〗由于第二型线积分计算有较多的简化方法,上述转化公式不太常用.

【两类面积分的关系】若 $\boldsymbol{n}$ **是曲面** S **的与指定定侧一致的单位法向量,则**

$$\iint_S P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y = \iint_S \{P,Q,R\} \cdot \boldsymbol{n}\mathrm{d}S.$$

〖点拨〗通常用等式右边的积分来计算左边的积分.

【格林公式】若函数 $P(x,y)$,$Q(x,y)$**在平面区域** D **及其边界** L**(逆时针方向)上有连续的一阶偏导数,**

则有

$$\oint_L P\mathrm{d}x + Q\mathrm{d}y = \iint_D (Q_x - P_y)\mathrm{d}x\mathrm{d}y.$$

〖点拨〗此公式主要用于计算闭曲线或(通过增补法)弓形曲线上的第二型线积分.

【斯托克斯公式】若函数 $P(x,y,z)$,$Q(x,y,z)$ 及 $R(x,y,z)$ 在定向曲面 S 及其边界曲线 L(L 的定向与 S 的定向成右手螺旋法则)上有连续的一阶偏导数,则有

$$\oint_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = \iint_D \mathbf{rot}\{P,Q,R\}\cdot \boldsymbol{n}\mathrm{d}S.$$

〖点拨〗此公式主要用于计算空间闭曲线上的第二型线积分.由于以 L 为边界的曲面 S 不惟一,通常选择平面或球面等作 S,以便将线积分化做简单的面积分.

【高斯公式】若函数 $P(x,y,z)$,$Q(x,y,z)$ 及 $R(x,y,z)$ 在空间区域 V 及其表面 S(外侧定向)上有连续的一阶偏导数,则有

$$\iint_S P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y$$
$$= \iiint_V (P_x + Q_y + R_z)\mathrm{d}V.$$

〖点拨〗此公式主要用于计算封闭曲面或(通过增补法)碗形曲面上的第二型曲面积分.

【线积分的N-L 公式】设曲线 L 以 A 为起点,B 为终点.在 L 上,函数 P,Q,R 连续.若有可微函数 $u(x,y,z)$,使得 $\mathrm{d}u=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$,则有计算公式

$$\int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = u(B) - u(A).$$

〖点拨1〗函数 $u(x,y,z)$ 称为线积分的原函数，并非每个线积分都存在这样的原函数. 存在原函数的必要条件(积分区域为线单连通时为充分条件)是

$$\mathbf{rot}\{P,Q,R\} = 0 \quad (\text{空间线积分情形}),$$

$$Q_x = P_y \quad (\text{平面线积分情形}).$$

〖点拨2〗在原函数存在时，可以用不定积分法或凑微分法求得函数 u.

例1　$\int_{(0,0,0)}^{(1,1,1)} x\mathrm{d}x + y\mathrm{d}y + z\mathrm{d}z$.

解　使用凑微分法找原函数.

$$\text{原式} = \int_{(0,0,0)}^{(1,1,1)} \mathrm{d}(x^2 + y^2 + z^2)/2$$

$$= \frac{1}{2}(x^2 + y^2 + z^2)\Big|_{(0,0,0)}^{(1,1,1)} = \frac{3}{2}.$$

【化作定积分计算两类线积分】若积分曲线 L 的参数方程为 $x=x(t), y=y(t), z=z(t)\ (a\leqslant t\leqslant b)$，则 L 上的两类线积分可依下式化做定积分.

〖第一型〗

$$\int_L f(x,y,z)\mathrm{d}s$$

$$=\int_a^b f(x(t),y(t),z(t))\sqrt{x'(t)^2+y'(t)^2+z'(t)^2}\,\mathrm{d}t.$$

〖第二型〗

$$\int_L P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$$

$$=\pm\int_a^b [Px'(t)+Qy'(t)+Rz'(t)]\mathrm{d}t.$$

其中 $P=P(x(t),y(t),z(t))$. Q,R 类似.

当 L 的起点与参数 $t=a$ 对应时取正号,否则取负号.

〖点拨〗以上公式亦适用于两类平面线积分的计算.

【改变积分路径计算第二型平面线积分】如图 6-2-1 所示,若在同向的曲线 L_1,L_2 之间的区域 D 上成立 $Q_x=P_y$,则有

$$\int_{L_1} P\mathrm{d}x + Q\mathrm{d}y = \int_{L_2} P\mathrm{d}x + Q\mathrm{d}y.$$

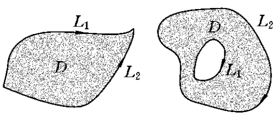


图 6-2-1

〖点拨〗利用此方法可以将线积分改变到一个较简单的曲线(折线段、圆或使被积函数变简单的曲线)上计算.

例　如图 6-2-2 所示,计算 $I=\oint_L \dfrac{y\mathrm{d}x-x\mathrm{d}y}{x^2+y^2}$,$L$ 是 $|x|+|y|=1$ 的正向.

解　除原点外可以验证 $Q_x=P_y$,于是可以将积分计算改变到 $C:x^2+y^2=1$(正向)上进行,故

$$I=\oint_C \frac{y\mathrm{d}x - x\mathrm{d}y}{x^2+y^2}$$

$$=\oint_C y\mathrm{d}x - x\mathrm{d}y \quad (x^2+y^2=1\ \text{代入})$$

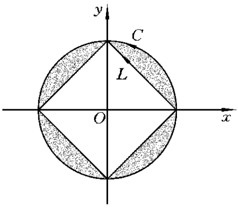


图 6-2-2

$$= \iint_D -2\mathrm{d}x\mathrm{d}y = -2\pi \quad \text{（格林公式）}.$$

〖点拨 1〗此题不可以直接使用格林公式，因为 P,Q 在 L 所围区域内的原点处无定义，但是在转换到曲线 C 上并借助等量代换法消化了分母 x^2+y^2 之后，便可以使用格林公式了.

〖点拨 2〗虽然函数 $u=\arctan\dfrac{x}{y}$ 是本题中被积函数式的原函数，但是不能由 N-L 公式，得

$$I=\int_L \frac{y\mathrm{d}x-x\mathrm{d}y}{x^2+y^2}=u(B)-u(A)=0,$$

错误原因在于在 L 上存在函数 u 无定义的点 $y=0$，因而不可以应用线积分 N-L 公式.

【改变积分曲面计算第二型面积分】如图 6-2-3 所示，若在同向的曲面 S_1 与 S_2 之间的空间区域 V 上成立

$$\operatorname{div}\{P,Q,R\}=0,$$

则有 $\iint\limits_{S_1} \boldsymbol{F} \cdot \boldsymbol{n} \mathrm{d}S = \iint\limits_{S_2} \boldsymbol{F} \cdot \boldsymbol{n} \mathrm{d}S \quad (\boldsymbol{F} = \{P, Q, R\})$.

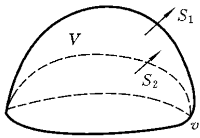


图 6-2-3

〖点拨〗利用此方法可以将面积分改变到一个使积分容易计算的曲面上进行.

例　计算引力场 $\boldsymbol{F} = -m\boldsymbol{r}/r^3$ 通过包围原点的闭曲面 S 的外侧的通量 Φ.

解　除原点外可验证 $\mathrm{div}\boldsymbol{F} = 0$,于是可以将积分运算改变到单位球面 S_0 的外侧上进行,注意到 S_0 的法向量 $\boldsymbol{n} = \boldsymbol{r}$,方程为 $r^2 = 1$,故有

$$\Phi = \iint\limits_{S} \boldsymbol{F} \cdot \boldsymbol{n} \mathrm{d}S = \iint\limits_{S_0} - m\boldsymbol{r} \cdot \boldsymbol{r} \mathrm{d}S$$

$$= - m \iint\limits_{S_0} \mathrm{d}S = - 4\pi.$$

【空间第二型线积分计算方法归纳】

$$I = \int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z.$$

(1) 化作定积分计算:这是基本方法,只需 P,Q,R 在 L 上连续便可,需要知道 L 的参数方程.

(2) 化作第一型面积分计算:要求 L 是闭曲线,选择一个适当的以 L 为边界的曲面 S,如平面或球

面,使用斯托克斯公式转化为第一型面积分(求出 $\boldsymbol{n}$ 并进行点乘)后计算.

(3) **化作平面线积分计算**:若在积分式中代入曲线 L 的方程后,可以消除变量 z,则空间线积分可以化作 L 在 Oxy 平面上的有向投影曲线 C 上的平面线积分,再依据平面线积分的求法计算.

(4) **改变积分路径计算**:若积分曲线 L 位于使得 $\mathbf{rot}\{P,Q,R\}=\mathbf{0}$ 的面单连通的区域中,则可以用改变路径法、N-L 公式法来计算积分.

例　设 L 是柱面 $|x|+|y|=1$ 和平面 $x+y+z=2$ 的交线,从 z 轴正向看,L 取逆时针向,求

$$I=\int_L (y^2-x^2)\mathrm{d}x+(2z^2-x^2)\mathrm{d}y+(3x^2-y^2)\mathrm{d}z.$$

解 1(斯托克斯公式)　选择以 L 为边界的平面 $S: x+y+z=2$ 的上侧,则有单位法向量 $\boldsymbol{n}_S=\{1,1,1\}/\sqrt{3}$,于是由斯托克斯公式得

$$\begin{aligned}I&=\frac{1}{\sqrt{3}}\iint_S\{-2y-4z,-2z-6x,-2x-2y\}\cdot\{1,1,1\}\mathrm{d}S\\&=-\frac{1}{\sqrt{3}}\iint_S(8x+4y+6z)\mathrm{d}S\\&\xlongequal{(1)}\frac{-1}{\sqrt{3}}\iint_S(6x+6y+6z)\mathrm{d}S\\&\xlongequal{(2)}-\frac{1}{\sqrt{3}}\iint_S 12\mathrm{d}S\\&\xlongequal{(3)}-\frac{12}{\sqrt{3}}\iint_D\sqrt{3}\,\mathrm{d}x\mathrm{d}y=-12|D|=-24.\end{aligned}$$

其中,(1)由于S关于x,y轮称,故$\iint\limits_S 2x\mathrm{d}S=\iint\limits_S 2y\mathrm{d}S$;(2)将$S$的方程$x+y+z=2$代入,被积式化简;(3)化作二重积分.

解2(化作平面线积分) 记C为L在Oxy平面上的投影曲线,其方向继承L,为逆时针方向,代入恒等关系式

$$z=2-x-y,$$

得
$$\begin{aligned}I=&\int_L[y^2-(2-x-y)^2-3x^2+y^2]\mathrm{d}x\\&+[2(2-x-y)^2-x^2-3x^2+y^2]\mathrm{d}y\\\xlongequal{(1)}&\iint\limits_C(-12-2x+2y)\mathrm{d}x\mathrm{d}y\\\xlongequal{(2)}&-12\iint\limits_D\mathrm{d}x\mathrm{d}y=-24.\end{aligned}$$

其中,(1)使用格林公式;(2)奇函数在对称区域的二重积分为零.

【平面第二型线积分计算方法归纳】

$$I=\int_L P\mathrm{d}x+Q\mathrm{d}y.$$

(1) 化作定积分计算:这是基本方法,只需要知道L的参数方程.

(2) 化作二重积分计算:当L是闭曲线且被积函数在其内部有连续偏导时,首先考虑应用格林公式求解;当L是弓形曲线时,可考虑"加边法",增补一段曲线C,使$L\cup C$为闭曲线,再调用格林公式计算其中的闭曲线积分,即

$$\int_L \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r}=\int_{L\cup C}\boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r}-\int_C \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r}.$$

（3）**借助原函数计算：若积分曲线** L **位于使得** $Q_x=P_y$ **的线单连通的区域中，则可以用** N-L **公式法来计算积分.**

（4）**改变路径计算：要求** $Q_x=P_y$ **在两条路径所围的闭区域上成立.**

【第二型面积分计算方法归纳】

$$\Phi=\iint_S P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y.$$

（1）**使用高斯公式化作三重积分：适合于** S **为闭曲面，要求被积函数在** S **及其内部有连续偏导；当** S **为碗形时，可考虑"加盖法"，增加一块曲面** π**，再调用高斯公式计算其中的闭曲面积分，即**

$$\iint_S \boldsymbol{F}\cdot \boldsymbol{n}\mathrm{d}S=\oiint_{S\cup\pi}\boldsymbol{F}\cdot \boldsymbol{n}\mathrm{d}S-\iint_\pi \boldsymbol{F}\cdot \boldsymbol{n}\mathrm{d}S.$$

（2）**化作第一型面积分：适合于** S **为单片曲面，求出** S **的单位法向量** $\boldsymbol{n}$**，与** $\boldsymbol{F}$ **点乘后便化作** $\boldsymbol{F}\cdot\boldsymbol{n}$ **的第一型面积分.**

（3）**化作二重积分：适合于向所对应的坐标平面投影为单影的曲面** S**. 通过分项法，可以化作曲面** S **在指定坐标平面上的投影域上的二重积分，转化公式分为三种.**

① $\displaystyle\iint_S P(x,y,z)\mathrm{d}y\mathrm{d}z$

$$=\pm\iint_{D_{yz}} P(x(y,z),y,z)\mathrm{d}y\mathrm{d}z,$$

要求S **能写成** $x=x(y,z),(y,z)\in D_{yz}$,**当** S **的定侧法向量与** x **轴成锐角时,取正号,否则取负号.**

② $\iint\limits_{S}Q(x,y,z)\mathrm{d}z\mathrm{d}x$

$$=\pm\iint\limits_{D_{zx}}Q(x,y(z,x),z)\mathrm{d}z\mathrm{d}x,$$

要求 S **能写成** $y=y(z,x),(z,x)\in D_{zx}$,S **的定侧法向量与** y **轴成锐角时取正号,否则取负号.**

③ $\iint\limits_{S}R(x,y,z)\mathrm{d}x\mathrm{d}y$

$$=\pm\iint\limits_{D_{xy}}R(x,y,z(x,y))\mathrm{d}x\mathrm{d}y,$$

要求 S **能写成** $y=y(z,x),(z,x)\in D_{zx}$,S **的定侧法向量与** y **轴成锐角时取正号,否则取负号.**

【第一型面积分的计算公式】可以根据曲面的特点,选择适当的投影坐标面,将面积分化作二重积分计算.针对投影区域的不同,有以下三种平行公式:

①**设曲面**S **与其在**Oxy **面上的投影域**D **是一一对应的,即**$S:z=z(x,y),(x,y)\in D$,**则**

$$\iint\limits_{S}f(x,y,z)\mathrm{d}S$$

$$=\iint\limits_{D}f(x,y,z(x,y))\sqrt{1+z_x^2+z_y^2}\mathrm{d}x\mathrm{d}y.$$

②**设曲面**$S:x=x(y,z),(y,z)\in D$,**则**

$$\iint\limits_{S}f(x,y,z)\mathrm{d}S$$

$$=\iint\limits_{D}f(x(y,z),y,z)\sqrt{1+x_y^2+x_z^2}\mathrm{d}y\mathrm{d}z.$$

③**设曲面** $S: y=y(z,x),(z,x)\in D$,**则**

$$\iint_S f(x,y,z)\mathrm{d}S$$

$$=\iint_D f(x,y(z,x),z)\sqrt{1+y_x^2+y_z^2}\mathrm{d}z\mathrm{d}x.$$

〖**点拨**〗**如果曲面**S **与其在坐标面上的投影不是一一对应的,可以考虑将** S **分割后,分片计算面积分.例如**

$$\iint_{x^2+y^2+z^2=R^2} f\mathrm{d}S=\iint_{\substack{x^2+y^2+z^2=R^2\\(z>0)}} f\mathrm{d}S+\iint_{\substack{x^2+y^2+z^2=R^2\\(z<0)}} f\mathrm{d}S.$$

【线面积分计算的快捷法】

(1)**等量代换法.将曲线** L **或曲面** S **的方程式代入积分式中进行化简.要注意的是,重积分中不能使用这一快捷法.**

(2)**几何意义法.** $\int_L \mathrm{d}s=L$ **的弧长,** $\iint_S \mathrm{d}S=S$ **的面积.**

(3)**奇偶对称法.若** $f(x,y,z)$ **关于** x **为奇函数(或偶函数),积分区域**L **和**S **关于**Oyz **面对称,则对第一型线积分和面积分,有(**L_1,S_1 **分别是** L,S **的一半)**

$$\int_L f\mathrm{d}s=0,\quad \iint_S f\mathrm{d}S=0$$

(**或**　$\int_L f\mathrm{d}s=2\int_{L_1} f\mathrm{d}s,\quad \iint_S f\mathrm{d}S=2\iint_{S_1} f\mathrm{d}S$).

类似地,可以写出关于y **和关于**z **为奇偶函数的结果(略去).**

(4)**轮换对称法.若积分区域** L **和** S **中** x,y,z **的**

地位对称,则在被积式中互换 x,y,z,结果不变,例如:

$$\int_L f(x)\mathrm{d}s=\int_L f(y)\mathrm{d}s=\int_L f(z)\mathrm{d}s,$$

$$\iint_S f(x)\mathrm{d}S=\iint_S f(y)\mathrm{d}S=\iint_S f(z)\mathrm{d}S.$$

例　曲线 L:$\begin{cases}x+y+z=1,\\x^2+y^2=1\end{cases}$ 中 x,y 位置对称,故有以下化简法:

$$\int_L(3x+y+2z)\mathrm{d}s=2\int_L(x+y+z)\mathrm{d}s=2\int_L \mathrm{d}s.$$

(5)重心坐标法. 记$(\bar{x},\bar{y},\bar{z})$为曲线 L 或曲面 S 的重心坐标. $|L|$和$|S|$为它们的弧长或面积,则有

$$\int_L x\mathrm{d}s=\bar{x}|L|,\quad \int_L y\mathrm{d}s=\bar{y}|L|,\quad \int_L z\mathrm{d}s=\bar{z}|L|,$$

$$\iint_S x\mathrm{d}S=\bar{x}|S|,\quad \iint_S y\mathrm{d}S=\bar{y}|S|,\quad \iint_S z\mathrm{d}S=\bar{z}|S|.$$

方法(2)~(5)只适合于第一型积分.

【线单连通区域】称 Ω 是线单连通区域,若 Ω 中任何闭曲线可以在 Ω 内连续地缩成一个点.

对于二维平面区域 D 来说,D 中无孔是线单连通的充要条件;对于三维空间区域 Ω 来说,内部仅含“气泡”(如球体中去掉球心)时为线单连通区域,但若有贯通的“孔”(如车胎体),则不是线单连通区域.

【面单连通区域】称空间区域 Ω 是面单连通区域,若 Ω 中任何闭曲面可以在 Ω 内连续地缩成一个点.

内部含有“气泡”的区域不是面单连通区域,但是仅含有贯通的“孔”的区域都是面单连通区域.

【梯度】 数量函数 $f(x,y,z)$ 的梯度是一个矢量：

$$\nabla f=\mathbf{grad}f=\{f_x,f_y,f_z\}.$$

【散度】 矢量场 $\boldsymbol{F}=\{P,Q,R\}$ 的散度是一个数量：

$$\nabla\cdot\boldsymbol{F}=\mathrm{div}\boldsymbol{F}=P_x+Q_y+R_z.$$

【旋度】 矢量场 $\boldsymbol{F}=\{P,Q,R\}$ 的旋度是一个矢量：

$$\nabla\times\boldsymbol{F}=\mathbf{rot}\boldsymbol{F}=\{R_y-Q_z,P_z-R_x,Q_x-P_y\}.$$

常见题型·应对

直接计算所给的线面积分；考察基本方法(化作定积分，重积分计算)；考察化简的方法(改变路径，使用对称性、轮称性、代入法等等)；考察三个重要的积分公式：格林公式、斯托克斯公式、高斯公式.

应对的策略很简单，只要将每一个独立的知识点掌握便可以，题型变化不是太大，但是一道题涉及的考点较多，比如面积分计算，可能会用到高斯公式、三重积分、定积分等等.

6.3　多元函数积分的应用

大纲要求

会用重积分、曲线积分及曲面积分求一些几何量与物理量(平面图形的面积、体积、曲面面积、弧长、质量、重心、转动惯量、引力、功及流量等).

注　本节数学二不要求.

知识点·点拨

几何应用公式

【曲线 L 的弧长】　$s=\int_L \mathrm{d}s.$

【曲面 S 的面积】　$S=\iint\limits_{S}\mathrm{d}S.$

【平面区域 D 的面积】　$S=\iint\limits_{D}\mathrm{d}x\mathrm{d}y.$

【空间区域 V 的体积】　$V=\iiint\limits_{V}\mathrm{d}x\mathrm{d}y\mathrm{d}z.$

〖点拨 1〗**利用定积分可以计算平面曲线弧长、曲边梯形面积、绕 x 轴及绕 y 轴旋转体体积和表面积、平行截面体体积(见 3.3 节),尽管这些公式能够由以上多元积分表示,但是定积分公式应当是首选.**

〖点拨 2〗**几何应用属基本内容,除了正确地套用公式、作图、定限之外,计算出积分值无疑是关键和考点所在.**

例 1　**求曲线 $y=xe^{x}$ 与 $y=ex$ 所围区域面积 S.**

解　由 $(xe^{x})'=e^{x}(x+1)>0$ 及 $(xe^{x})''=e^{x}(x+2)>0$ 知,曲线 $y=xe^{x}$ 单调增加、下凸,且过原点,交直线 $y=ex$ 于 $x=0$ 及 $x=1$ 处,如图6-3-1 所示. 于是有

$$S=\iint\limits_{D}\mathrm{d}x\mathrm{d}y=\int_{0}^{1}\mathrm{d}x\int_{xe^{x}}^{ex}\mathrm{d}y$$

$$=\int_{0}^{1}(ex-xe^{x})\mathrm{d}x=\frac{e}{2}-1.$$

例 2　**求曲线 $y^{2}=x^{2}(a^{2}-x^{2})$ 所围区域的面积 S $(a>0)$.**

解　由于曲线方程中只含 x^{2} 及 y^{2},故曲线所围区域关于 x 轴及 y 轴对称,由于 $a^{2}-x^{2}\geqslant0$,故曲线的范围为 $-a\leqslant x\leqslant a$. 于是只用考虑在第一象限中的区

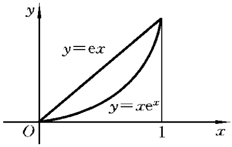


图 6-3-1

域的面积,使用定积分公式计算,有

$$S=4\int_0^a y(x)\mathrm{d}x=4\int_0^a x\sqrt{a^2-x^2}\mathrm{d}x=\frac{4}{3}a^3.$$

例 3　求曲线 $r=a(1+\cos\theta)$ 所围平面区域的面积 S.

解　从 $r\geqslant 0$ 来求出 θ 的范围是 $0\leqslant\theta\leqslant 2\pi$,由 $r(-\theta)=r(\theta)$ 知,曲线 $r=r(\theta)$ 关于 x 轴对称,故

$$S=\iint_D \mathrm{d}x\mathrm{d}y=2\int_0^{\pi}\mathrm{d}\theta\int_0^{a(1+\cos\theta)}r\mathrm{d}r$$

$$=\int_0^{\pi}a^2(1+\cos\theta)^2\mathrm{d}\theta=\frac{3}{2}\pi a^2.$$

物理应用公式

下文中 Ω 代表曲线 L、曲面 S、平面区域 D、空间区域 V.

【质量】密度为 $\rho(P)$ 的物体 Ω 的质量为 $m_\Omega=\int_\Omega\rho\mathrm{d}\Omega$. 当 Ω 依次取曲线 L、曲面 S、平面区域 D、空间区域 V 时,有

$$m_L=\int_L\rho\mathrm{d}s,\quad m_S=\iint_S\rho\mathrm{d}S,$$

$$m_D = \iint_D \rho \mathrm{d}\sigma, \quad m_V = \iiint_V \rho \mathrm{d}V.$$

【重心】密度为$\rho(P)$的物体Ω的重心为$M(\overline{x},\overline{y},\overline{z})$，其中

$$\overline{x} = \frac{1}{m}\int_\Omega x\rho \mathrm{d}\Omega, \quad \overline{y} = \frac{1}{m}\int_\Omega y\rho \mathrm{d}\Omega,$$

$$\overline{z} = \frac{1}{m}\int_\Omega z\rho \mathrm{d}\Omega.$$

m是Ω的质量，如$\Omega=D$为二维区域时，有

$$\overline{x} = \frac{1}{m}\iint_D x\rho \mathrm{d}\sigma, \quad \overline{y} = \frac{1}{m}\iint_D y\rho \mathrm{d}\sigma.$$

【转动惯量】物体Ω绕着定轴L的转动惯量为

$$J_L = \int_\Omega d(P,L)^2 \rho g \mathrm{d}\Omega.$$

其中，ρ是密度函数，$d(P,L)$是点$P\in\Omega$到直线L的距离，例如：

(1)设Ω为平面区域，取L为y轴，则点$P(x,y)$到L的距离的平方为x^2，从而物体D关于y轴的转动惯量为

$$J_y = \iint_D x^2 \rho g \mathrm{d}x\mathrm{d}y.$$

类似地，物体D关于x轴的转动惯量为

$$J_x = \iint_D y^2 \rho g \mathrm{d}x\mathrm{d}y.$$

(2)设$\Omega=V$为空间区域，取L为z轴，则$d(P,L)^2=x^2+y^2$，物体V关于z轴的转动惯量为

$$J_z = \iiint_V (x^2+y^2)\rho g \mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

类似地,可以写出关于其他坐标轴的转动惯量.

【引力】三维物体 Ω 对位于点 $P_0(x_0,y_0,z_0)$ 的单位质点所产生的引力为

$$\boldsymbol{F}=\iiint_{\Omega}\frac{k\rho g}{r^3}\boldsymbol{r}\mathrm{d}V,$$

其中,ρ 是 Ω 的密度,$\boldsymbol{r}=\overrightarrow{P_0P}$,$r=|\boldsymbol{r}|$,$k$ 为比例系数.

此处,向量函数的重积分约定为对分量的重积分:

$$\iiint_{\Omega}\{P,Q,R\}\mathrm{d}V=\left\{\iiint_{\Omega}P\mathrm{d}V,\iiint_{\Omega}Q\mathrm{d}V,\iiint_{\Omega}R\mathrm{d}V\right\}.$$

【变力做功】在力场 $\boldsymbol{F}=\{P,Q,R\}$ 作用下,物体沿有向曲线 L 移动过程中系统做功大小为

$$W=\int_L\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}=\int_L P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z,$$

对于平面力场 $\boldsymbol{F}=\{P,Q\}$,则有

$$W=\int_L P\mathrm{d}x+Q\mathrm{d}y.$$

〖点拨 1〗若 $\boldsymbol{F}$ 是线单连通区域 Ω 内的梯度场,$\boldsymbol{F}=\mathbf{grad}u$,从 A 到 B 的曲线 L_{AB} 位于 Ω 内,则做功 W 与路径无关,即有函数 $u(x,y,z)$,使得

$$\int_{L_{AB}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}=u(B)-u(A),$$

特别地,
$$\oint_L\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}=0.$$

〖点拨 2〗若在应用问题中没有写出 $\boldsymbol{F}$ 的表示式,则可以依据条件,通过分别计算向量的模与单位方向来构建 $\boldsymbol{F}=|\boldsymbol{F}|\boldsymbol{F}^0$.

例 设平面力场 $\boldsymbol{F}$ 指向原点，$\boldsymbol{r}=x\boldsymbol{i}+y\boldsymbol{j}$，$|\boldsymbol{F}|$ 与 $r=|\boldsymbol{r}|$ 成正比. 一质点沿椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 从点 $(a,0)$ 移动至点 $(0,b)$，求 $\boldsymbol{F}$ 对质点所做的功 W.

解 依题意有 $\boldsymbol{F}=-k\boldsymbol{r}$，$k$ 是正常数，于是

$$W=\int_L \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r}=-k\int_L x\mathrm{d}x+y\mathrm{d}y$$

$$=-\frac{k}{2}\int_L \mathrm{d}(x^2+y^2)=\frac{k}{2}(x^2+y^2)\Big|_{(0,b)}^{(a,0)}$$

$$=\frac{k}{2}(a^2-b^2).$$

【通量】 在流速场 $\boldsymbol{V}=\{P,Q,R\}$ 中，通过简单曲面 S 一侧的流量大小为

$$\Phi=\iint_S \boldsymbol{F}\cdot\boldsymbol{n}^0\mathrm{d}S=\iint_S P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}y+R\mathrm{d}x\mathrm{d}y.$$

〖点拨〗 若 $\boldsymbol{V}$ 是面单连通域 Ω 内的旋度场，$\boldsymbol{V}=\mathbf{rot}\boldsymbol{F}$，曲面 S 位于 Ω 内，则通量 Φ 只与 S 的边界相关，即

$$\Phi=\iint_S \boldsymbol{V}\cdot\boldsymbol{n}^0\mathrm{d}S=\oint_L \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r},$$

特别地，
$$\oiint_S \boldsymbol{V}\cdot\boldsymbol{n}^0\mathrm{d}S=0.$$

第7章　无穷级数

7.1　数项级数

大纲要求

理解常数项级数收敛、发散以及收敛级数的和的概念，掌握级数的基本性质及收敛的必要条件，掌握几何级数及 p-级数的收敛与发散的条件. 掌握正项级数收敛性的比较判别法和比值判别法，会用根值判别法. 掌握交错级数的莱布尼兹判别法. 了解任意项级数绝对收敛与条件收敛的概念以及绝对收敛与收敛的关系.

注　数学二不要求本章.

知识点·点拨

【部分和】级数 $\sum\limits_{n=1}^{\infty} a_n$ 的前 n 项之和 $S_n=\sum\limits_{i=1}^{n} a_i$ 称为此级数的部分和.

【收敛与发散】若部分和数列 S_n 收敛于实数 S，则称级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛，和为 S. 否则便称级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

【收敛的必要条件】若 $\sum\limits_{n=1}^{\infty} a_n$ 收敛，则 $\lim\limits_{n\to\infty} a_n=0$，进而数列 a_n 是有界数列.

〖点拨〗若级数的通项不趋于零，则级数一定发散.

【常用级数及其敛散性】

(1) **摆动级数**：$\sum\limits_{n=0}^{\infty}(-1)^n$，**发散**；

(2) **调和级数**：$\sum\limits_{n=1}^{\infty}\dfrac{1}{n}$，**发散**；

(3) **等比级数**：$\sum\limits_{n=0}^{\infty}r^n$，**当** $|r|<1$ **时绝对收敛，其他情形发散**；

(4) p-**级数**：$\sum\limits_{n=1}^{\infty}\dfrac{1}{n^p}$，**当** $p>1$ **时收敛，其他情形发散.**

【收敛级数的基本性质】设级数 $\sum\limits_{n=1}^{\infty}a_n$ **与** $\sum\limits_{n=1}^{\infty}b_n$ **皆收敛，其和分别为** A **与** B.

性质Ⅰ　**对常数** α,β，$\sum\limits_{n=1}^{\infty}(\alpha a_n+\beta b_n)$**收敛，且其和为** $\alpha A+\beta B$.

性质Ⅱ　**若** $a_n\leqslant b_n(n=1,2,\cdots)$，**则**

$$\sum_{n=1}^{\infty}a_n\leqslant\sum_{n=1}^{\infty}b_n.$$

性质Ⅲ(结合律)　**收敛级数** $a_1+a_2+\cdots+a_n+\cdots$ **中可任意地加括号而不影响其收敛性与和.**

〖**点拨**〗**发散级数加括号后有可能收敛，如摆动级数.**

性质Ⅳ　**改变(或增减)收敛级数** $\sum a_n$ **的有限项，不影响其收敛性(但其和可能改变). 因此，级数敛散性的研究可以从某项之后开始.**

【级数加括号问题】对级数中的各项实行加括号，会产生新的级数. 新级数与原级数的敛散关系如下：

(1) 原级数收敛⇒加括号级数收敛；

(2) 加括号级数发散⇒原级数发散；

(3) 加括号级数收敛且原级数通项趋于零⇔原级数收敛.

【正项级数及其特点】当 $\sum a_n$ 中每一项为非负时称此级数为正项级数. 它的部分和数列 $\{S_n\}$ 是单调增加的.

【正项级数·收敛的充要条件】正项级数 $\sum a_n$ 收敛的充分必要条件是 $\{S_n\}$ 有上界.

〖点拨〗相对来说，判定有界性较容易，由此得到一系列的敛散判别法.

〖子级数〗$\{a_n\}$ 的子列构成的无穷级数称为 $\sum a_n$ 的子级数. 收敛的正项级数的子级数也收敛.

【正项级数·比较判别法】若自某项之后有 $0 \leqslant a_n \leqslant b_n$，则当 $\sum b_n$ 收敛时，$\sum a_n$ 也收敛；当 $\sum a_n$ 发散时，$\sum b_n$ 也发散.

简言之，大的收敛⇒小的收敛，小的发散⇒大的发散.

【正项级数·比阶法】通项是同阶无穷小量的两个正项级数有相同的敛散性.

〖点拨〗变号级数没有这一性质. 例如，$a_n = \dfrac{(-1)^n}{\sqrt{n}}$ 与 $b_n = \dfrac{(-1)^n}{\sqrt{n}} + \dfrac{1}{n}$ 是等价无穷小量，但是

$\sum a_n$ 收敛，$\sum b_n$ 发散(作为收敛级数与发散级数的和).

【正项级数· 积分判别法】设函数 $f(x)$ 连续，则通项 $a_n=f\left(\frac{1}{n}\right)$ 单调减少时，正项级数 $\sum a_n$ 与广义积分 $\int_1^{+\infty} f(x)\mathrm{d}x$ 有相同的敛散性.

【正项级数· 比值法】若 $\lim\limits_{n\to\infty}\frac{a_{n+1}}{a_n}=l$，则正项级数 $\sum a_n$ 当 $l<1$ 时收敛，当 $l>1$ 时发散.

【正项级数· 根值法】若 $\lim\limits_{n\to\infty}\sqrt[n]{a_n}=l$，则正项级数 $\sum a_n$ 当 $l<1$ 时收敛，当 $l>1$ 时发散.

〖点拨 1〗当上述两个判别法中的 $l=1$ 时，级数既可能是收敛的，如 $\sum\frac{1}{n^2}$，也可能是发散的，如 $\sum\frac{1}{n}$.

〖点拨 2〗通常，比值法适合于通项中出现阶乘记号的级数. 根值法适合于出现 n 次幂的情形.

【变号级数】含有无限多个正项及无限多个负项的级数.

【绝对收敛】若级数 $\sum|a_n|$ 收敛，则称级数 $\sum a_n$ 绝对收敛.

【绝对收敛定理】若级数 $\sum a_n$ 绝对收敛，即 $\sum|a_n|$ 收敛，则级数 $\sum a_n$ 收敛.

〖点拨〗若根据比值法得出 $\sum|a_n|$ 发散，则可以断定 $\sum a_n$ 发散(因为 $|a_n|$ 单调增加，使得 a_n 不趋

于0).

【条件收敛】若级数$\sum |a_n|$发散,同时级数$\sum a_n$收敛,则称$\sum a_n$为条件收敛.

【变号级数的叠加性质】

绝对收敛+绝对收敛=绝对收敛;

绝对收敛+条件收敛=条件收敛;

条件收敛+条件收敛=收敛.

【莱布尼兹判别法】当a_n单调减少且趋于零时,交错级数$\sum (-1)^n a_n$为收敛级数.

常见题型·应对

【判定级数敛散的一般流程】首先观察通项是否趋于零,是则继续,否则级数发散;然后观察通项是否变号,不变号则归于正项级数问题,否则归于变号级数问题;最后再分别判定.

【正项级数判敛散】如果通项中含有阶乘、n次幂,可以考虑比值法或根值法;如果通项是$\dfrac{1}{n}$的函数,可以考虑比阶法转化成已知的p-级数问题;还可以用比较判别法,进行适当的放大或缩小处理.

【变号级数判敛散】首先考虑对应的绝对值级数的敛散性,收敛时级数为绝对收敛,发散时则不能下结论. 如果是交错级数,便调用莱布尼兹判别法,若判得收敛,便结合绝对值级数的发散性,说此级数条件收敛. 如果不是交错级数,则可以考虑加括号判别法或者部分和方法.

【适合于部分和方法的问题】

(1) **如果要计算数项级数的和,便不允许有误差,此时多使用部分和法.**

例 1 $\sum\limits_{n=2}^{\infty}\ln\left(1-\frac{1}{n^2}\right)$

$=\lim\limits_{n\to\infty}\sum\limits_{k=2}^{n}[\ln(k+1)+\ln(k-1)-2\ln k]$

$=\lim\limits_{n\to\infty}[\ln(n+1)-\ln n+\ln 2]$

$=\ln 2.$

(2) **如果通项中出现相邻两项的差,则部分和比较容易化简. 例如,级数** $\sum\limits_{n=1}^{\infty}(x_{n+1}-x_n)$ **当数列** $\{x_n\}$ **收敛时收敛.**

【适合于分项法的问题】将通项分为几项之和,逐项考虑敛散性. 例如通项为

$$a_n=(-1)^n\frac{\sqrt{n}-(-1)^n}{n-1}$$

$$=(-1)^n\frac{\sqrt{n}}{n-1}-\frac{1}{n-1}$$

的级数,由于级数 $\sum(-1)^n\frac{\sqrt{n}}{n-1}$ **收敛,** $\sum\frac{1}{n-1}$ **发散,考虑到"收敛+发散=发散",故级数** $\sum a_n$ **发散.**

【通项中含有定积分的问题】直接求出定积分往往较难. 可先考虑对被积函数进行放大或缩小处理,再使用正项级数的比较判别法进行判定.

例 1 **判定** $\sum\int_0^{\frac{\pi}{n}}\frac{\sin x}{1+x}\mathrm{d}x$ **的敛散性.**

解 **被积函数为分式,缩小分母变形:**

$$0\leqslant \int_0^{\frac{\pi}{2}} \frac{\sin x}{1+x}\mathrm{d}x \leqslant \int_0^{\frac{\pi}{n}} \sin x\mathrm{d}x$$

$$= 1-\cos\frac{\pi}{n} \sim \frac{\pi^2}{2}\cdot\frac{1}{n^2}.$$

由于 $\sum\frac{1}{n^2}$ 收敛,故与之同阶的正项级数 $\sum\left(1-\cos\frac{\pi}{n}\right)$ 也收敛,从而通项较小的原级数也收敛.

例2 判定级数 $\sum\int_0^1 x^n\arctan x\mathrm{d}x$ 的敛散性.

解 直接积分与直接放大都不易进行,先做等式变形:

$$\int_0^1 x^n\arctan x\mathrm{d}x$$

$$=\int_0^1 \arctan x\mathrm{d}\frac{x^{n+1}}{n+1}$$

$$=\frac{x^{n+1}}{n+1}\arctan x\Big|_0^1-\int_0^1\frac{x^{n+1}}{1+x^2}\cdot\frac{1}{n+1}\mathrm{d}x$$

$$=\frac{\pi}{4}\cdot\frac{1}{n+1}-\frac{1}{n+1}\int_0^1\frac{x^{n+1}}{1+x^2}\mathrm{d}x.$$

于是所论级数的通项分为两项之和,一项是

$$a_n=\frac{\pi}{4}\cdot\frac{1}{n+1},$$

它对应的级数发散;另一项是

$$b_n=\frac{1}{n+1}\int_0^1\frac{x^{n+1}}{1+x^2}\mathrm{d}x.$$

由于

$$0<\frac{1}{n+1}\int_0^1\frac{x^{n+1}}{1+x^2}\mathrm{d}x<\frac{1}{n+1}\int_0^1 x^{n+1}\mathrm{d}x\sim\frac{1}{n^2},$$

故 b_n 对应的级数收敛,从而原级数发散.

7.2 幂 级 数

大纲要求

了解函数项级数的收敛域及和函数的概念.理解幂级数收敛半径的概念,并掌握幂级数的收敛半径、收敛区间及收敛域的求法.了解幂级数在其收敛区间内的一些基本性质(和函数的连续性、逐项微分和逐项积分),会求一些幂级数在收敛区间内的和函数,并会由此求出某些数项级数的和.了解函数展开为泰勒级数的充分必要条件.掌握 e^x,$\ln(1+x)$,$\sin x$,$\cos x$,$(1+x)^\alpha$ 的麦克劳林展开式,会用它们将一些简单函数间接展开成幂级数.

知识点·点拨

【幂级数】一般形式为 $\sum\limits_{n=0}^{\infty}a_n(x-x_0)^n$,标准形式为 $\sum\limits_{n=0}^{\infty}a_nx^n$.通过代换 $t=x-x_0$,可以将一般形式化作标准形式.

【收敛点与发散点】若 $x=a$ 时幂级数收敛,则称 a 为幂级数的收敛点,否则称 a 为发散点.

【收敛域】幂级数收敛点的全体构成的集合.

【收敛半径与收敛区间】若幂级数 $\sum\limits_{n=0}^{\infty}a_n(x-x_0)^n$ 的收敛域不是 $\{x_0\}$,也不是 $(-\infty,+\infty)$,则存在正数 R(称为收敛半径)和开区间 (x_0-R,x_0+R)(称为收敛区间),使得

(1)在收敛区间内幂级数是绝对收敛;

(2)在收敛区间的端点处,其敛散性不定,如果幂级数在点 $x=a$ 处是条件收敛,则该点一定是收敛区间的端点;

(3)在区间 $[x_0-R, x_0+R]$ 之外幂级数发散,如图7-1-1所示.

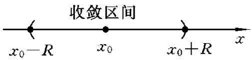


图 7-1-1

〖点拨〗收敛区间不一定与收敛域相同.

【收敛半径的计算】收敛半径 R 的计算公式为

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|=\lim_{n\to\infty}\frac{1}{\sqrt[n]{|a_n|}}.$$

〖点拨1〗对形为 $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ 和 $\sum_{n=0}^{\infty}a_nx^{2n}$ 的幂级数,应当作代换 $y=x-x_0$ 和 $y=x^2$,化作标准形式的幂级数后求得关于 y 的收敛范围,再还原成 x 的范围.

〖点拨2〗使用比值法讨论幂级数的绝对收敛范围也可以得到收敛半径.

【和函数】幂级数在收敛域上的和,记作 $S(x)$,其定义域便是收敛域.

【逐项求导公式】设 D 是收敛区间,则

$$\left(\sum_{n=0}^{\infty}a_nx^n\right)'=\sum_{n=1}^{\infty}na_nx^{n-1}\text{,其中 } x\in D.$$

【逐项求积公式】设 D 是收敛域,则

$$\int_a^b\left(\sum_{n=0}^{\infty}a_nx^n\right)\mathrm{d}x=\sum_{n=0}^{\infty}\int_a^b a_nx^n\mathrm{d}x,$$

其中,$[a,b]\subseteq D$.

【收敛半径的不变性】幂级数 $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ 经过求导或积分之后收敛半径不会改变.

【收敛端点的可变性】幂级数积分之后,发散的端点可能会收敛;而求导之后收敛的端点可能会发散.例如,$\sum_{n=1}^{\infty}(-1)^n\frac{x^n}{n}$在收敛区间的端点$x=1$处收敛,但是其导数$\sum_{n=1}^{\infty}(-1)^nx^{n-1}$在该点发散.

【泰勒级数】设函数 $f(x)$在包含 x_0 的某区间内无限次可微,则级数

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

称为 $f(x)$在点 x_0 的泰勒级数.

〖点拨〗泰勒级数不一定收敛.

【麦克劳林级数】即 $x_0=0$ 时的泰勒级数.

【幂级数展开定理】设函数 $f(x)$在包含 x_0 的某个邻域 I 内无限次可微,则 $f(x)$在该邻域内能展开成泰勒级数的充分必要条件是

$$\lim_{n\to\infty}\frac{f^{(n+1)}[x_0+\theta(x-x_0)]}{(n+1)!}(x-x_0)^{n+1}=0,$$

$$x\in I,\quad 0<\theta<1,$$

此条件满足时,在邻域 I 上成立

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

常见题型·应对

【收敛域概念问题】直接计算幂级数的收敛区间或收敛域.难点在于讨论端点的敛散性.记住,条件收敛只能在收敛区间的端点发生,收敛区间内部的点都是绝对收敛点.以下两个填空题比较典型.

例1 若幂级数 $\sum\limits_{n=1}^{\infty}a_n(x-1)^n$ 在 $x=-1$ 处收敛,则此级数在 $x=2$ 处绝对收敛.

例2 若幂级数 $\sum\limits_{n=1}^{\infty}a_n(x-x_0)^n$ 在点 $x=1,x=5$ 处都是条件收敛,则 $[1,5]$ 为收敛域,且 $x_0=3$.

【展开为幂级数·借助已有展开式和积分】

例
$$\begin{aligned}\ln(2+x)&=\int_0^x\frac{1}{2+t}\mathrm{d}t+\ln2\\&=\frac{1}{2}\int_0^x\frac{1}{1-(-t/2)}\mathrm{d}t+\ln2\\&=\frac{1}{2}\int_0^x\sum_{n=0}^{\infty}\left(\frac{-t}{2}\right)^n\mathrm{d}t+\ln2\\&=\frac{1}{2}\sum_{n=0}^{\infty}\int_0^x\left(\frac{-t}{2}\right)^n\mathrm{d}t+\ln2\\&=\sum_{n=0}^{\infty}(-1)^n\cdot\frac{x^{n+1}}{2^{n+1}(n+1)}+\ln2,\\&\qquad\qquad -2<x\leqslant2.\end{aligned}$$

【展开为幂级数·借助已有展开式和导数】

例
$$\begin{aligned}\frac{1}{(1-x)^2}&=\left(\frac{1}{1-x}\right)'=\left(\sum_{n=1}^{\infty}x^n\right)'\\&=\sum_{n=1}^{\infty}nx^{n-1},\quad |x|<1.\end{aligned}$$

【展开为幂级数 · 借助已有展开式和因式分解】

例　$$\frac{1}{2+x-x^2}=\frac{1}{3}\left(\frac{1}{1+x}+\frac{1}{2-x}\right)$$

$$=\frac{1}{3}\cdot\frac{1}{1+x}+\frac{1}{6}\cdot\frac{1}{1-(x/2)}$$

$$=\frac{1}{3}\sum_{n=0}^{\infty}(-1)^n x^n+\frac{1}{6}\sum_{n=0}^{\infty}\left(\frac{x}{2}\right)^n$$

$$=\sum_{n=0}^{\infty}\frac{1}{3}\left[(-1)^n+\frac{1}{2^{n+1}}\right]x^n,\quad |x|<1.$$

【在非零点的展开式 · 平移代换】通过代换 $t=x-x_0$**,可以将麦克劳林公式应用到非零点的展开问题中.**

例 1　**求** $\ln x$ **在** $x=3$ **的幂级数展开式.**

解　**代换** $t=x-3$(**使** $x=3$ **时**,$t=0$),**得**

$$\ln x=\ln(3+t)=\ln 3\left(1+\frac{t}{3}\right)=\ln 3+\ln\left(1+\frac{t}{3}\right)$$

$$=\ln 3+\sum_{n=0}^{\infty}(-1)^n\frac{1}{n+1}\left(\frac{t}{3}\right)^{n+1}\quad\left(\left|\frac{t}{3}\right|<1\right)$$

$$=\ln 3\sum_{n=1}^{\infty}\frac{(-1)^n}{n+1}\frac{1}{3^{n+1}}(x+3)^{n+1}\quad(|x-3|<3)$$

$$=\ln 3+\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n\cdot 3^n}(x-3)^n,\quad 0<x\leqslant 6.$$

例 2　**求** $\frac{1}{1+x}$ **在** $x=3$ **的幂级数展开式.**

解　**代换** $t=x-3$,**得**

$$\frac{1}{1+x}=\frac{1}{4+t}=\frac{1}{4}\cdot\frac{1}{1-(-t/4)}$$

$$=\frac{1}{4}\sum_{n=0}^{\infty}\frac{(-1)^n}{4^n}t^n\quad\left(\left|\frac{t}{4}\right|<1\right)$$

$$= \frac{1}{4}\sum_{n=0}^{\infty}\frac{(-1)^n}{4^n}(x-3)^n,\quad -1<x<7.$$

【幂级数求和 · 借助已有公式和导数】

例 **求** $\sum\limits_{n=1}^{\infty} nx^n$ **的和函数** $S(x)$.

解 $$S(x)=\sum_{n=1}^{\infty}x(x^n)'=x\left(\sum_{n=1}^{\infty}x^n\right)'$$

$$=x\left(\sum_{n=0}^{\infty}x^n-1\right)'=x\left(\frac{1}{1-x}-1\right)'$$

$$=\frac{x}{(1-x)^2},\quad |x|<1.$$

【幂级数求和 · 借助已有公式与积分】

例 **求** $\sum\limits_{n=1}^{\infty}\frac{x^{2n}}{n}$ **的和函数** $S(x)$.

解 $$S(x)=\sum_{n=1}^{\infty}\int_0^{x^2}t^{n-1}\mathrm{d}t=\int_0^{x^2}\left(\sum_{n=2}^{\infty}t^{n-1}\right)\mathrm{d}t$$

$$=\int_0^{x^2}\frac{1}{1-t}\mathrm{d}t=-\ln(1-t)\Big|_0^{x^2}$$

$$=-\ln(1-x^2),\quad -1<x<1.$$

【幂级数求和 · 系数拆分】

例 1 **求** $\sum\limits_{n=1}^{\infty}n^2x^2$ **的和函数** $S(x)$.

解 **作分解** $n^2=(n+1)n-n$**,则**

$$S(x)=x\sum_{n=1}^{\infty}(n+1)nx^{n-1}-x\sum_{n=1}^{\infty}nx^{n-1}$$

$$=x\left(\sum_{n=1}^{\infty}x^{n+1}\right)''-\left(x\sum_{n=1}^{\infty}x^n\right)'$$

$$= x\left(\frac{1}{1-x} - 1 - x\right)'' - x\left(\frac{1}{1-x} - 1\right)'$$

$$= \frac{x(1+x)}{(1-x)^3},\quad |x| < 1.$$

例 2　求 $\sum_{n=2}^{\infty} \frac{x^n}{n(n-1)}$ **的和函数** $S(x)$.

解　**作分解** $\frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}$ **,则**

$$S(x) = \sum_{n=2}^{\infty} \frac{x^n}{n-1} - \sum_{n=2}^{\infty} \frac{x^n}{n}$$

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

$$= (x-1)[-\ln(1-x)] + x.$$

【幂级数求和 · 微分方程】

例　**求** $S(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$.

解　$S(x) = 1 + \frac{1}{2}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \cdots,$

$$S'(x) = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots,$$

$$S''(x) = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots,$$

故 $S''(x) = S(x)$ **,且** $S(0) = 1, S'(0) = 0$ **,故得**

$$S(x) = \frac{1}{2}(e^x + e^{-x}).$$

注　**构建** $S'(x) + S(x) = e^x, S(0) = 1$ **,也可以求得和函数.**

【数项级数求和问题】通常是结合幂级数求和或者三角级数展开问题出现. 代入收敛域内适当的值便

可以由和函数中求得数项级数的和.

【常用幂级数展开式】

$$\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n,\qquad |x|<1;$$

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!},\qquad |x|<+\infty;$$

$$\sin x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!},\qquad |x|<+\infty;$$

$$\cos x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!},\qquad |x|<+\infty;$$

$$\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n},\quad -1<x\leqslant 1;$$

$$\arctan x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{2n+1},\quad -1\leqslant x\leqslant 1.$$

【常用幂级数求和公式】

$$\sum_{n=0}^{\infty}x^n=\frac{1}{1-x},\quad |x|<1;$$

$$\sum_{n=1}^{\infty}nx^n=\frac{x}{(1-x)^2},\quad |x|<1;$$

$$\sum_{n=1}^{\infty}\frac{x^n}{n}=-\ln(1-x),\quad -1\leqslant x<1;$$

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}=e^x,\quad -\infty<x<+\infty.$$

7.3 傅里叶级数

大纲要求

了解傅里叶级数的概念和狄利克雷收敛定理，会将定义在$[-l,l]$上的函数展开为傅里叶级数，会将定义在$[0,l]$上的函数展开为正弦级数与余弦级数，会写出傅里叶级数的和的表达式.

知识点·点拨

【三角级数】 以$T=2l$为周期的正弦函数和余弦函数构成的如下函数项级数：

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{l}+b_n\sin\frac{n\pi x}{l}\right).$$

【函数$f(x)$的傅里叶系数】 设$f(x)$在区间$[-l,l]$上可积. 以下系数称为$f(x)$的傅里叶系数：

$$a_n=\frac{1}{l}\int_{-l}^{l}f(x)\cos\frac{n\pi x}{l}\mathrm{d}x,\quad n\geqslant 0,$$

$$b_n=\frac{1}{l}\int_{-l}^{l}f(x)\sin\frac{n\pi x}{l}\mathrm{d}x,\quad n\geqslant 1.$$

【函数$f(x)$的傅里叶级数】 使用$f(x)$的傅里叶系数构成的三角级数：

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{l}+b_n\sin\frac{n\pi x}{l}\right).$$

〖点拨〗由傅里叶系数看出：奇函数的傅里叶级数中只出现正弦函数，偶函数的傅里叶级数中只出现余弦函数.

【狄利克雷收敛定理】 设$f(x)$在$[-l,l]$上①分段单

调，②分段连续，则其傅里叶级数处处收敛，且傅里叶级数的和函数$S(x)$以$T=2l$为周期. 在$[-l,l]$上有

$$S(x)=\begin{cases} f(x), & -l<x<l, f\text{ 在 } x \text{ 连续;} \\ \dfrac{f(x^-)+f(x^+)}{2}, & -l<x<l, f\text{ 在 } x \text{ 间断;} \\ \dfrac{f(-l^+)+f(l^-)}{2}, & x=\pm l. \end{cases}$$

【三种形式的傅里叶级数展开问题】

(1)**标准展开.** 将对称区间$[-l,l]$上函数$f(x)$展开成标准形式

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{l}+b_n\sin\frac{n\pi x}{l}\right).$$

(2)**正弦展开.** 将$[0,l]$上$f(x)$展开成正弦级数

$$\sum_{n=1}^{\infty}b_n\sin\frac{n\pi x}{l}.$$

(3)**余弦展开.** 将$[0,l]$上$f(x)$展开成余弦级数

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}a_n\cos\frac{n\pi x}{l}.$$

〖点拨1〗三种展开方式的系数计算公式相同. 其中，正弦展开是将$[0,l]$上函数$f(x)$奇延拓至$[-l,l]$上后再展开，从而

$$a_n=0,\quad b_n=\frac{2}{l}\int_0^l f(x)\sin\frac{n\pi x}{l}\mathrm{d}x,$$

余弦展开是将$[0,l]$上函数$f(x)$偶延拓至$[-l,l]$上后再展开，从而

$$b_n=0,\quad a_n=\frac{2}{l}\int_0^l f(x)\cos\frac{n\pi x}{l}\mathrm{d}x.$$

〖点拨2〗三种展开方式的和函数$S(x)$一定要

在对称区间$[-l,l]$上对被展函数 $f(x)$或延拓函数依狄利克雷收敛定理建立(借助几何作图).

常见题型 · 应对

【考察收敛定理】要求写出和函数$S(x)$在$[-l,l]$内或者(利用$S(x)$以 $2l$ 为周期的周期性)在$[-l,l]$外某点的函数值.

〖点拨〗根据条件弄清是哪一种展开方式,在对称区间上依据狄利克雷收敛定理解决.

例 1　设函数 $f(x)=\begin{cases}2, & -1<x\leqslant 0,\\ x^2, & 0<x\leqslant 1,\end{cases}$ 其图形如图 7-3-1 所示. $S(x)$是 $f(x)$在$[-1,1]$上的傅里叶级数 $\dfrac{a_0}{2}+\sum\limits_{n=1}^{\infty}(a_n\cos n\pi x+b_n\sin n\pi x)$的和函数,则 $S(0)=1$,$S(1)=S(-1)=\dfrac{3}{2}$,$S\left(\dfrac{5}{2}\right)=S\left(\dfrac{1}{2}\right)=\dfrac{1}{4}$(利用和函数的周期性).

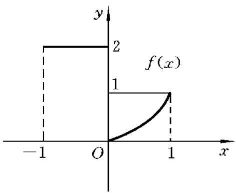


图 7-3-1

例 2　设 $f(x)=x^2$,$0<x<1$,其图形如图7-3-2

所示.

$$b_n=\frac{1}{2}\int_0^1 f(x)\sin n\pi x\mathrm{d}x,\quad n\geqslant 1,$$

$$S(x)=\sum_{n=1}^{\infty}b_n\sin n\pi x,$$

则可推知 $S(x)$ 是 $f(x)$ 在 $(0,1)$ 内的正弦级数. 从而

$$S\left(-\frac{1}{2}\right)=-\frac{1}{4},\quad S(0)=S(1)=0.$$

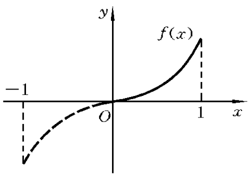


图 7-3-2

【单个系数计算】要求计算某个 a_n 或 b_n.

例　设 $f(x)=\pi x+x^2(-\pi\leqslant x\leqslant\pi)$ 的展开式为 $\dfrac{a_0}{2}+\sum\limits_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$，求 b_3.

解　$b_3=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(x)\sin 3x\mathrm{d}x$

$$=\frac{1}{\pi}\int_{-\pi}^{\pi}\pi x\sin 3x\mathrm{d}x+0=\frac{2\pi}{3}.$$

【展开为傅里叶级数】

例　将 $f(x)=\begin{cases}0, & -\pi\leqslant x<0,\\ x, & 0<x\leqslant\pi\end{cases}$ 在 $[-\pi,\pi]$ 上展开为傅里叶级数.

解 $a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\mathrm{d}x$

$$=\int_0^{\pi}x\cos nx\mathrm{d}x=\frac{(-1)^n-1}{n^2\pi},$$

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\mathrm{d}x=\frac{1}{\pi}x\mathrm{d}x=\frac{\pi}{2},$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\mathrm{d}x$$

$$=\frac{1}{\pi}\int_0^{\pi}x\sin nx\mathrm{d}x=\frac{(-1)^{n-1}}{n},$$

和函数为 $S(x)=\begin{cases}f(x), & |x|<\pi,\\ \pi/2, & x=\pm\pi.\end{cases}$

故所求展开式为

$$f(x)=\frac{\pi}{4}+\sum_{n=1}^{\infty}\left[\frac{(-1)^n-1}{n^2\pi}\cos nx+\frac{(-1)^{n-1}}{n}\sin nx\right]\quad(|x|<\pi).$$

【展开为正弦级数】

例 **将**$[0,\pi]$**上函数** $f(x)=x+1$ **展开为仅有正弦函数的傅里叶级数(称为正弦级数)**.

解 **构造**$[-\pi,\pi]$**上的奇函数** $F(x)$**,使它在**$(0,\pi]$**上与** $f(x)$**一致.再求**$[-\pi,\pi]$**上** $F(x)$**的傅里叶级数,由于** $a_n=0$**,故得到的三角级数中无余弦成分.写出其和函数** $S(x)$**后,截取在**$[0,\pi]$**上的结果,亦即**

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}F(x)\sin nx\mathrm{d}x=\frac{2}{\pi}\int_0^{\pi}F(x)\sin x\mathrm{d}x$$

$$=\frac{2}{\pi}\int_0^{\pi}f(x)\sin nx\mathrm{d}x=\frac{2}{\pi}\int_0^{\pi}(x+1)\sin nx\mathrm{d}x$$

$$=\frac{2}{n\pi}[1-(\pi+1)(-1)^n],\quad n\geqslant1.$$

在$[0,\pi]$上的和函数为

$$S(x)=\begin{cases}x+1, & 0<x<\pi,\\ 0, & x=0,\pi.\end{cases}$$

故所求的展开式为

$$f(x)=\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1-(\pi+1)(-1)^n}{n}\sin nx \quad (0<x<\pi)$$

【展开为余弦级数】

例　将$[0,\pi]$上函数$f(x)=x+1$展开为仅有余弦函数的傅里叶级数(称为余弦级数).

解　构造$[-\pi,\pi]$上的偶函数$F(x)$,使它在$(0,\pi]$上与$f(x)$一致.再求$[-\pi,\pi]$上$F(x)$的傅里叶级数,由于$b_n=0$,故得到的级数是余弦级数,写出其和函数后,截取在$[0,\pi]$上的结果,亦即

$$\begin{aligned}a_n&=\frac{1}{\pi}\int_{-\pi}^{\pi}F(x)\cos nx\mathrm{d}x=\frac{2}{\pi}\int_{0}^{\pi}F(x)\cos nx\mathrm{d}x\\&=\frac{2}{\pi}\int_{0}^{\pi}f(x)\cos nx\mathrm{d}x=\frac{2}{\pi}\int_{0}^{\pi}(x+1)\cos nx\mathrm{d}x\\&=\frac{2[(-1)^n-1]}{n^2\pi},\quad n\geqslant1;\end{aligned}$$

$$a_0=\frac{2}{\pi}\int_0^{\pi}(x+1)\mathrm{d}x=\pi+2.$$

在$[0,\pi]$上的和函数为$S(x)=f(x)$.

故所求展开式为

$$f(x)=1+\frac{\pi}{2}-\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\cos(2n-1)x \quad (0\leqslant x\leqslant\pi).$$

第 8 章　常微分方程

大纲要求

了解微分方程及其解、阶、通解、初始条件和特解等概念.掌握变量可分离的微分方程及一阶线性微分方程的解法.会解齐次微分方程、贝努利方程和全微分方程,会用简单的变量代换解某些微分方程.会用降阶法解下列微分方程:

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

理解线性微分方程解的性质及解的结构.掌握二阶常系数齐次线性微分方程的解法,会解某些高于二阶的常系数齐次线性微分方程.会解自由项为多项式、指数函数、正弦函数、余弦函数,以及它们的和与积的二阶常系数非齐次线性微分方程.会解欧拉方程.会用微分方程解决一些简单的应用问题.

注　数学二不要求贝努利方程、全微分方程.

8.1　一阶微分方程

知识点·点拨

【一阶微分方程】　含有未知函数 $y(x)$ 的导数 $y'(x)$ 的等式,如

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

或
$$a(x,y)\mathrm{d}x+b(x,y)\mathrm{d}y=0.$$

常见的一阶微分方程有:可分离变量方程、齐次方程、线性方程、贝努利方程、全微分方程.

【一阶微分方程的通解】满足所给方程的，含有一个任意常数的可微函数：

$$y = y(x, C).$$

【一阶微分方程的定解】将定解条件 $y(x_0)=y_0$ 代入通解后可求出任意常数的值，从而得出的解.

【一阶可分离方程及其求解】形式为

$$y' = f(x)g(y).$$

〖求通解〗分离变量使方程为 $a(x)\mathrm{d}x=b(y)\mathrm{d}y$，作不定积分

$$\int a(x)\mathrm{d}x = \int b(y)\mathrm{d}y,$$

即得通解.

〖求定解〗若定解条件为 $y(x_0)=y_0$，则分离变量后作变限积分

$$\int_{x_0}^{x} a(x)\mathrm{d}x = \int_{y_0}^{y} b(y)\mathrm{d}y,$$

即得定解.

【一阶线性微分方程及其求解】标准形式为

$$y' + a(x)y = b(x).$$

〖求通解〗取 $A(x)$，使 $A'(x)=a(x)$，则通解为

$$y(x) = \mathrm{e}^{-A(x)}\left[\int b(x)\mathrm{e}^{A(x)}\mathrm{d}x + C\right]$$

或

$$y(x) = \mathrm{e}^{-A(x)}\left[\int_{x_0}^{x} b(t)\mathrm{e}^{A(t)}\mathrm{d}t + C\right].$$

〖定解〗若 $y(x_0)=y_0$，则定解可写为

$$y = \mathrm{e}^{-A(x)}\left[\int_{x_0}^{x} b(t)\mathrm{e}^{A(t)}\mathrm{d}t + y_0\mathrm{e}^{A(x_0)}\right].$$

〖点拨 1〗如果以 y 为因变量时方程不是线性的，换作 x 为因变量便有可能是线性的. 例如，微分方程

$$y'(x+\mathrm{e}^y)=1,$$

交换 x,y 的地位后，便是函数 $x(y)$ 的线性微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}y}-x=\mathrm{e}^y.$$

〖点拨 2〗若所给方程既是线性方程，又是其他类型时，应当首选线性方程的通解公式法.

〖点拨 3〗在需要进一步讨论解函数的极限、不等式时，如果通解公式中积分不能求出，则使用变上限积分表示通解.

例　求方程 $y'+ay=f(x)$，$y(0)=0$ 的解，当 $|f(x)|\leqslant k\ (x>0)$ 时，估计解函数的大小.

解　取 $A(x)=ax$，则定解为

$$\begin{aligned}y&=\mathrm{e}^{-ax}\left[\int_0^x f(x)\mathrm{e}^{ax}\mathrm{d}x+0\right]\\&=\mathrm{e}^{-ax}\int_0^x f(x)\mathrm{e}^{ax}\mathrm{d}x,\end{aligned}$$

由于 $|f(x)|\leqslant k(x>0)$，故有

$$\begin{aligned}|y(x)|&\leqslant \mathrm{e}^{-ax}\int_0^x |f(x)|\mathrm{e}^{ax}\mathrm{d}x\\&\leqslant k\mathrm{e}^{-ax}\int_0^x \mathrm{e}^{ax}\mathrm{d}x\\&=\frac{k}{a}\mathrm{e}^{-ax}(\mathrm{e}^{ax}-1)=\frac{k}{a}(1-\mathrm{e}^{-ax}).\end{aligned}$$

【齐次方程及其求解】形式为　$y'=f\left(\dfrac{y}{x}\right)$.

〖求解〗引入新的函数变量 $u=\dfrac{y}{x}$ 后，变成可分离方程

$$\frac{\mathrm{d}u}{f(u)-u}=\frac{\mathrm{d}x}{x}.$$

【贝努利方程及其求解】标准形式为

$$y'+a(x)y=b(x)y^n\quad(n\neq 0,1).$$

〖求解〗引入新的函数变量 $u=y^{1-n}$ 后，可变成线性微分方程

$$\frac{1}{1-n}u'+a(x)u=b(x).$$

【全微分方程及其求解】标准形式为

$$p(x,y)\mathrm{d}x+q(x,y)\mathrm{d}y=0,$$

其中 $$q_x(x,y)=p_y(x,y).$$

〖解法1〗通过凑微分，求得函数 $u(x,y)$，使得

$$\mathrm{d}u(x,y)=p(x,y)\mathrm{d}x+q(x,y)\mathrm{d}y,$$

通解为 $u(x,y)=C$.

〖解法2〗通过变上限线积分，求得函数

$$u(x,y)=\int_{(x_0,y_0)}^{(x,y)}p(x,y)\mathrm{d}x+q(x,y)\mathrm{d}y,$$

通解为 $$u(x,y)=C.$$

【积分方程及其求解】含有未知函数的变限积分的方程称为积分方程.

〖求解要点〗(1)通过求导来消除积分，转换为微分方程后求解；(2)求导数之前尽可能将积分孤立，以便求导后能够消去积分；(3)取 x 为特殊的值，试着找初始条件.

例 1　对于积分方程 $f(x)+2\int_0^x f(x)\mathrm{d}x=x^2$，首先找到初始条件 $f(0)=0$，再对积分方程求导，得线性微分方程

$$f'(x)+2f(x)=2x,$$

从中可以求得定解　$f(x)=x-\frac{1}{2}+\frac{1}{2}\mathrm{e}^{-2x}$.

例 2　对于积分方程

$$f(x)=\mathrm{e}^x+\mathrm{e}^x\int_0^x f^2(x)\mathrm{d}x,$$

两边乘以 e^{-x}，然后求导

$$[\mathrm{e}^{-x}f(x)]'=\left[1+\int_0^x f^2(x)\mathrm{d}x\right]',$$

所得方程是贝努利方程

$$f'(x)-f(x)=\mathrm{e}^x f^2(x),\quad f(0)=1,$$

可以求得定解为　$f(x)=2\mathrm{e}^x/(3-\mathrm{e}^{2x})$.

8.2　二阶可降阶微分方程

【二阶微分方程】含有未知函数 $y(x)$ 的二阶导数 $y''(x)$ 的等式，如

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

常见的二阶微分方程有：常系数线性方程、可降阶方程和欧拉方程.

【二阶微分方程的通解】满足所给方程的含有两个任意常数的可微函数，即

$$y=y(x,C_1,C_2).$$

【定解条件】二阶方程的定解条件为

$$y(x_0)=y_0,\quad y'(x_0)=y_1.$$

【可降阶方程之一:缺因变量及其一阶导数】方程形式:

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

〖求解〗直接积分两次即得通解.

例　对于 $y''=\cos x$,积分两次,得

$$y'=\sin x+C_1,\quad y=-\cos x+C_1x+C_2.$$

【可降阶方程之二:缺因变量】方程形式:

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

〖求解〗令 $p=y'$,便化做未知函数 $p(x)$ 的一阶微分方程

$$p'=f(x,p).$$

例　对于 $xy''+y'=4x$,令 $p=y'$,化做 $p'+p/x=4$,求得 $p=y'=2x+C_1/x$,再积分得通解为

$$y=x^2+C_1\ln x+C_2.$$

【可降阶方程之三:缺自变量】方程形式:

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

〖求解〗令 $p=y'$,便化做未知函数 $p(y)$ 的一阶微分方程 $pp'=f(y,p)$.

例　对于 $y''+y'^2=y$, $y(0)=3/2$, $y'(0)=1$,令 $p=y'$,化做定解问题

$$pp'+p^2=y,\quad p(3/2)=1,$$

可看做是 p^2 的线性方程,求解得 $p=\sqrt{y-1/2}$,再解定解问题

$$y'=\sqrt{y-1/2},\quad y(0)=3/2,$$

即得所求问题的解

$$y=\frac{1}{4}(x+2)^2+\frac{1}{2}.$$

8.3 线性微分方程

【二阶线性微分方程】

齐次： $y''+a(x)y'+b(x)y=0$；

非齐次： $y''+a(x)y'+b(x)y=f(x)$.

【基本解 · 基本解组】二阶齐次线性微分方程

$$y''+a(x)y'+b(x)y=0$$

的两个不成比例的解函数称为该方程的基本解,它们构成方程的基本解组.

【二阶齐次线性微分方程通解的结构】通解是两个基本解 y_1, y_2 **的线性组合**：

$$y_{齐}=C_1y_1+C_2y_2,$$

其中, C_1, C_2 **是任意常数.**

【二阶非齐次线性微分方程通解的结构】通解是相应的齐次线性方程的通解与非齐次线性方程的任何一个解函数的组合：

$$y_{非齐}=C_1y_1+C_2y_2+y^*,$$

其中, $y_i(x)(i=1,2)$ **是齐次线性方程的基本解**, y^* **是非齐次线性方程的特解.**

【解的叠加原理】

〖**非齐次**〗**设函数** $y_i(x)(i=1,2)$ **是方程**

$$y''+a(x)y'+b(x)y=f_i(x),\quad i=1,2$$

的解, p,q **为常数,则** $py_1(x)+qy_2(x)$ **是方程**

$$y''+a(x)y'+b(x)y=pf_1(x)+qf_2(x)$$

的解.

〖**齐次**〗**设函数** $y_i(x)(i=1,2)$ **是齐次方程**

$$y'' + a(x)y' + b(x)y = 0$$

的解,则 $py_1(x)+qy_2(x)$还是同一方程的解.

〖非齐次方程的解与对应的齐次方程的解〗

非齐次线性方程的解+齐次线性方程的解

=非齐次线性方程的解;

非齐次线性方程的解−非齐次线性方程的解

=齐次线性方程的解.

【特征方程·特征根】常系数线性微分方程

$$y'' + ay' + by = f(x)$$

的特征方程为 $\lambda^2+a\lambda+b=0$,它的根称为特征值或特征根.

【二阶齐次线性微分方程的基本解求法】

特征根	基本解组
互异实根 λ_1,λ_2	$e^{\lambda_1 x},e^{\lambda_2 x}$
相等实根 λ,λ	$e^{\lambda x},xe^{\lambda x}$
共轭复根 $\alpha\pm i\beta$	$e^{\alpha x}\cos\beta x,e^{\alpha x}\sin\beta x$

〖点拨 1〗若 $\lambda=a$ 是特征根,则函数 $y=e^{ax}$便是齐次微分方程的解函数.

〖点拨 2〗三阶以上的方程考的较少,其特征根和基本解组可以类似地写出.

【二阶非齐次线性微分方程的特解求法】根据非齐次项 $f(x)$的特点,写出特解 y^* 的待定形式并代入到微分方程中,比较同次项的系数,列出方程求出待定系数即得.

【写 y^* 待定形式的六字诀】

同类型:y^* 与 $f(x)$ 是同类型的函数,如下表所示.

$f(x)$ 的类型	y^* 的形式
$2-3x+x^2$	$A+Bx+Cx^2$
$x\mathrm{e}^{3x}$	$(A+Bx)\mathrm{e}^{3x}$
$\sin 3x$	$A\sin 3x+B\cos 3x$
$x\cos 3x$	$(A+Bx)\sin 3x+(C+Dx)\cos 3x$
$\mathrm{e}^{2x}\cos 3x$	$\mathrm{e}^{2x}(A\sin 3x+B\cos 3x)$

后调整:如果 y^* 包含了对应的齐次方程的基本解,则需要通过乘以 x 来避开基本解,如果没有成功便再乘以 x.

经过这两步之后,便得到待定的二阶非齐次线性微分方程的特解待定形式.

〖点拨 1〗y^* 中包含的多项式的系数是待定的,y^* 中正弦函数部分与余弦函数部分是成对出现的,缺一不可.

〖点拨 2〗自由项函数 $f(x)$ 必须是多项式、指数函数、正弦函数、余弦函数,或者是它们的和与积.若 $f(x)$ 是以上形式的函数 $f_1(x)$ 与 $f_2(x)$ 的和,则可以针对 $f_i(x)$ 写出 y_i^* $(i=1,2)$,再以 $y^*=y_1^*+y_2^*$ 作为特解的待定形式.

【二阶欧拉方程】以下形式的二阶变系数线性微分方程

$$ax^2y''+bxy'+cy=f(x)$$

称为欧拉方程.

〖求解要点〗通过自变量代换 $x=\mathrm{e}^t$ 使得

$$xy' = \dot{y}, \quad x^2y'' = \ddot{y} - \dot{y},$$

函数 $y(x)$ 的变系数线性微分方程便可以化作函数 $y(t)$ 的常系数线性微分方程（关于 t 的导数记作 $\dot{y}$, $\dot{y}$）

$$a\ddot{y} + (b-a)\dot{y} + cy = f(\mathrm{e}^t).$$

例 1　$x^2y''+4xy'+2y=0 \quad (x>0)$.

解　令 $x=\mathrm{e}^t$，得到齐次常系数线性微分方程

$$\dot{y} + (4-1)\dot{y} + 2y = 0,$$

使用特征方程，可以求得通解为

$$y = C_1\mathrm{e}^{-t} + C_2\mathrm{e}^{-2t} = C_1\frac{1}{x} + C_2\frac{1}{x^2}.$$

例 2　$x^2y''+2y'=12\ln x$.

解　令 $x=\mathrm{e}^t$，得到非齐次常系数线性微分方程

$$\dot{y} + \dot{y} = 12t,$$

由 $\lambda^2+\lambda=0$，求得特征根为

$$\lambda_1=0, \quad \lambda_2=-1,$$

设 $y^*=(At+B)t$，代入方程后得

$$2A + B + 2At = 12t,$$

故 $A=6$, $B=-12$，方程通解为

$$\begin{aligned} y &= C_1 + C_2\mathrm{e}^{-t} + (6t-12)t \\ &= C_1 + C_2\frac{1}{x} + 6(\ln x)^2 - 12\ln x. \end{aligned}$$

常见题型 · 应对

【变量互换问题】

例　将 $x''+(y+\sin x)(x')^3=0$ 换作以 y 为因变

量的微分方程.并求在条件 $y(0)=0, y'(0)=\frac{3}{2}$ 下方程的解.

解　将 x', x'' 分别用 y', y'' 表示：$x'=\frac{1}{y'}$，$x''=\frac{-y''}{y'^3}$（参见 2.1 节），回代后得线性微分方程

$$y''-y=\sin x,$$

通解为 $$y=C_1e^x+C_2e^{-x}-\frac{1}{2}\sin x,$$

定解为 $$y=e^x-e^{-x}-\frac{1}{2}\sin x.$$

【由解函数构造二阶线性微分方程问题】

例 1　已知二阶非齐次线性微分方程的三个解为 $y_1=e^{-x}, y_2=2e^{-2x}+e^{-x}, y_3=3e^{-3x}+e^{-x}$，求此二阶非齐次线性微分方程及其通解.

解　应用叠加原理，求出齐次方程的基本解组，再设法推算出特征值 λ_i，得到特征方程后便可写出齐次线性微分方程，再代入一个解函数，便可以定出非齐次项 $f(x)$.

由 $y_2-y_1=2e^{-2x}$ 得齐次线性微分方程的解函数及特征值：$e^{-2x}, \lambda=-2$. 由 $y_3-y_1=3e^{-3x}$ 得齐次线性微分方程的另一个解函数及特征值：$e^{-3x}, \lambda=-3$. 于是，特征方程为

$$(\lambda+2)(\lambda+3)=\lambda^2+5\lambda+6=0,$$

所求微分方程为

$$y''+5y'+6y=f(x),$$

最后代入 y_1，得出非齐次项为

$$f(x)=e^{-x}-5e^{-x}+6e^{-x}=2e^{-x}.$$

依据通解结构知，所求通解为

$$y=C_1e^{-2x}+C_2e^{-3x}+e^{-x}.$$

例 2　已知 $y_1=x, y_2=1/x$ 是某二阶齐次线性微分方程的解函数，建立此微分方程.

解　$y=C_1x+C_2/x$，构成通解，故本题相当于由通解求方程. 通解两边对 x 求导两次，即

$$y'=C_1-C_2/x^2,\quad y''=2C_2/x^3.$$

消去 C_1, C_2 便得到所求的二阶线性(变系数)微分方程

$$x^2y''+xy'-y=0.$$

【微分方程的几何应用】将导数或变限积分应用到题给的几何关系中便可得到未知函数的方程.

〖点拨 1〗说曲线 $y=y(x)$ 在点 $(1,2)$ 与直线 $y=3x-1$ 相切，意味着 $y(1)=2, y'(1)=3$.

〖点拨 2〗过点 (x,y) 的平面曲线 $y=y(x)$ 的切线方程 L 为

$$Y-y=y'(x)(X-x).$$

例 1　求下凸曲线 $y=y(x)$，使它在点 $(1,1)$ 有水平切线且在任意一点的曲率 $\kappa=\sqrt{1+(y')^2}/2y^2$.

解　曲线下凸 $\Leftrightarrow y''\geqslant 0$；在点 $(1,1)$ 有水平切线 $\Leftrightarrow y(1)=1, y'(1)=0$. 从条件知，曲率

$$\kappa=\frac{|y''|}{(1+y'^2)^{3/2}}=\frac{\sqrt{1+y'^2}}{2y^2},$$

于是微分方程化作

$$2y^2y''=(1+y'^2)^2,$$

属于缺自变量 x 的可降阶方程. 使用降阶法，可以求得解函数为

$$y=1+\frac{1}{4}(x-1)^2.$$

例2　设曲线 $y=f(x)$ 在任意点 $(x,y)(x>0)$ 的切线在 y 轴上的截距为 $f(x)$ 在区间 $[0,x]$ 上的平均值，建立该函数的微分方程.

解　曲线的切线在 y 轴上的截距为 $y-xy'$，未知函数在区间 $[0,x]$ 上的平均值为 $\frac{1}{x}\int_0^x y(t)\mathrm{d}t$，故所求方程为

$$y-xy'=\frac{1}{x}\int_0^x f(t)\mathrm{d}t$$

或

$$y'+xy''=0.$$

【力学问题中的微分方程】依据牛顿第二定律 $F=ma$ 建立的微分方程有以下三种：

求函数 $v(t)$：　$F=mv'(t)$；

求函数 $s(t)$：　$F=ms''(t)$；

求函数 $v(s)$：　$F=mv\dfrac{\mathrm{d}v}{\mathrm{d}s}$.

已经考过的力学问题有：

下滑问题、浮力-下降问题、振动问题、阻力运动问题.

【依据告知的定律应用微分方程】题目会将所用的定律告知，要做的是写成微分方程及求解. 常见的定律如下：

〖冷却定律〗　$\dfrac{\mathrm{d}T}{\mathrm{d}t}=k(T_0-T)$，　$k>0$.

〖衰减定律〗　$\frac{\mathrm{d}m}{\mathrm{d}t}=-km,\quad k>0.$

〖增长定律〗　$\frac{\mathrm{d}m}{\mathrm{d}t}=km,\quad k>0.$

〖传播定律〗　$\frac{\mathrm{d}N}{\mathrm{d}t}=kN(L-N),\quad k>0.$

【依据微元法建立微分方程】在时间微区间或空间微区间上建立变量微元之间的恒等式,便得到所求的微分方程.

例　某湖储水量为 V,每年流入含 A 的污水量是 $V/6$,流入不含 A 的污水量是 $V/6$,流出的水量为 $V/3$. 已知 1999 年湖中含 A 为 $5m_0$,超过了国家规定标准. 从 2000 年初起,国家规定排入的含 A 的污水浓度不得高于 m_0/V. 问:至多经过多少年,湖中 A 的含量降至 m_0 以内?

解　设在时刻 t 湖中含 A 的量为 $m(t)$,则在时间变量微元段 $[t,t+\mathrm{d}t]$,湖水含 A 的浓度近似为 $m(t)/V$,于是考虑到湖水的流动,得出微元 $\mathrm{d}m$ 的均衡关系式

$$\begin{aligned}\mathrm{d}m(t)&=\frac{m_0}{V}\frac{V}{6}\mathrm{d}t-\frac{m}{V}\frac{V}{3}\mathrm{d}t\\&=\left(\frac{m_0}{6}-\frac{m}{3}\right)\mathrm{d}t,\end{aligned}$$

$$m(0)=5m_0,$$

减少量 $=$ 进入量 $-$ 排出量.

解得　$$m=\frac{m_0}{2}\left(1+9\mathrm{e}^{-\frac{t}{3}}\right),$$

令 $m(T)=m_0$,得 $T=6\ln 3\approx 6.5$(年),即至多经过 $6\ln 3$ 年,湖中 A 的含量可降至 m_0 以内.

第2篇　线性代数

第9章　行　列　式

大纲要求

了解行列式的概念,掌握行列式的基本性质.会用行列式的性质和按行(列)展开定理计算行列式.

知识点·点拨

【二阶行列式】称

$$D=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$$

为二阶行列式.

【三阶行列式】称

$$\begin{aligned} D &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32} \end{aligned}$$

为三阶行列式.

【对角线】从左上角到右下角的对角线叫做主对角线,从右上角到左下角的对角线叫做副对角线.

【排列】自然数$1,2,\cdots,n$按一定次序排成一排,称为

一个 n 元排列，记为 $i_1 i_2 \cdots i_n$. $1\ 2\ \cdots\ n$ 称为自然排列. n 元排列总共有 $n!$ 个.

例如，自然数 1,2,3 共有 $3!=6$ 个排列，我们用 $i_1 i_2 i_3$ 表示这 6 个排列中的某一个.

【逆序 · 逆序数】 在一个 n 元排列 $i_1 i_2 \cdots i_n$ 中，若一个大的数排在一个小的数的前面，则称这两个数构成一个逆序. 一个排列中逆序个数的总和就称为这个排列的逆序数，记为 $\tau[i_1 i_2 \cdots i_n]$.

例　依次从后向前，数出前面比其大的数字的个数，相加之后便是逆序数. 于是

$$\tau[3\ 2\ 4\ 1]=3+0+1+0=4,$$

$$\tau[1\ 3\ 5\ 2\ 4]=1+2+0+0+0=3,$$

$$\begin{aligned}&\tau[n\ n-1\ \cdots\ 2\ 1]\\&=(n-1)+(n-2)+\cdots+2+1+0\\&=\frac{n(n-1)}{2}.\end{aligned}$$

【奇排列 · 偶排列】 若排列的逆序数为奇(偶)数，则称此排列为奇(偶)排列.

【对换 · 邻换】 一个排列中的某两个数的位置互换，其余的数不动，就得到一个新的排列，称这样的变换为一次对换，其中相邻两个数的对换称为邻换.

【对换性质】 一次对换改变排列的奇偶性. 此性质有如下推论：

(1) 任意一个 n 元排列都可以经过一定次数的对换变成自然排列，并且所作对换的次数与该排列的逆序数有相同的奇偶性.

(2) 在全体 n 元排列的集合中，奇排列与偶排列

各占一半.

【n 阶行列式】把 n^2 个数 $a_{ij}(i=1,2,\cdots,n;j=1,2,\cdots,n)$排成$n$ 行 n 列,按照下式

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

$$=\sum_{[i_1 i_2 \cdots i_n]}(-1)^{\tau[i_1 i_2 \cdots i_n]}a_{1i_1}a_{2i_2}\cdots a_{ni_n}$$

计算得到的一个数,称为 n 阶行列式,简记为 $D=\det(a_{ij})$或$D=|(a_{ij})|$,其中 $\sum\limits_{[i_1 i_2 \cdots i_n]}$ 表示对所有n 元排列求和.

〖点拨〗定义式右边的每一项 $a_{1i_1}a_{2i_2}\cdots a_{ni_n}$ 中的每 一个元取自 D 中不同行不同列. 当行下标按自然顺序排列时,相应的列下标是$1\ 2\ \cdots\ n$ 的一个n 元排列 $i_1 i_2 \cdots i_n$,若它是偶排列,则该排列对应的项取正号;若是奇排列,则取负号,用$(-1)^{\tau[i_1 i_2 \cdots i_n]}$表示.

例　下三角行列式

$$D=\begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=a_{11}a_{22}\cdots a_{nn}.$$

上三角行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{vmatrix}=a_{11}a_{22}\cdots a_{nn}.$$

对角形行列式

$$\begin{vmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{vmatrix} = a_1 a_2 \cdots a_n.$$

$$\begin{vmatrix} & & & a_1 \\ & & a_2 & \\ & \iddots & & \\ a_n & & & \end{vmatrix} = (-1)^{1+2+\cdots+n} a_1 a_2 \cdots a_n.$$

【行列式的转置】如下式所示,把行列式

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

的行与列互换,得到新的行列式,记为

$$D^{\mathrm{T}} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix},$$

称 D^{T} 为 D 的转置行列式. 显然

$$(D^{\mathrm{T}})^{\mathrm{T}} = D.$$

〖点拨〗行列式与转置行列式相等.

【行列式的基本性质】以下对行列式的变形称为初等变换.

〖性质 1〗用一个数 k 乘行列式,等于将行列式的某一行(列)元素都乘以 k,即

$$k\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

也可以说,若行列式某行(列)有公因子 k,则可以把它提到行列式外面.

〖性质 2〗**若对换行列式的任意两行(列),则行列式变号.**

〖性质 3〗**把行列式的第 j 行(列)元素的 k 倍加到第 i 行(列)的对应元素上,行列式的值不变.**

【拆项性质】若行列式的第 i 行(列)的每一个元素都可表示为两数之和,则该行列式可表示为两个行列式之和,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

【代数余子式· 余子式】在 n 阶行列式 D 中,划去元素 a_{ij} 所在的第 i 行和第 j 列的元素,剩余的元素按原

次序构成一个 $n-1$ 阶行列式，称为 a_{ij} 的余子式，记为 M_{ij}. 称 $(-1)^{i+j}M_{ij}$ 为 a_{ij} 的代数余子式，记为 A_{ij}.

例如，对行列式 $\begin{vmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 0 & 2 & 1 & 0 \end{vmatrix}$，元素 $a_{32}=3$ 的余子式和代数余子式分别为

$$M_{32}=\begin{vmatrix} 1 & 0 & 4 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{vmatrix}=-1,$$

$$A_{32}=(-1)^{3+2}M_{32}=1.$$

〖代数余子式重要公式〗

$$\sum_{t=1}^{n} a_{it}A_{jt}=\begin{cases} D, & i=j, \\ 0, & i\neq j. \end{cases}$$

【按行（列）展开定理】行列式等于任一行（或列）的每个元素与其对应的代数余子式乘积的和，即

$$D=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}$$

（或 $D=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}$.）

〖点拨〗沿着零元素较多的行或列展开行列式，可以减少计算量. 因此我们总是先运用行列式的性质，将某一行（列）元素尽可能多地化为零，然后再按该行（列）展开.

例 1 计算行列式

$$D=\begin{vmatrix} -3 & 1 & 2 & 5 \\ 0 & 5 & 3 & 2 \\ 5 & -3 & 2 & 4 \\ 1 & -1 & 0 & 4 \end{vmatrix}.$$

解　先把第1行与第4行对换,再利用基本性质3把第1列中除第1个元素外的其他元素都化为零.

$$D=-\begin{vmatrix} 1 & -1 & 0 & 4 \\ 0 & 5 & 3 & 2 \\ 5 & -3 & 2 & 4 \\ -3 & 1 & 2 & 5 \end{vmatrix}$$

$$=-\begin{vmatrix} 1 & -1 & 0 & 4 \\ 0 & 5 & 3 & 2 \\ 0 & 2 & 2 & -16 \\ 0 & -2 & 2 & 17 \end{vmatrix}$$

$$=-\begin{vmatrix} 5 & 3 & 2 \\ 2 & 2 & -16 \\ -2 & 2 & 17 \end{vmatrix}=-340.$$

例2　计算$n+1$阶行列式

$$D_{n+1}=\begin{vmatrix} -a_1 & a_1 & 0 & \cdots & 0 \\ 0 & -a_2 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -a_n & a_n \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix}.$$

解　将第1列至第n列都加到第$n+1$列上去,然后由第$n+1$列展开.

$$D_{n+1}=(n+1)\begin{vmatrix} -a_1 & a_1 & & & \\ & -a_2 & a_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & a_{n-1} \\ & & & & -a_n \end{vmatrix}$$

$$= (-1)^n (n+1) a_1 a_2 \cdots a_n.$$

【行列式为零的充分条件】

(1)**有两行(或两列)相同;**

(2)**有两行(或两列)元素成比例.**

【方阵的行列式公式】设 A,B **为** n **阶方阵,** I **为单位矩阵,** λ **为实数,** $|A|$ **表示** A **中元素构成的行列式,则**

$$|AB| = |A||B|, \quad |\lambda A| = \lambda^n |A|,$$

$$|A^{\mathrm{T}}| = |A|, \quad |A^{-1}| = |A|^{-1} \quad (A\text{ 可逆}),$$

$$|A^*| = |A|^{n-1}, \quad |I| = 1.$$

【特殊矩阵的行列式】

(1)**三角形矩阵的行列式等于主对角线上的元素的乘积;**

(2)**如果** n **阶方阵** A **有特征值** $\lambda_1,\lambda_2,\cdots,\lambda_n$**,则**

$$|A| = \lambda_1 \lambda_2 \cdots \lambda_n;$$

(3)A **有特征值** $a \Leftrightarrow |A - aI| = 0$;

(4)**设** A **是正交阵,则** $|A|^2 = 1$, $|A| = \pm 1$;

(5)**设** A **是** m **阶方阵,** B **是** n **阶方阵,则**

$$\begin{vmatrix} A & 0 \\ * & B \end{vmatrix} = \begin{vmatrix} A & * \\ 0 & B \end{vmatrix} = |A||B|,$$

$$\begin{vmatrix} 0 & A \\ B & * \end{vmatrix} = \begin{vmatrix} * & A \\ B & 0 \end{vmatrix} = (-1)^{mn} |A||B|.$$

【范德蒙行列式】

$$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-2} & a_2^{n-2} & \cdots & a_n^{n-2} \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1)$$
$$(a_3 - a_2)(a_4 - a_2)\cdots(a_n - a_2)\cdots(a_n - a_{n-1}).$$

常见题型·应对

【行列式计算法概述】计算行列式的方法较多,有

(1)**利用初等变换化作三角形行列式;**

(2)**利用展开公式化作低阶行列式;**

(3)**利用矩阵关系式计算方阵的行列式;**

(4)**利用列向量或行向量属性计算行列式.**

【利用矩阵关系计算方阵的行列式】设法得到所求行列式对应矩阵的乘积关系式,再利用方阵行列式公式计算.

例1 设$\boldsymbol{A},\boldsymbol{B}$为$n$阶方阵,$|\boldsymbol{A}|=2$,$|\boldsymbol{B}|=-3$,求$|2\boldsymbol{A}^*\boldsymbol{B}^{-1}|$.

解 $|2\boldsymbol{A}^*\boldsymbol{B}^{-1}|=2^n|\boldsymbol{A}^*||\boldsymbol{B}^{-1}|$

$$=2^n|\boldsymbol{A}|^{n-1}|\boldsymbol{B}|^{-1}=-2^{2n-1}/3.$$

例2 设$\boldsymbol{ABA}^*=2\boldsymbol{BA}^*+\boldsymbol{I}$,$\boldsymbol{A}=\begin{pmatrix}2&1&0\\1&2&0\\0&0&1\end{pmatrix}$,求$|\boldsymbol{B}|$.

解 将已知等式化成$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{BA}^*=\boldsymbol{I}$,取行列式得$|\boldsymbol{A}-2\boldsymbol{I}||\boldsymbol{B}||\boldsymbol{A}^*|=1$;由于$|\boldsymbol{A}-2\boldsymbol{I}|=1$,$|\boldsymbol{A}|=3$,故$|\boldsymbol{A}^*|=|\boldsymbol{A}|^{3-1}=9$,$|B|=\dfrac{1}{9}$.

例3 设$\boldsymbol{A}$为n阶方阵,$|\boldsymbol{A}|=a\neq0$,求

$$D=|\boldsymbol{A}^{-1}+2\boldsymbol{A}^*|.$$

解 分解因式,有

$$\boldsymbol{A}^{-1}+2\boldsymbol{A}^*=\boldsymbol{A}^{-1}(\boldsymbol{I}+2\boldsymbol{AA}^*)$$

$$= \boldsymbol{A}^{-1}(\boldsymbol{I} + 2|\boldsymbol{A}|\boldsymbol{I})$$
$$= \boldsymbol{A}^{-1}(1 + 2|\boldsymbol{A}|)\boldsymbol{I},$$

故 $D = |\boldsymbol{A}|^{-1}(2|\boldsymbol{A}| + 1)^n = a^{-1}(2a + 1)^n.$

例 4 设方阵 $A=(\alpha_1,\alpha_2,\alpha_3)$ **的行列式为** a，**求** A **的列向量构造的矩阵**

$$\boldsymbol{B} = (\alpha_1, 4\alpha_1 + 2\alpha_2, 5\alpha_1 + 6\alpha_2 + 3\alpha_3)$$

的行列式之值.

解 $\boldsymbol{B}=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1 & 4 & 5\\ 0 & 2 & 6\\ 0 & 0 & 3\end{pmatrix},$

故 $$|\boldsymbol{B}|=|\boldsymbol{A}|\cdot\begin{vmatrix}1 & 4 & 5\\ 0 & 2 & 6\\ 0 & 0 & 3\end{vmatrix}=6a.$$

【利用列向量之间的关系化简行列式】通常用来比较两个以列向量形式出现的抽象行列式的值.

例 1 设 4 阶方阵

$$\boldsymbol{A}=(\xi,\alpha,\beta,\gamma),\quad \boldsymbol{B}=(\eta,\beta,\gamma,\alpha),$$
$$|\boldsymbol{A}|=1,\quad |\boldsymbol{B}|=2,$$

求 $|\boldsymbol{A}+\boldsymbol{B}|$.

解

$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{B}| &= |\xi+\eta,\alpha+\beta,\beta+\gamma,\gamma+\alpha|\\
&=|\xi+\eta,2(\alpha+\beta+\gamma),\beta+\gamma,\gamma+\alpha| && ①\\
&=2|\xi+\eta,\alpha+\beta+\gamma,-\alpha,-\beta| && ②\\
&=2|\xi+\eta,\gamma,-\alpha,-\beta| && ③\\
&=2(|\xi,\gamma,-\alpha,-\beta|+|\eta,\gamma,-\alpha,-\beta|) && ④\\
&=2(|\xi,\alpha,\beta,\gamma|+|\eta,\beta,\gamma,\alpha|) && ⑤\\
&=2(|\boldsymbol{A}|+|\boldsymbol{B}|)=6.\end{aligned}$$

其中:①第3、第4列加至第2列;②提因子2后,第2列乘以-1加到第3、第4列;③第3、第4列加至第2列;④行列式拆项性质;⑤两列互换.

例2　设$\alpha,\beta,\gamma,\xi,\eta$是4维列向量,$|\alpha,\beta,\gamma,\xi|=m$,$|\alpha,\beta,\eta,\gamma|=n$,求$D=|\gamma,\beta,\alpha,\xi+\eta|$.

解　直接使用拆项性质以及对换变形,得

$$\begin{aligned}D&=|\gamma,\beta,\alpha,\xi|+|\gamma,\beta,\alpha,\eta|\\&=-|\alpha,\beta,\gamma,\xi|+|\alpha,\beta,\eta,\gamma|\\&=n-m.\end{aligned}$$

【化成三角形计算行列式】利用行或列初等变换将行列式化成三角形后,主对角线上元素之积便是行列式之值.

【使用降阶法计算行列式】应用展开公式

$$|\boldsymbol{A}|=\sum_{k=1}^{n}a_{ik}A_{ik}=\sum_{k=1}^{n}a_{kj}A_{kj}$$

可以将n阶行列式$|\boldsymbol{A}|$的计算归于$n-1$阶行列式A_{ij}的计算.如果选择展开的第i行或第j列中零元较多,则可以简化行列式计算,得到递归公式或求得行列式之值.

例　以下行列式依第1列展开计算,得

$$D_n=\begin{vmatrix}a&b&0&\cdots&0\\0&a&b&\ddots&\vdots\\\vdots&\ddots&\ddots&\ddots&0\\0&\cdots&0&a&b\\b&0&\cdots&0&a\end{vmatrix}$$

$$= a\begin{vmatrix} a & b & \\ & \ddots & \ddots \\ & & a \end{vmatrix} + b\cdot(-1)^{n+1}\begin{vmatrix} b & & \\ a & \ddots & \\ & \ddots & b \end{vmatrix}$$

$$= a^n - (-b)^n.$$

【适合于降阶法的行列式特征】

(1)**某行(列)的零元素较多；**

(2)**某行(列)的元素全相同,从而可以将该列化简为至多一个非零元素；**

(3)**各行(列)的元素之和相同,从而可以统一加至第 1 列(行)变作**(2)；

(4)**能够用拆项法产生特征**(1)～(3).

例 1　**求** $D=\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$.

解　$D=2(x+y)\begin{vmatrix} 1 & y & x+y \\ 1 & x+y & x \\ 1 & x & y \end{vmatrix}$(**统加**)

$$=2(x+y)\begin{vmatrix} 1 & y & x+y \\ 0 & x & -y \\ 0 & x-y & -x \end{vmatrix}\text{(三角化)}$$

$$=2(x+y)\begin{vmatrix} x & -y \\ x-y & -x \end{vmatrix}$$

$$=-2(x^3+y^3).$$

例 2　**求** $D=\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1-a & 1 & 1 \\ 1 & 1 & 1+b & 1 \\ 1 & 1 & 1 & 1-b \end{vmatrix}$.

解

$$D=\begin{vmatrix}1&1&1&1\\1&1-a&1&1\\1&1&1+b&1\\1&1&1&1-b\end{vmatrix}+\begin{vmatrix}a&1&1&1\\0&1-a&1&1\\0&1&1+b&1\\0&1&1&1-b\end{vmatrix}\text{(分拆第一列)}$$

$$=\begin{vmatrix}1&1&1&1\\0&-a&0&0\\0&0&b&0\\0&0&0&-b\end{vmatrix}+a\begin{vmatrix}1-a&1&1\\1&1+b&1\\1&1&1-b\end{vmatrix}$$

$$=ab^2+a\left(\begin{vmatrix}1&1&1\\1&1+b&1\\1&1&1-b\end{vmatrix}+\begin{vmatrix}-a&1&1\\0&1+b&1\\0&1&1-b\end{vmatrix}\right)$$

$$=ab^2+a\left(\begin{vmatrix}1&1&1\\0&b&0\\0&0&-b\end{vmatrix}-a\begin{vmatrix}1+b&1\\1&1-b\end{vmatrix}\right)$$

$$=ab^2+a(-b^2+ab^2)=a^2b^2.$$

【箭头形行列式计算示范】

例 1 **求** $D_4=\begin{vmatrix}a_0&b_1&b_2&b_3\\c_1&a_1&0&0\\c_2&0&a_2&0\\c_3&0&0&a_3\end{vmatrix}\quad(a_i\neq 0).$

解　提因子使得主对角线上出现三个 1,再用第 2、第 3、第 4 列元素改变第 1 列产生三个 0:

$$D_4 = a_1a_2a_3\begin{vmatrix} a_0 & b_1 & b_2 & b_3 \\ \dfrac{c_1}{a_1} & 1 & 0 & 0 \\ \dfrac{c_2}{a_2} & 0 & 1 & 0 \\ \dfrac{c_3}{a_3} & 0 & 0 & 1 \end{vmatrix}$$

$$= a_1a_2a_3\begin{vmatrix} a_0 - \dfrac{b_1c_1}{a_1} & b_1 & b_2 & b_3 \\ 0 & 1 & 0 & 0 \\ \dfrac{c_2}{a_2} & 0 & 1 & 0 \\ \dfrac{c_3}{a_3} & 0 & 0 & 1 \end{vmatrix}$$

$$= \cdots = a_1a_2a_3\begin{vmatrix} s & b_1 & b_2 & b_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$= a_1a_2a_3s \quad \left(s = a_0 - \sum_{i=1}^{3}\frac{b_ic_i}{a_i}\right).$$

例 2　**求** $D_4 = \begin{vmatrix} x+1 & x & x & x \\ x & x+2 & x & x \\ x & x & x+3 & x \\ x & x & x & x+4 \end{vmatrix}$.

解　第 1 行乘以 -1 加至其他行即产生箭头形行列式. 可求出 $D_4 = 50x + 24$.

【三对角行列式计算示范】

例 1　**计算** $D_n=\begin{vmatrix} a & y & & & \\ x & a & y & & \\ & x & \ddots & \ddots & \\ & & \ddots & a & y \\ & & & x & a \end{vmatrix}$.

解　**依第** 1 **行展开,便可以得到递推公式**

$$D_n = aD_{n-1} - xyD_{n-2},\quad n > 2.$$

从而可以由低到高或者由高到低计算.

例 2　$D_3=\begin{vmatrix} 1-a & a & \\ -1 & 1-a & a \\ & -1 & 1-a \end{vmatrix}$

$$=(1-a)D_2+aD_1=1-a+a^2-a^3.$$

【依定义计算行列式中的某一项】依据定义,行列式中的各项是由不同行且不同列的元素相乘构成的,一些试题专门考察这一点,尤其是符号的确定法则.

例　$D_4=\begin{vmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{vmatrix}$ **的展开式中包含** x^4 **与** x^3 **的项是什么?**

解　**分析含** x^4 **项,它只能由主对角线元素构成,故为** $10x^4$**;而含** x^3 **项只可以是**

$$a_{21}a_{12}a_{33}a_{44}\cdot(-1)^{\tau(2134)}=-2x^3$$

以及　$(-1)^{\tau(4231)}a_{41}a_{22}a_{33}a_{14}=-3x^3.$

故所求项为　$-5x^3+10x^4.$

【利用代数余子式表示行列式】设 D_n **为** n **阶行列式,**

$A_{ij}(1\leqslant i,j\leqslant n)$是其中元素的代数余子式,则有以下考点:

(1)某行(或某列)的代数余子式之线性组合,如

$$b_1A_{11}+b_2A_{12}+\cdots+b_nA_{1n}$$

恰是 D_n 中以$(b_1,b_2,\cdots,b_n)$替代第 1 行元素之后的新的行列式之值.

(2)主对角线上的代数余子式之和,如

$$A_{11}+A_{22}+\cdots+A_{nn}$$

恰是伴随矩阵 A^* 之特征值之和. 如果 A 可逆, $\lambda_i(1\leqslant i\leqslant n)$是 A 的特征值,则$\frac{|A|}{\lambda_i}$是 A^* 的特征值,从而

$$\begin{aligned}&A_{11}+A_{22}+\cdots+A_{nn}\\&\quad=|A|\left(\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\cdots+\frac{1}{\lambda_n}\right).\end{aligned}$$

【计算多项式矩阵的行列式】方阵 A 的多项式矩阵的行列式通常使用特征值方法来求.

例 1　设 A 为三阶方阵,α 为三阶向量,向量组 $\alpha, A\alpha, A^2\alpha$ 线性无关,且 $A^3\alpha=4A\alpha-3A^2\alpha$,求 $|2A^2+3I|$.

解　条件推出 $A(A+I)(A-4I)\alpha=\mathbf{0}$,若 $|A|\neq 0$,则左乘 A^{-1} 得

$$(A+I)(A-4I)\alpha=-4\alpha-3A\alpha+A^2\alpha=\mathbf{0},$$

与 $\alpha, A\alpha, A^2\alpha$ 线性无关矛盾,故 $|A|=0$. 类似地推出 $|A+I|=0$ 及 $|A-4I|=0$,于是 A 有三个特征值 $0,-1,4$,从而 $2A^2+3I$ 有特征值 $3,5,35$,它们的积便是所求行列式的值,为 525.

注　特征值的传递性质见 13.1 节.

例 2　求矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_1b_1+x & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2+x & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}b_n \\ a_nb_1 & \cdots & a_nb_{n-1} & a_nb_n+x \end{pmatrix}$$

的行列式,其中 $\sum\limits_{i=1}^{n} a_ib_i \neq 0$.

解　将所给矩阵进行分解,得 $\boldsymbol{A}=x\boldsymbol{I}+\boldsymbol{B}$,其中 $\boldsymbol{B}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}(b_1,b_2,\cdots,b_n)$ 是对称矩阵,依条件知 $r(\boldsymbol{B})=1$,从而它只有一个非零特征值,记为 λ,由于特征值的和等于矩阵的迹,故

$$\lambda=\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(\boldsymbol{B})=\sum_{i=1}^{n} a_ib_i,$$

因而 $\boldsymbol{B}$ 的特征值为 $\mathrm{tr}(\boldsymbol{B}),0,\cdots,0$;从而(依据特征值的传递性质)$A$ 的特征值为 $\mathrm{tr}(\boldsymbol{B})+x,x,\cdots,x$,它们的积即为所求:$|\boldsymbol{A}|=(\mathrm{tr}(\boldsymbol{B})+x)x^{n-1}$.

【快速计算特征值的一种方法】三阶矩阵的特征值计算是经常遇到的行列式运算.可以考虑使用 $|\boldsymbol{A}-\lambda\boldsymbol{I}|$ 的第 2 行(或第 2 列),将位于(1,3)及(3,1)的元素化作 0,之后观察行列式是否出现一次因式 $\lambda-a$.

例 求矩阵$A=\begin{pmatrix}1 & 2 & -3\\ -1 & 4 & -3\\ 1 & -2 & 5\end{pmatrix}$的特征值.

解 第2行乘以-1加到第1行，第2行直接加到第3行，得

$$\begin{aligned}|\boldsymbol{A}-\lambda\boldsymbol{I}| &= \begin{vmatrix}1-\lambda & 2 & -3\\ -1 & 4-\lambda & -3\\ 1 & -2 & 5-\lambda\end{vmatrix}\\ &= \begin{vmatrix}2-\lambda & \lambda-2 & 0\\ -1 & 4-\lambda & -3\\ 0 & 2-\lambda & 2-\lambda\end{vmatrix}\\ &= (\lambda-2)^2\begin{vmatrix}-1 & 1 & 0\\ -1 & 4-\lambda & -3\\ 0 & -1 & -1\end{vmatrix}\\ &= (\lambda-2)^2(6-\lambda).\end{aligned}$$

故A的特征值为2，2，6.

第10章　矩　　阵

大纲要求

理解矩阵的概念，了解单位矩阵、数量矩阵、对角矩阵、三角矩阵、对称矩阵和反对称矩阵，以及它们的性质. 掌握矩阵的线性运算、乘法、转置，以及它们的运算规律，了解方阵的幂以及方阵乘积的行列式. 理解逆矩阵的概念，掌握逆矩阵的性质，以及矩阵可逆的充分必要条件，理解伴随矩阵的概念，会用伴随矩阵求逆矩阵. 掌握矩阵的初等变换，了解初等矩阵的性质和矩阵等价的概念，理解矩阵的秩的概念，掌握用初等变换求矩阵的逆矩阵和秩的方法，了解分块矩阵及其运算.

注　数学二不要求分块矩阵，由于也不要求第14章，故所要求的正交矩阵放在这一章.

知识点·点拨

10.1　矩阵的基本概念

【矩阵概念】数域 F（例如实数全体）上的 $m\times n$ 个数 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 排成 m 行 n 列的数表

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称为 $m\times n$ 矩阵，简记为 $A=(a_{ij})_{m\times n}$，其中 a_{ij} 是矩阵 A 的第 i 行第 j 列的元素. 元素属于实数域的矩阵叫做实矩阵；属于复数域的矩阵叫做复矩阵.

【同型矩阵】若两矩阵的行数、列数分别相等，则称它们是同型矩阵.

【方阵】行数与列数相等的矩阵称为方阵. 当行数和列数均为 n 时，也称为 n 阶方阵.

注　本书中未指明阶数的方阵默认为 n 阶方阵.

【方阵行列式】由 n 阶方阵 A 的元素所构成的行列式叫做方阵 A 的行列式，记为 $|A|$ 或 $\det A$.

【矩阵的相等】同型且对应位置的元素都相等的矩阵.

【零矩阵】所有元素都为零的矩阵.

【单位矩阵】主对角线上元素都为 1，其余元素都为零的方阵：

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

记为 I 或 E，其中，未写出的元素默认为零元.

【对角矩阵】主对角线以外的元素皆为零的 n 阶方阵，记作

$$\Lambda=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

【上(下)三角矩阵】主对角线的下(上)方元素都为零的方阵,如$\boldsymbol{A}$($\boldsymbol{B}$)所示:

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix},$$

$$\boldsymbol{B}=\begin{pmatrix} b_{11} & & & \\ b_{21} & b_{22} & & \\ \vdots & \vdots & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

【行矩阵或行向量】只有一行的矩阵,记为

$$(a_1,a_2,\cdots,a_n).$$

【列矩阵或列向量】只有一列的矩阵,记为

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

或　$$(a_1,a_2,\cdots,a_m)^{\mathrm{T}}.$$

【转置矩阵】将$m\times n$矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

的行与列互换后得到如下的一个$n\times m$矩阵,称为$\boldsymbol{A}$的转置矩阵,记为$\boldsymbol{A}^{\mathrm{T}}$,即

$$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

显然A^{T}的第i行第j列的元素等于A的第j行第i列的元素.

【矩阵转置运算律】

(1)$(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}$;

(2)对方阵$\boldsymbol{A}$,有$|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$;

(3)$(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$,$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$;

(4)$(k\boldsymbol{A})^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}$,$k$为数域中的数;

(5)$(\boldsymbol{A}^{\mathrm{T}})^{*}=(\boldsymbol{A}^{*})^{\mathrm{T}}$;

(6)$(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}$($\boldsymbol{A}$为可逆矩阵).

【矩阵的加法】两个同型矩阵$\boldsymbol{A}=(a_{ij})_{m\times n}$与$\boldsymbol{B}=(b_{ij})_{m\times n}$的和规定为矩阵

$$\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix},$$

记作
$$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}.$$

【矩阵加法运算律】

(1)交换律:$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$;

(2)结合律:$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$;

(3)$\boldsymbol{A}+\boldsymbol{0}=\boldsymbol{0}+\boldsymbol{A}=\boldsymbol{A}$,其中$\boldsymbol{0}$为与$\boldsymbol{A}$同型的零矩阵.

【矩阵的数乘】矩阵A与数k的乘积规定为矩阵

$$\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix},$$

记作
$$k\boldsymbol{A} = (ka_{ij})_{m\times n}.$$

【数量矩阵】单位矩阵与数量k的积,$k\boldsymbol{I}$.

【负矩阵】称$(-1)\boldsymbol{A}$为$\boldsymbol{A}$的负矩阵,记作$-\boldsymbol{A}$.

【矩阵的减法】两个同型矩阵$\boldsymbol{A}$与$\boldsymbol{B}$的减法定义为

$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B}).$$

【线性运算律】矩阵的加法与数乘称为矩阵的线性运算.它们满足下面的运算律.

(1)结合律:$(kl)\boldsymbol{A}=k(l\boldsymbol{A})$.

(2)分配律:$k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$;

$$(l+k)\boldsymbol{A} = l\boldsymbol{A} + k\boldsymbol{A}.$$

例　设矩阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 2 \end{pmatrix},\quad \boldsymbol{B} = \begin{pmatrix} -2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix},$$

则

$$\begin{aligned} 2\boldsymbol{A} - 3\boldsymbol{B} &= 2\begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 2 \end{pmatrix} - 3\begin{pmatrix} -2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & 2 \\ -2 & 0 & 4 \end{pmatrix} + \begin{pmatrix} 6 & -3 & 0 \\ -3 & -3 & -6 \end{pmatrix} \\ &= \begin{pmatrix} 8 & -3 & 2 \\ -5 & -3 & -2 \end{pmatrix}. \end{aligned}$$

【矩阵的乘法】设$\boldsymbol{A}$是$m\times k$矩阵,$\boldsymbol{B}$是$k\times n$矩阵,

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix},$$

$$B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix},$$

A 与 B 的乘积定义为 $m\times n$ 矩阵 C，即

$$C=AB=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix},$$

其中，$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ik}b_{kj}=\sum\limits_{t=1}^{k}a_{it}b_{tj}$，即乘积矩阵 C 的第 i 行第 j 列的元素等于矩阵 A 的第 i 行元素与矩阵 B 的第 j 列对应元素乘积之和.

〖点拨〗矩阵乘法不满足交换律. 若 A 是 3×2 矩阵，B 是 2×4 矩阵，则 AB 为 3×4 矩阵，而 BA 则无意义. 此外，即便 A，B 是两个同阶方阵，也不一定成立 $AB=BA$.

【矩阵乘法运算律】

(1)**结合律**：$(AB)C=A(BC)$.

(2)**左分配律**：$A(B+C)=AB+AC$；

右分配律：$(B+C)A=BA+CA$.

(3)**数乘交换律**：$k(AB)=(kA)B=A(kB)$.

(4)与单位矩阵$\boldsymbol{I}$相乘:如果乘式有意义,则

$$\boldsymbol{IA}=\boldsymbol{A},\boldsymbol{AI}=\boldsymbol{A},\boldsymbol{AIB}=\boldsymbol{AB}.$$

〖点拨〗

(1)矩阵乘法不满足交换律,即一般$\boldsymbol{AB}\neq\boldsymbol{BA}$.因此我们称乘积$\boldsymbol{AB}$为$\boldsymbol{A}$左乘$\boldsymbol{B}$,或$\boldsymbol{B}$右乘$\boldsymbol{A}$.

(2)在数量运算中,若$ab=0$,必有$a=0$或$b=0$,但在矩阵乘积运算中,若$\boldsymbol{AB}=\boldsymbol{0}$,则未必有$\boldsymbol{A}=\boldsymbol{0}$或$\boldsymbol{B}=\boldsymbol{0}$,例如

$$\boldsymbol{A}=\begin{pmatrix}1&0\\1&0\end{pmatrix}\neq\boldsymbol{0},\quad \boldsymbol{B}=\begin{pmatrix}0&0\\1&1\end{pmatrix}\neq\boldsymbol{0},$$

但 $$\boldsymbol{AB}=\begin{pmatrix}1&0\\1&0\end{pmatrix}\begin{pmatrix}0&0\\1&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=\boldsymbol{0}.$$

(3)矩阵乘法不成立消去律.就是说,若$\boldsymbol{AB}=\boldsymbol{AC}$,则未必有$\boldsymbol{B}=\boldsymbol{C}$,例如

$$\boldsymbol{A}=\begin{pmatrix}1&-1\\-2&2\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}3&1\\2&-1\end{pmatrix},$$

$$\boldsymbol{C}=\begin{pmatrix}2&-1\\1&-3\end{pmatrix},$$

有$\boldsymbol{AB}=\boldsymbol{AC}=\begin{pmatrix}1&2\\-2&-4\end{pmatrix}$,而$\boldsymbol{B}\neq\boldsymbol{C}$.

【方阵的迹】方阵$\boldsymbol{A}$的主对角线上的元素的和称为$\boldsymbol{A}$的迹,记作$\operatorname{tr}(\boldsymbol{A})=\sum\limits_{i=1}^{n}a_{ii}$.

【乘积矩阵的行列式和迹】设$\boldsymbol{A},\boldsymbol{B}$为$n$阶方阵,则有

$$|\boldsymbol{AB}|=|\boldsymbol{BA}|=|\boldsymbol{A}||\boldsymbol{B}|,\quad \operatorname{tr}(\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BA}).$$

【方阵的正整数幂】设$\boldsymbol{A}$为方阵,k为正整数,则记

$$A^0 = I, \underbrace{AA\cdots A}_{k} = A^k,$$

注　一般地，$(AB)^k \neq A^k B^k$.

【方阵的多项式】设x的m次多项式为

$$h(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_{m-1}x + a_m,$$

则称矩阵$a_0 A^m + a_1 A^{m-1} + \cdots + a_{m-1}A + a_m I$为方阵$A$的$m$次多项式，记作$h(A)$.

【逆矩阵 · 可逆】设A为n阶方阵，若存在n阶方阵B，使得

$$AB = BA = I,$$

则称B为A的逆矩阵，记A的逆为$B = A^{-1}$，并称A为可逆矩阵或满秩矩阵，或非奇异矩阵.

〖点拨〗若A可逆，则A的逆矩阵是惟一的.

【伴随矩阵】n阶方阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

中元素a_{ij}的代数余子式A_{ij}构成如下的n阶方阵

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

称之为矩阵A的伴随矩阵.

〖点拨〗对任何方阵A：$AA^* = A^* A = |A| I$.

【逆矩阵公式】方阵A可逆的充要条件是$|A| \neq 0$，且

$$A^{-1} = A^* / |A|.$$

【逆矩阵运算性质】可逆矩阵有如下重要性质:

(1) $(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$;

(2) $(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}$,$(\boldsymbol{A}^{-1})^{*}=(\boldsymbol{A}^{*})^{-1}$;

(3) $(k\boldsymbol{A})^{-1}=\dfrac{1}{k}\boldsymbol{A}^{-1}$($k$ 为非零的数);

(4) $|\boldsymbol{A}^{-1}|=\dfrac{1}{|\boldsymbol{A}|}=|\boldsymbol{A}|^{-1}$;

(5) $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$;

(6) 如果方阵 A 满足多项式方程

$$\boldsymbol{A}^{m}+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{I}=\boldsymbol{0}\quad(a_m\neq 0),$$

则 $$\boldsymbol{A}^{-1}=-\frac{1}{a_m}(\boldsymbol{A}^{m-1}+a_1\boldsymbol{A}^{m-2}+\cdots+a_{m-1}\boldsymbol{I}).$$

【分块矩阵】用贯穿于矩阵的纵线和横线分割矩阵 A 为若干块,每小块叫做矩阵 A 的子块(子矩阵),以子块为元素的矩阵叫做分块矩阵. 例如

$$\boldsymbol{A}=\left(\begin{array}{cc:ccc}1&0&1&1&2\\0&1&-1&2&3\\ \hdashline 3&0&5&1&0\\2&3&1&2&1\end{array}\right)=\left(\begin{array}{c:c}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}\\ \hdashline \boldsymbol{A}_{21}&\boldsymbol{A}_{22}\end{array}\right),$$

其中,子块

$$\boldsymbol{A}_{11}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \boldsymbol{A}_{12}=\begin{pmatrix}1&1&2\\-1&2&3\end{pmatrix},$$

$$\boldsymbol{A}_{21}=\begin{pmatrix}3&0\\2&3\end{pmatrix},\quad \boldsymbol{A}_{22}=\begin{pmatrix}5&1&0\\1&2&1\end{pmatrix}.$$

在向量和方程问题中常把 A 按列分块:

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n),$$

其中　$\alpha_i=(a_{1i},a_{2i},\cdots,a_{mi})^{\mathrm{T}}$　$i=1,2,\cdots,n$.

【分块矩阵的加法】

设$\boldsymbol{A},\boldsymbol{B}$为$m\times n$矩阵,用相同分法把$\boldsymbol{A}$与$\boldsymbol{B}$分块为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{pmatrix},$$

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1s}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{r1} & \boldsymbol{B}_{r2} & \cdots & \boldsymbol{B}_{rs}\end{pmatrix},$$

其中每一个$\boldsymbol{A}_{ij}$与$\boldsymbol{B}_{ij}$是同型子块矩阵,则

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12} & \cdots & \boldsymbol{A}_{1s}+\boldsymbol{B}_{1s}\\ \boldsymbol{A}_{21}+\boldsymbol{B}_{21} & \boldsymbol{A}_{22}+\boldsymbol{B}_{22} & \cdots & \boldsymbol{A}_{2s}+\boldsymbol{B}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1}+\boldsymbol{B}_{r1} & \boldsymbol{A}_{r2}+\boldsymbol{B}_{r2} & \cdots & \boldsymbol{A}_{rs}+\boldsymbol{B}_{rs}\end{pmatrix}.$$

这里我们把每个小块看作新矩阵的元素进行运算.

【数乘分块矩阵】数乘分块矩阵:

$$k\boldsymbol{A}=\begin{pmatrix}k\boldsymbol{A}_{11} & k\boldsymbol{A}_{12} & \cdots & k\boldsymbol{A}_{1s}\\ k\boldsymbol{A}_{21} & k\boldsymbol{A}_{22} & \cdots & k\boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ k\boldsymbol{A}_{r1} & k\boldsymbol{A}_{r2} & \cdots & k\boldsymbol{A}_{rs}\end{pmatrix}.$$

【分块矩阵的转置】转置一个分块矩阵时,在分块矩阵中除了作行、列位置互换外,还要对每一个子矩阵做转置.例如

$$A = \left(\begin{array}{cc:cc:c} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 4 & 5 & 2 \\ \hdashline 1 & 4 & 3 & 5 & 6 \end{array}\right),$$

则

$$A^{\mathrm{T}} = \left(\begin{array}{cc:c} 1 & 0 & 1 \\ 0 & 1 & 4 \\ \hdashline 2 & 4 & 3 \\ 1 & 5 & 5 \\ \hdashline 1 & 2 & 6 \end{array}\right).$$

【分块矩阵的乘法】设矩阵 A 左乘 B 有意义,如果 A 的列分块方法与 B 的行分块分法一致,例如:

$$A = \left(\begin{array}{cc:c} 2 & 0 & 5 \\ \hdashline 4 & 2 & 1 \\ 0 & -1 & 2 \end{array}\right), \quad B = \left(\begin{array}{c:cc:c} 1 & 2 & 4 & -1 \\ 5 & 3 & 1 & 0 \\ \hdashline -1 & 0 & 2 & 3 \end{array}\right).$$

则可以将每个小块看作矩阵元素依普通矩阵乘法规则作乘积.例如常用的一种分块乘积为

$$A(\alpha_1,\alpha_2,\cdots,\alpha_n) = (A\alpha_1,A\alpha_2,\cdots,A\alpha_n)$$

10.2　几种常用矩阵的性质归纳

对以下几种常见的 n 阶方阵,我们从基本特征、行列式、特征值、逆矩阵、典型例子、运算保持性质等方面进行归纳并对难点给予适当的解释.

【对称矩阵】设 A 为 n 阶方阵,若 $A^{\mathrm{T}}=A$,则称 A 为对称矩阵,例如

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 2 & 0 & -3 \end{pmatrix}.$$

〖元素特征〗n 阶矩阵 A 为对称矩阵的充分必要条件是其元素关于主对角线对称：$a_{ij}=a_{ji}$.

〖运算保持性〗

(1)若矩阵 A 对称，则它的转置矩阵、伴随矩阵、逆矩阵仍然是对称矩阵；

(2)两个同阶对称矩阵的线性组合仍然是对称矩阵；

(3)两个可交换顺序的对称矩阵的乘积矩阵仍然是对称矩阵.

〖特征值〗实对称矩阵的特征值是实数.

〖特征向量〗实对称矩阵的属于不同特征值的特征向量正交.

〖对角化〗实对称矩阵一定可以用正交矩阵相似成一个由特征值构成的对角矩阵.

【反对称矩阵】设 A 为 n 阶方阵，若 $A^{\mathrm{T}}=-A$，则称 A 为反对称矩阵.

〖元素特征〗n 阶矩阵为反对称矩阵的充分必要条件是 $a_{ij}=-a_{ji}$. 特别地，反对称矩阵的主对角线上元素都为零，例如：

$$\boldsymbol{B}=\begin{pmatrix}0 & 1 & -2\\ -1 & 0 & 3\\ 2 & -3 & 0\end{pmatrix}.$$

〖运算〗反对称矩阵的平方是对称矩阵.

〖特征值〗反对称矩阵的特征值为零或纯虚数.

〖二次型〗对任何向量 α，$\alpha^{\mathrm{T}}A\alpha=0$.

〖可逆问题〗若实数 $a\neq0$，则 $\boldsymbol{I}+aA$ 可逆.

〖行列式〗若n为奇数,则n阶方阵A的行列式为0,矩阵不可逆.

【幂等矩阵】满足$A^2=A$的方阵,例如单位矩阵.

〖特征值〗特征值只能取0或1.

〖对角化〗可以相似于对角阵$\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix}$.

【对合矩阵】满足$A^2=\boldsymbol{I}$的方阵,例如$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$.

〖行列式〗行列式为-1或1,从而一定可逆.

〖特征值〗特征值只能取-1或1.

〖秩关系〗 $r(\boldsymbol{I}-A)+r(\boldsymbol{I}+A)=n$.

【正交矩阵】满足$AA^{\mathrm{T}}=\boldsymbol{I}$的方阵,例如单位矩阵.

〖元素特征〗如果某个元素为1或-1,则它所在行和列的所有其他元素都为零.

〖正交性〗任何两个列(行)向量的内积为零.

〖规范性〗每一列(行)向量的元素的平方和为1.

〖行列式〗正交矩阵的行列式为1或-1.

〖可逆性〗正交矩阵是可逆矩阵:$A^{-1}=A^{\mathrm{T}}$.

〖特征值〗特征值不为0.

〖保持性〗

(1)正交矩阵的转置矩阵、逆矩阵、伴随矩阵还是正交矩阵;

(2)两个同阶的正交矩阵的积还是正交矩阵.

〖充要条件〗方阵A是正交矩阵$\Leftrightarrow A^{-1}=A^{\mathrm{T}}\Leftrightarrow$列向量组为标准正交(彼此正交且模为1).

【幂等、对合、正交之关系】由实方阵中幂等、对合、

正交三个性质中任意两条可以推出第三条.

【正定矩阵】对任何非零实向量α均有$\alpha^{\mathrm{T}}A\alpha>0$的实对称方阵$A$.

例　对角线上都是正实数的对角矩阵.

〖行列式〗正定矩阵的行列式为正数.

〖可逆性〗正定矩阵为可逆矩阵.

〖特征值〗正定矩阵的特征值都为正数.

〖保持性〗

(1)正定矩阵的转置矩阵、逆矩阵、伴随矩阵还是正定矩阵;

(2)两个同阶的正定矩阵的正系数线性组合还是正定矩阵;

(3)两个可交换次序的正定矩阵的乘积矩阵还是正定矩阵.

〖充要条件〗实对称方阵:正定$\Leftrightarrow$合同于单位矩阵$\Leftrightarrow$所有顺序主子式大于0$\Leftrightarrow$特征值都为正数.

10.3　矩阵初等变换与初等矩阵

【矩阵的初等变换】下列三种变换为初等行变换.

(1)对换变换:互换矩阵第i行与第j行的位置,记作$r_i\leftrightarrow r_j$;

(2)数乘变换:用一个非零常数k乘以矩阵的第i行,记作kr_i;

(3)倍加变换:将矩阵的第j行元素的k倍加到第i行上,记作r_i+kr_j.

若把定义中的行换成列,则是矩阵的初等列变换,相应地记作

$$c_i \leftrightarrow c_j, kc_i \text{ 和 } c_i + kc_j.$$

矩阵的初等行变换和初等列变换统称为矩阵的初等变换.

【行阶梯形 · 行标准形 · 标准形】设 A 是一个 $m \times n$ 矩阵,通过初等行变换可以把 A 化为如下类型的矩阵,称为行阶梯形:

$$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

若要求非零行的第一个元素为1,则得到行标准形,例如

$$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

进一步,通过初等行变换和初等列变换,可以把矩阵化为如下的形式,称为矩阵的标准形

$$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

【初等矩阵】单位矩阵 I 经过一次初等变换得到的矩阵称为初等矩阵,共有三种. 例如,将 I_3 的第一、二行互换,第三行乘以非零的数 k 以及第一行的 k 倍加到第三行所产生的初等矩阵分别为

$$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix},\quad\begin{pmatrix}1&0&0\\0&1&0\\0&0&k\end{pmatrix},\quad\begin{pmatrix}1&0&0\\0&1&0\\k&0&1\end{pmatrix}.$$

依次记作$\boldsymbol{R}_{12}$,$\boldsymbol{R}_{3(k)}$,$\boldsymbol{R}_{(3,1(k))}$,类似地理解对单位矩阵作初等列变换产生的初等矩阵,记作$\boldsymbol{C}_{ij}$,$\boldsymbol{C}_{i(k)}$,$\boldsymbol{C}_{(i,j(k))}$.

【初等矩阵的逆矩阵】初等矩阵是可逆矩阵,其逆矩阵为

$$\boldsymbol{R}_{ij}^{-1}=\boldsymbol{R}_{ij},\quad \boldsymbol{R}_{i(k)}^{-1}=\boldsymbol{R}_{i\left(\frac{1}{k}\right)},\quad \boldsymbol{R}_{(i,j(k))}^{-1}=\boldsymbol{R}_{(i,j(-k))};$$

$$\boldsymbol{C}_{ij}^{-1}=\boldsymbol{C}_{ij},\quad \boldsymbol{C}_{i(k)}^{-1}=\boldsymbol{C}_{i\left(\frac{1}{k}\right)},\quad \boldsymbol{C}_{(i,j(k))}^{-1}=\boldsymbol{C}_{(i,j(-k))}.$$

【对换产生的初等矩阵的性质】两行或两列对换产生的初等矩阵$\boldsymbol{R}_{ij}$和$\boldsymbol{C}_{ij}$比较特别,满足以下性质:

(1)对等性 $\boldsymbol{R}_{ij}=\boldsymbol{C}_{ij}$;

(2)逆矩阵 $\boldsymbol{R}_{ij}^{-1}=\boldsymbol{R}_{ij}$,$\boldsymbol{C}_{ij}^{-1}=\boldsymbol{C}_{ij}$;

(3)转置矩阵 $\boldsymbol{R}_{ij}^{\mathrm{T}}=\boldsymbol{R}_{ij}$.

在使用对换手段对矩阵A进行合同变换时,产生的结果为$\boldsymbol{B}=\boldsymbol{R}_{ij}\boldsymbol{A}\boldsymbol{C}_{ij}$,由于$\boldsymbol{C}_{ij}^{-1}=\boldsymbol{R}_{ij}$,故此时矩阵$A$和$B$不仅合同而且相似.

【初等变换与初等矩阵】对矩阵A施行某种初等行(列)变换,相当于用相应的行(列)初等矩阵左(右)乘矩阵A.细节如下:

(1) $\boldsymbol{R}_{ij}\boldsymbol{A}(\boldsymbol{A}\boldsymbol{C}_{ij})\Leftrightarrow$交换$A$的$i$,$j$两行(列);

(2) $\boldsymbol{R}_{i(k)}\boldsymbol{A}(\boldsymbol{A}\boldsymbol{C}_{i(k)})\Leftrightarrow$将$A$的第$i$行(列)乘非零数$k$;

(3) $\boldsymbol{R}_{(i,j(k))}\boldsymbol{A}(\boldsymbol{A}\boldsymbol{C}_{(i,j(k))})\Leftrightarrow$将$A$的第$j$行(列)的$k$倍加到第$i$行(列)上.

【矩阵的等价关系】若矩阵 A 经过初等变换化为矩阵 B,则称 A 与 B 等价.

【矩阵的等价定理】设 $m\times n$ 矩阵 A 的秩为 r,则存在 m 阶可逆矩阵 $\boldsymbol{P}$ 和 n 阶可逆矩阵 $\boldsymbol{Q}$,使 A 等价于标准形,即

$$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.$$

特别地,若 A 是一个可逆的 n 阶方阵,则 A 等价于单位矩阵 $\boldsymbol{I}_n$.

【可逆矩阵的初等矩阵分解】可逆矩阵 A 可表示成有限个初等矩阵的乘积.

【初等变换法求逆矩阵】若通过初等行变换使得矩阵 A 变为单位矩阵:$(\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m)A=\boldsymbol{I}$,则逆矩阵为

$$\boldsymbol{A}^{-1}=(\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m)\boldsymbol{I}=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m.$$

流程记作

$$(\boldsymbol{A}\ \vdots\ \boldsymbol{I})\xrightarrow{\text{初等行变换}}(\boldsymbol{I}\ \vdots\ A^{-1}).$$

10.4　矩 阵 的 秩

【矩阵的秩】矩阵 A 中不等于零的子式的最高阶数称为矩阵 A 的秩,记作 $r(A)$. 零矩阵的秩规定为 0.

〖点拨〗若 A 是 B 的子块,则 $r(A)\leqslant r(\boldsymbol{B})$.

【矩阵按秩分解定理】

(1)设矩阵 A 的秩为 r,则 A 可表示为 r 个秩为 1 的矩阵之和.

(2)设 $m\times n$ 矩阵 A 的秩为 r,则存在两个秩为 r 的 $m\times r$ 矩阵 B 和 $r\times n$ 矩阵 C,使得 $A=BC$.

〖推论〗设矩阵A的秩为1,则存在非零列向量α和β,使得$A=\alpha^{\mathrm{T}}\beta$.

【关于秩的基本公式】

$r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})$;
$r(k\boldsymbol{A})=r(\boldsymbol{A})\ (k\neq 0)$;
$r(\boldsymbol{AB})\leqslant\min(r(\boldsymbol{A}),\ r(\boldsymbol{B}))$;
$r(\boldsymbol{AB})\geqslant r(\boldsymbol{A})+r(\boldsymbol{B})-n$　(n是A的列数);
$\boldsymbol{AB}=\boldsymbol{0}$时$r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant n$　(n是A的列数);
$r(\boldsymbol{A}_{m\times n})\leqslant\min\{m,n\}$;
$r(\boldsymbol{A}^*)=\begin{cases}n, & r(\boldsymbol{A})=n,\\ 1, & r(\boldsymbol{A})=n-1,\\ 0, & r(\boldsymbol{A})<n-1,\end{cases}$其中$A$是$n$阶方阵;
若$A\neq\boldsymbol{0}$,则$r(\boldsymbol{A})\geqslant 1$;
$r\begin{pmatrix}\boldsymbol{A} & \\ & \boldsymbol{B}\end{pmatrix}=r(\boldsymbol{A})+r(\boldsymbol{B})$;
$r(\boldsymbol{A}+\boldsymbol{B})\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$;
$\max(r(\boldsymbol{A}),r(\boldsymbol{B}))\leqslant r(\boldsymbol{A}\ \vdots\ \boldsymbol{B})\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$.

【关于秩的基本性质】

(1)对矩阵A进行初等变换不改变$r(A)$的大小.

(2)若$\boldsymbol{P}$可逆,则$r(\boldsymbol{AP})=r(\boldsymbol{A})=r(\boldsymbol{PA})$.

(3)若$r(A)=r$,则A至少有一个r阶子式不为零,且所有高于r阶的子式全为零.

【三秩相等定理】矩阵A的秩与A的行向量组的秩及列向量组的秩相等.

【线性表示与矩阵秩的关系】若矩阵A的列向量组可以由矩阵B的列向量组线性表示,即$BX=A$,则它

们的秩关系为$r(A) \leqslant r(B)=r(A \vdots B)$.

【线性方程组与矩阵秩的关系】若齐次方程$BX=0$的解都是齐次方程$AX=0$的解,则它们的秩关系为$r(A) \leqslant r(B)$.

常见题型·应对

【求矩阵秩问题·解法概述】判定向量组的线性相关、线性表示、极大无关组和方程组的解的讨论等问题中都需要计算矩阵的秩的大小.主要方法如下:

(1)用初等行变换将A化做行阶梯形后,其中非零行个数即$r(A)$.

(2)使用秩的基本公式,设法证明$r \leqslant r(A) \leqslant r$,从而$r(A)=r$.

例 设A为n阶方阵,且$A^2=A$,证明

$$r(A)+r(A-I)=n.$$

证 由$A(A-I)=0$,知$r(A)+r(A-I) \leqslant n$;又由$A+(I-A)=I$知

$$\begin{aligned} n = r(I) &\leqslant r(A)+r(I-A) \\ &= r(A)+r(A-I). \end{aligned}$$

故有 $$r(A)+r(A-I)=n.$$

【求逆矩阵问题·解法概述】判定矩阵的可逆性以及计算逆矩阵的题目比较基本,可以是单独的问题,也可以是矩阵运算问题中的一个考点.必须知道以下常规的应对方法.当n阶方阵A可逆时,其求法有以下几种.

(1)设法导出$AB=I$或$BA=I$,从而$A^{-1}=B$.

(2)使用伴随矩阵,$A^{-1}=A^*/|A|$,如

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(3)采用行初等变换,$(\boldsymbol{A} \,\vdots\, \boldsymbol{I}) \to (\boldsymbol{I} \,\vdots\, \boldsymbol{A}^{-1})$.

(4)借助因式分解,若 $\boldsymbol{A}=\boldsymbol{PQ}$,则 $\boldsymbol{A}^{-1}=\boldsymbol{Q}^{-1}\boldsymbol{P}^{-1}$.

(5)借助分块矩阵,如

$$\begin{pmatrix} \boldsymbol{P} & \\ & \boldsymbol{Q} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{P}^{-1} & \\ & \boldsymbol{Q}^{-1} \end{pmatrix},$$

$$\begin{pmatrix} & \boldsymbol{P} \\ \boldsymbol{Q} & \end{pmatrix}^{-1} = \begin{pmatrix} & \boldsymbol{Q}^{-1} \\ \boldsymbol{P}^{-1} & \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{B} & \boldsymbol{D} \\ & \boldsymbol{C} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{B}^{-1} & -\boldsymbol{B}^{-1}\boldsymbol{D}\boldsymbol{C}^{-1} \\ & \boldsymbol{C}^{-1} \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{B} & \\ \boldsymbol{D} & \boldsymbol{C} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{B}^{-1} & \\ -\boldsymbol{C}^{-1}\boldsymbol{D}\boldsymbol{B}^{-1} & \boldsymbol{C}^{-1} \end{pmatrix},$$

特别地,

$$\begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & & \\ & a_2^{-1} & \\ & & a_3^{-1} \end{pmatrix},$$

$$\begin{pmatrix} & & a_1 \\ & a_2 & \\ a_3 & & \end{pmatrix}^{-1} = \begin{pmatrix} & & a_3^{-1} \\ & a_2^{-1} & \\ a_1^{-1} & & \end{pmatrix}.$$

(6) 伴随矩阵 $\boldsymbol{A}^*$ 的逆:$(\boldsymbol{A}^*)^{-1}=\boldsymbol{A}/|\boldsymbol{A}|$;

正交矩阵 A 的逆:$\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$.

【由定义确定逆矩阵】从问题给出的矩阵出发,设法变形成 $\boldsymbol{AB}=\boldsymbol{I}$ 或 $\boldsymbol{BA}=\boldsymbol{I}$ 的形式,则 $\boldsymbol{A}^{-1}=\boldsymbol{B}$.

例1　设$\boldsymbol{A},\boldsymbol{B}$ 是n 阶方阵,$\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$,证明$\boldsymbol{A}-\boldsymbol{I}$可逆,并求其逆.

证　$\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B} \Rightarrow \boldsymbol{AB}-\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{I}=\boldsymbol{I}$

$$\Rightarrow (\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{I},$$

故 $\boldsymbol{A}-\boldsymbol{I}$ 可逆，$(\boldsymbol{A}-\boldsymbol{I})^{-1}=\boldsymbol{B}-\boldsymbol{I}$.

例 2　设 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2+2\boldsymbol{A}-3\boldsymbol{I}=\boldsymbol{0}$，证明 $\boldsymbol{A}$，$\boldsymbol{A}\pm 2\boldsymbol{I}$ 以及 $\boldsymbol{A}+4\boldsymbol{I}$ 可逆.

证　由于 $\boldsymbol{A}(\boldsymbol{A}+2\boldsymbol{I})=3\boldsymbol{I}$ 及 $3\boldsymbol{I}$ 的可逆性知，$\boldsymbol{A}$ 及 $\boldsymbol{A}+2\boldsymbol{I}$ 为可逆阵. 为说明 $\boldsymbol{A}+4\boldsymbol{I}$ 可逆，先写出待定方程式

$$(\boldsymbol{A}+4\boldsymbol{I})(\boldsymbol{A}+a\boldsymbol{I})=b\boldsymbol{I},$$

展开后与题给等式比较系数得

$$(\boldsymbol{A}+4\boldsymbol{I})(\boldsymbol{A}-2\boldsymbol{I})=-5\boldsymbol{I},$$

可见 $\boldsymbol{A}+4\boldsymbol{I}$，$\boldsymbol{A}-2\boldsymbol{I}$ 可逆.

例 3　设 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^3=\boldsymbol{0}$，证明 $\boldsymbol{I}-\boldsymbol{A}$ 可逆，并求其逆.

解　由条件得

$$\boldsymbol{I}-\boldsymbol{A}^3=(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{A}^2+\boldsymbol{A}+\boldsymbol{I})=\boldsymbol{I},$$

故
$$(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{A}^2+\boldsymbol{A}+\boldsymbol{I}.$$

【分解因式法判定矩阵可逆】设法将所论矩阵写成可逆矩阵的乘积.

例 1　设 $\boldsymbol{A}$，$\boldsymbol{B}$ 是 n 阶方阵，$\boldsymbol{B}$ 与 $\boldsymbol{I}+\boldsymbol{AB}$ 可逆，证明 $\boldsymbol{I}+\boldsymbol{BA}$ 可逆.

证　$\boldsymbol{I}+\boldsymbol{BA}=\boldsymbol{B}(\boldsymbol{B}^{-1}+\boldsymbol{A})=\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{AB})\boldsymbol{B}^{-1}$.

例 2　设 $\boldsymbol{A}$，$\boldsymbol{B}$ 是 n 阶可逆阵，$\boldsymbol{A}+\boldsymbol{B}$ 可逆，证明 $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 可逆.

证
$$\begin{aligned}\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}&=\boldsymbol{A}^{-1}(\boldsymbol{I}+\boldsymbol{A}\boldsymbol{B}^{-1})\\&=\boldsymbol{A}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{B}^{-1}.\end{aligned}$$

例 3　设 $\boldsymbol{A}=\begin{pmatrix}1 & -1\\ 2 & 3\end{pmatrix}$，证明 $\boldsymbol{B}=\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{I}$ 可

逆.

证 因式分解，$B=(A-2I)(A-I)$，易知

$$|A-2I|\neq 0,|A-I|\neq 0,$$

即B是两可逆阵$A-2I,A-I$之乘积，故B可逆.

【通过分块矩阵求逆阵】

例 设$A=\begin{pmatrix}0&a_1&0&0\\0&0&a_2&0\\0&0&0&a_3\\a_4&0&0&0\end{pmatrix}$，$a_1a_2a_3a_4\neq 0$，求$A^{-1}$及$(A^*)^{-1}$.

解 视A为分块矩阵$\begin{pmatrix}\mathbf{0}&P\\Q&\mathbf{0}\end{pmatrix}$，则

$$A^{-1}=\begin{pmatrix}0&0&0&a_4^{-1}\\a_1^{-1}&0&0&0\\0&a_2^{-1}&0&0\\0&0&a_3^{-1}&0\end{pmatrix},$$

$(A^*)^{-1}=A/|A|=\dfrac{(-1)^{\tau(4123)}}{a_1a_2a_3a_4}A=A/(-a_1a_2a_3a_4)$.

【矩阵方程求解问题·解法概述】类比于数的方程问题，未知矩阵满足一个方程，要求结合矩阵的基本运算求出未知矩阵. 其常规求解方法如下：

(1)将含有未知矩阵X的方程进行化简，使之形为$AX=B$或$XA=B$. 遇到A^*，A^{T}，A^{-1}这类矩阵时，要尽可能转换到A.

(2)若A可逆，则$AX=B$的解是$X=A^{-1}B$；$XA=B$的解是$X=BA^{-1}$，A的可逆性要论证或指明.

(3)若A不可逆，则$AX=B$的求解归于线性方

程组求解问题,其中 X 的行数与 A 相同,列数与 B 相同.

【A 不可逆时 $AX=B$ 的求解法】

例 　设 $A=\begin{pmatrix}1&2&-3\\1&1&-1\end{pmatrix}$, $B=\begin{pmatrix}3&-1\\2&0\end{pmatrix}$,解方程 $AX=B$.

解 　记 $B=[B_1 \vdots B_2]$,分别求解方程组 $AX=B_1$ 和 $AX=B_2$,得到通解

$$X_1=\begin{pmatrix}1-c_1\\1+2c_1\\c_1\end{pmatrix}\quad 和\quad X_2=\begin{pmatrix}c_2\\1-2c_2\\1-c_2\end{pmatrix},$$

于是 $X=\begin{pmatrix}1-c_1&c_2\\1+2c_1&1-2c_2\\c_1&1-c_2\end{pmatrix}$, c_1,c_2 为任意实数.

【A 可逆时 $AX=B$ 的求解法】

例 1 　设 $A=\begin{pmatrix}1&1&-1\\0&1&1\\0&0&-1\end{pmatrix}$,求解方程

$$A^2-AX=I.$$

解 　方程即 　$AX=A^2-I$,

左乘 A^{-1},得 　$X=A-A^{-1}$,

故

$$X=\begin{pmatrix}0&2&1\\0&0&0\\0&0&0\end{pmatrix}.$$

例 2 　设 $A=\begin{pmatrix}2&0&1\\0&3&0\\2&0&2\end{pmatrix}$, $B=\begin{pmatrix}1&&\\&-1&\\&&0\end{pmatrix}$,且

有 $\boldsymbol{A}\boldsymbol{X}+2\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}+2\boldsymbol{X}$，计算 $\boldsymbol{X}^4$.

解　由　$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}=\boldsymbol{B}(\boldsymbol{A}-2\boldsymbol{I})$

得　$$\boldsymbol{X}=(\boldsymbol{A}-2\boldsymbol{I})^{-1}\boldsymbol{B}(\boldsymbol{A}-2\boldsymbol{I})$$

于是

$$\begin{aligned}\boldsymbol{X}^4 &= (\boldsymbol{A}-2\boldsymbol{I})^{-1}\boldsymbol{B}^4(\boldsymbol{A}-2\boldsymbol{I}) \\ &= \begin{pmatrix} & & 1 \\ & 1 & \\ 2 & & \end{pmatrix}^{-1}\begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}\begin{pmatrix} & & 1 \\ & 1 & \\ 2 & & \end{pmatrix} \\ &= \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}.\end{aligned}$$

【矩阵行和均为 a 时的常规推理】如果矩阵 $\boldsymbol{A}$ 的每一行元素之和均为 a，记 $\boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}}$，则有

$$\boldsymbol{A}\boldsymbol{e}=a\boldsymbol{e}.$$

这意味着有以下系列结果：

(1)列向量 $\boldsymbol{e}$ 是 $\boldsymbol{A}\boldsymbol{X}=a\boldsymbol{e}$ 的非零解.

(2)a 是矩阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{e}$ 是相应的特征向量.

(3)当矩阵 $\boldsymbol{A}$ 可逆时，得 $\boldsymbol{A}^{-1}\boldsymbol{e}=\dfrac{1}{a}\boldsymbol{e}$，从而逆矩阵 $\boldsymbol{A}^{-1}$ 的行和均为 $1/a$.

(4)类似地推出，矩阵 $\boldsymbol{A}^m$ 的行和均为 a^m.

【$|\boldsymbol{A}|=0$ 时的常规推理】如果 n 阶矩阵 $\boldsymbol{A}$ 的行列式满足 $|\boldsymbol{A}|=0$，则成立以下结果.

(1)$\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{0}$，$r(\boldsymbol{A})+r(\boldsymbol{A}^*)\leqslant n$.

(2)$\boldsymbol{A}^*$ 的列向量都是齐次方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{0}$ 的解；$\boldsymbol{A}$ 的列向量都是齐次方程组 $\boldsymbol{A}^*\boldsymbol{X}=\boldsymbol{0}$ 的解.

(3)$\lambda=0$ 是矩阵 $\boldsymbol{A}$ 的特征值；若还有 $r(\boldsymbol{A})=$

$n-1$,则有以下两个结果:

①$r(\boldsymbol{A}^*)=1$,从而 $\boldsymbol{A}^*$ 的列向量成比例.

②齐次方程 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间维数 $\dim N(\boldsymbol{A})=1$.

【$r(\boldsymbol{A})=1$ 的常规推理】如果 n 阶方阵 $\boldsymbol{A}$ 满足 $r(\boldsymbol{A})=1$,则有以下结论:

(1)存在两个非零列向量 α,β:$\boldsymbol{A}=\alpha\beta^{\mathrm{T}}$.

(2)$\boldsymbol{A}$ 有特征值 $\lambda=\alpha^{\mathrm{T}}\beta$ 和 0.

(3)方程 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间的维数为 $n-1$.

(4)$\boldsymbol{A}$ 的任何两行(列)成比例.

第11章 向　　量

大纲要求

理解n维向量、向量的线性组合与线性表示的概念.理解向量组线性相关、线性无关的定义,掌握向量组线性相关、线性无关的有关性质及判别法.理解向量组的极大线性无关组和向量组的秩的概念,会求向量组的极大线性无关组和秩.理解向量组等价的概念,以及矩阵的秩与其行(列)向量组的秩之间的关系.

了解n维向量空间、子空间、基底、维数、坐标的概念.了解基变换和坐标变换公式,会求过渡矩阵.

了解内积的概念,掌握线性无关向量组正交规范化的施密特方法.

注　数学二不要求上述第二段的内容.

知识点·点拨

11.1　线性相关·线性无关

【n维向量】在数域F上的n个数$a_1,a_2,\cdots,a_n$构成有序数组(也可以看做$1\times n$矩阵),记作

$$\alpha=(a_1,a_2,\cdots,a_n),$$

称之为F上的一个n维向量.其中a_i称为α的第i个分量(或坐标).称α为行向量,称$\beta=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$为列向量,称$\mathbf{0}=(0,0,\cdots,0)$为零向量.

【向量的线性表示】对一组n维向量$\alpha_1,\alpha_2,\cdots,\alpha_s$和

β,若存在一组常数$k_1,k_2,\cdots,k_s$,使得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s,$$

则称β可以由向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性表示.

例 证明向量$\beta=(2,-3,4)^{\mathrm{T}}$可以由向量组$\alpha_1=(1,1,1)^{\mathrm{T}},\alpha_2=(2,1,0)^{\mathrm{T}},\alpha_3=(3,0,0)^{\mathrm{T}}$线性表示.

证 设有线性表示式$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta$,即

$$\begin{cases} x_1+2x_2+3x_3=2, \\ x_1+x_2=-3, \\ x_1=4. \end{cases}$$

解得$x_1=4,x_2=-7,x_3=4$. 故$\beta=4\alpha_1-7\alpha_2+4\alpha_3$.

【向量的线性组合】称$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$为向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$的一个线性组合,常数$k_1,k_2,\cdots,k_s$称为线性组合的组合系数.

例 $3\alpha_1+5\alpha_2-4\alpha_3$为向量组$\alpha_1,\alpha_2,\alpha_3$的一个线性组合;$0\alpha_1+0\alpha_2+0\alpha_3$也是向量组$\alpha_1,\alpha_2,\alpha_3$的线性组合.

【线性相关·线性无关】给定一组n维向量$\alpha_1,\alpha_2,\cdots,\alpha_s$,如果存在不全为零的数$k_1,k_2,\cdots,k_s$,使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = \mathbf{0},$$

则称向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性相关;否则称之线性无关.

〖点拨〗

(1)只含一个向量α_1的向量组线性相关$\Leftrightarrow\alpha_1=\mathbf{0}$.

(2)向量组α_1,α_2线性相关$\Leftrightarrow\alpha_1$与α_2的对应分量成比例.

(3)由向量组$\alpha_1,\alpha_2,\alpha_3$中任意两个线性无关推不出该向量组线性无关,例如平面三角形中的三条边向量组.

【线性相关与线性无关的性质】

(1)如果向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性相关,则增加新的向量后依然线性相关.

(2)如果向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关,则减去其中一些向量后依然线性无关.

(3)增加线性无关向量组中向量的分量的个数不改变其无关性.

(4)减少线性相关向量组中向量的分量的个数不改变其相关性.

【判别向量组线性相关的方法归纳】设$\alpha_1,\alpha_2,\cdots,\alpha_m$是一组$n$维列向量,则在以下任何一种情形下,它们是线性相关的:

(1)矩阵$A=(\alpha_1,\alpha_2,\cdots,\alpha_m)$的秩$r(A)<m$;

(2)在$m=n$的情形,行列式$|\alpha_1,\alpha_2,\cdots,\alpha_m|=0$;

(3)矩阵$A=(\alpha_1,\alpha_2,\cdots,\alpha_m)$对应的齐次方程$AX=\mathbf{0}$有非零解;

(4)有矩阵分解式$(\alpha_1,\alpha_2,\cdots,\alpha_m)=AB$,其中$r(A)<m$或$r(B)<m$;

(5)$\alpha_1,\alpha_2,\cdots,\alpha_m$中有一个可以由其余向量线性表示;

(6)$\alpha_1,\alpha_2,\cdots,\alpha_m$中包含线性相关的子向量组;

(7)$\alpha_1,\alpha_2,\cdots,\alpha_m$中包含有零向量;

(8)$\alpha_1,\alpha_2,\cdots,\alpha_m$中每个向量可以由某个向量组

$\beta_1,\beta_2,\cdots,\beta_t(t<m)$线性表示;

(9)有非零矩阵$\boldsymbol{B}$使$(\alpha_1,\alpha_2,\cdots,\alpha_m)\boldsymbol{B}=\boldsymbol{0}$.

【判定向量组线性无关的方法归纳】在以下所列条件之一满足时,所论n维列向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关:

(1)矩阵$\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_m)$的秩$r(\boldsymbol{A})=m$;

(2)在$m=n$的情形,行列式$|\alpha_1,\alpha_2,\cdots,\alpha_m|\neq 0$;

(3)矩阵$\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_m)$对应的齐次方程$\boldsymbol{AX}=\boldsymbol{0}$只有零解$\boldsymbol{X}=\boldsymbol{0}$;

(4)存在线性无关的向量组$\beta_1,\beta_2,\cdots,\beta_m$及可逆阵$\boldsymbol{B}$,使$(\alpha_1,\alpha_2,\cdots,\alpha_m)=(\beta_1,\beta_2,\cdots,\beta_m)\boldsymbol{B}$;

(5)$\alpha_1,\alpha_2,\cdots,\alpha_m$中任何一个向量都不能由其余向量线性表示;

(6)有非零向量β,它能由$\alpha_1,\alpha_2,\cdots,\alpha_m$线性表示,且表示式惟一.

【确定向量的线性表示的方法归纳】以下条件之一成立时可以推出β能由$\alpha_1,\alpha_2,\cdots,\alpha_m$线性表示:

(1) $r(\alpha_1,\alpha_2,\cdots,\alpha_m,\beta)=r(\alpha_1,\alpha_2,\cdots,\alpha_m)$;

(2)方程$(\alpha_1,\alpha_2,\cdots,\alpha_m)\boldsymbol{X}=\beta$有解;

(3)$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关,而向量组$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性相关,此时线性表示的形式惟一.

【极大线性无关组·向量组的秩】向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$中的一个线性无关的子组$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$称为极大线性无关组,如果该子组再增加一个向量组中的成员便线性相关.此时称向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$的秩为$r$,记作$r(\alpha_1,\alpha_2,\cdots,\alpha_m)=r$.

〖点拨〗一个向量组的秩是惟一的.但是它的极大线性无关组并不惟一,除非该向量组自身是线性无关的.

【极大无关组求法】采用矩阵法求解.将所给向量组逐行排成一个矩阵,然后做初等行变换,得到的行阶梯形中非零行对应的向量便构成一个极大无关组.

例 设有行向量组

$$\alpha_1=(1,1,2,2),\quad \alpha_2=(2,5,3,4),$$
$$\alpha_3=(0,3,2,3),\quad \alpha_4=(2,2,1,1),$$

求出一个极大无关组及该向量组的秩.

解 将向量组按行排成矩阵,采用行初等变换化作行阶梯形(第五列是为了同步记录向量变动的过程):

$$\boldsymbol{A}=\left(\begin{array}{cccc:c}1&1&2&2&\alpha_1\\2&5&3&4&\alpha_2\\0&3&2&3&\alpha_3\\2&2&1&1&\alpha_4\end{array}\right)$$
$$\rightarrow\left(\begin{array}{cccc:c}1&1&2&2&\alpha_1\\0&3&-1&0&\alpha_2-2\alpha_1\\0&0&3&3&\alpha_3-\alpha_2+2\alpha_1\\0&0&0&0&\alpha_4-\alpha_2+\alpha_3\end{array}\right).$$

由以上结果得:①因为虚线左边的矩阵的秩是3,故$r(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=3$;②观察最后一行,得$\alpha_4=\alpha_2-\alpha_3$,从而$\alpha_1,\alpha_2,\alpha_3$构成一个极大无关组.

【向量组的线性表示·定义】设有两个向量组

$$(\text{Ⅰ})\alpha_1,\alpha_2,\cdots,\alpha_s,\quad (\text{Ⅱ})\beta_1,\beta_2,\cdots,\beta_t,$$

如果向量组(Ⅰ)中的每个向量都可由向量组(Ⅱ)线性表示,则称向量组(Ⅰ)可由向量组(Ⅱ)线性表示.

【向量组的等价·定义】如果向量组(Ⅰ)和向量组(Ⅱ)可以互相线性表示,则称这两个向量组等价,记作(Ⅰ)$\cong$(Ⅱ).

【极大线性无关组的等价性】向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$与它的极大线性无关组等价,从而它的任意两个极大线性无关组等价.

【向量组等价关系的性质】向量组的等价关系具有以下三个性质:

(1)自反性,即(Ⅰ)$\cong$(Ⅰ);

(2)对称性,若(Ⅰ)$\cong$(Ⅱ),则(Ⅱ)$\cong$(Ⅰ);

(3)传递性,若(Ⅰ)$\cong$(Ⅱ),(Ⅱ)$\cong$(Ⅲ),则(Ⅰ)$\cong$(Ⅲ).

上面(Ⅰ),(Ⅱ),(Ⅲ)均表示向量组,下同.

【向量组等价与矩阵等价的区别】等价的向量组必然有相等的秩;但是反过来不对,例如,三维向量空间中的向量(1,0,0)和(0,1,0)分别作为含一个向量的向量组,秩都是1,但是它们不能互相线性表示,因此不等价.而对于矩阵来说,两个同型矩阵等价的充分必要条件便是它们的秩相等.

【向量组的线性表示和等价的判定方法】如果记$\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_s)$,$\boldsymbol{B}=(\beta_1,\beta_2,\cdots,\beta_t)$,则

(1)向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$可由向量组$\beta_1,\beta_2,\cdots,\beta_t$线性表示$\Leftrightarrow$存在矩阵$\boldsymbol{K}_{t\times s}$,使得$\boldsymbol{A}=\boldsymbol{B}\boldsymbol{K}_{t\times s}$

$\Leftrightarrow r(\boldsymbol{A} \mid \boldsymbol{B})=r(\boldsymbol{A})$;

(2)向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$与向量组$\beta_1,\beta_2,\cdots,\beta_t$等价$\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{A}\ \vdots\ \boldsymbol{B})=r(\boldsymbol{B})$.

【向量组的线性表示与线性相关】若向量组(Ⅰ)可由向量个数较少的向量组(Ⅱ)线性表示,则向量组(Ⅰ)线性相关.

〖推论〗若线性无关的向量组$\beta_1,\beta_2,\cdots,\beta_t$可由向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性表示,则$t\leqslant s$.

【向量组的线性表示与向量组的秩】设向量组(Ⅰ)可由向量组(Ⅱ)线性表示,则r(Ⅰ)$\leqslant r$(Ⅱ).

【向量组的秩与矩阵的秩·三秩相等定理】矩阵A的行向量组的秩等于矩阵A的列向量组的秩,也等于矩阵A的秩.

〖点拨〗常常利用这个定理把讨论向量组的秩的问题转化成讨论相应的矩阵的秩的问题(后者有较多的公式可用).

11.2　向量空间·坐标·基变换

【向量空间】设V是数域F上的n维向量构成的非空集合,且满足

(1)若$\alpha,\beta\in V$,则$\alpha+\beta\in V$,

(2)若$\alpha\in V,k\in F$,则$k\alpha\in V$,

便称V为数域F上的向量空间.

例　仅有零向量的集合构成向量空间,称为零空间.齐次线性方程组的解向量全体构成一个向量空间.

【子空间】设U和V都是向量空间,且$U\subseteq V$,则称U

是 V 的子空间.

【生成子空间】由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的线性组合构成的集合是一个向量空间,记作

$$L\{\alpha_1,\alpha_2,\cdots,\alpha_r\}=\{k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r \mid k_i\in\mathbf{R}\},$$

称 L 为由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 生成的向量空间.

【基·维数】设 V 是向量空间,若向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n\in V$ 满足以下两条:

(1) $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关;

(2) V 中的任一向量都可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示,则称 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为空间 V 的一组基,n 称为 V 的维数,记作 $\dim V=n$,并称 V 是 n 维向量空间.

例　$\mathbf{R}^n$ 中的一组基是

$$\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\quad \boldsymbol{e}_2=(0,1,0,\cdots,0)^{\mathrm{T}},\cdots,$$
$$\boldsymbol{e}_n=(0,0,\cdots,0,1)^{\mathrm{T}},$$

因为对任意 $\alpha=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$,有

$$\alpha=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n,$$

故
$$\dim\mathbf{R}^n=n.$$

〖点拨〗若把向量空间 V 看做向量组,则 V 的基就是向量组 V 中的极大线性无关组,V 的维数就是向量组 V 的秩.

【基的判定】n 维向量空间 V 的任意 n 个线性无关的向量都可作为向量空间 V 的基.

【向量的坐标】设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 n 维空间 V 的一组基,则空间 V 中任一向量 α 可惟一地表示为

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,$$

称有序数组 $x_1,x_2,\cdots,x_n$ 为向量 α 关于基 $\alpha_1,\alpha_2,\cdots,$

α_n 的坐标,在基不变的空间中,人们常用 α 的坐标向量 $X=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 来代表向量 α.

【基变换·过渡矩阵】n 维空间 V 的任意两组基 α_1, $\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 是彼此等价的:

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots\alpha_n)C\quad(\text{基变换公式}),$$

表示式中的矩阵 C 称为从 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到 $\beta_1,\beta_2,\cdots$, β_n 的过渡矩阵.

【坐标变换公式】设向量 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标是 $X=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,同时在基 $\beta_1,\beta_2,\cdots,\beta_n$ 下的坐标是 $Y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$,$C$ 为从 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵,则有 $X=CY$.

〖点拨〗可以通过矩阵运算来记忆:

$$\alpha=(\beta_1,\beta_2,\cdots,\beta_n)Y=(\alpha_1,\alpha_2,\cdots,\alpha_n)CY,$$

又　$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)X$,对比即知 $X=CY$.

11.3　内积·正交·标准正交基

【向量的内积】定义 $\mathbf{R}^n$ 中向量 $\alpha=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 与 $\beta=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$ 的内积为实数(记作(α,β)),

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n=\alpha^{\mathrm{T}}\beta.$$

〖点拨〗$n=3$ 时,$(\alpha,\beta)=a_1b_1+a_2b_2+a_3b_3$,与向量代数中的内积(即点乘)的定义是一致的.

【内积的基本性质】

(1)交换性:$(\alpha,\beta)=(\beta,\alpha)$.

(2)线性性:$(\alpha_1+\alpha_2,\beta)=(\alpha_1,\beta)+(\alpha_2,\beta)$,

$(k\alpha,\beta)=k(\alpha,\beta)$,其中 k 为实数.

(3) 非负性:$(\alpha,\alpha)=\sum\limits_{i=1}^{n}a_i^2\geqslant 0$ 且$(\alpha,\alpha)=0$ 的

充要条件是 $\alpha=\mathbf{0}$.

(4)柯西不等式:$|(\alpha,\beta)|\leqslant|\alpha|\,|\beta|$.

(5)三角不等式:$|\alpha+\beta|\leqslant|\alpha|+|\beta|$.

【向量的长度】称非负实数 $\sqrt{(\alpha,\alpha)}=\sqrt{\sum_{i=1}^{n}a_i^2}$ 为向量 α 的长度,记作 $|\alpha|$.

【单位向量】称长度等于 1 的向量为单位向量.

【向量的单位化】如果 α 是非零向量,则 $\alpha^0=\alpha/|\alpha|$ 是单位向量,称 α^0 为 α 的单位向量,也记为 $\boldsymbol{e}_\alpha$.

【向量的夹角】设 α,β 为 $\mathbf{R}^n$ 中的向量. 定义 $\mathbf{R}^n$ 中两向量 α,β 夹角的余弦为

$$\cos\theta=\frac{(\alpha,\beta)}{|\alpha|\,|\beta|},$$

于是向量的夹角为

$$\theta=\arccos\frac{(\alpha,\beta)}{|\alpha|\,|\beta|}.$$

【向量的正交】设 $\alpha,\beta\in\mathbf{R}^n$,若有 $(\alpha,\beta)=0$,则称向量 α 与 β 是正交的.

【正交向量组 · 标准正交向量组】若一个向量组中的向量两两正交,则称这个向量组为正交向量组;若正交向量组中每一个向量还是单位向量,则称此正交向量组为标准正交向量组.

【标准正交向量组判定】向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是标准正交向量组的充要条件是

$$(\alpha_i,\alpha_j)=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

【正交向量组与线性无关向量组】若正交向量组 α_1,

$\alpha_2,\cdots,\alpha_r$ 中不含零向量,则 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 为线性无关向量组.标准正交向量组必定为线性无关向量组.

【标准正交基】在欧几里德空间 $\mathbf{R}^n$ 中,若一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 满足标准正交向量组的条件,则称之为标准正交基.

【正交矩阵】若 C 满足

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{I},$$

则称 n 阶实矩阵 C 为正交矩阵.

【正交矩阵的向量组】矩阵 C 是正交矩阵 $\Leftrightarrow C$ 的列(行)向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是标准正交向量组.

【正交变换】设 C 为 n 阶正交矩阵,X 和 Y 是欧几里德空间 $\mathbf{R}^n$ 中的 n 维向量,则线性变换 $X=CY$ 是 $\mathbf{R}^n$ 上的正交变换.

【正交变换的性质】设 $X=CY$ 是欧几里德空间 $\mathbf{R}^n$ 上的线性变换,则下列命题等价:

(1)线性变换 $X=CY$ 为正交变换.

(2)在线性变换 $X=CY$ 下,向量的内积不变,即

$$\boldsymbol{X}_1=\boldsymbol{C}\boldsymbol{Y}_1,\boldsymbol{X}_2=\boldsymbol{C}\boldsymbol{Y}_2$$

时,$(\boldsymbol{X}_1,\boldsymbol{X}_2)=(\boldsymbol{Y}_1,\boldsymbol{Y}_2)$.

(3)线性变换 $X=CY$ 把 $\mathbf{R}^n$ 中的标准正交基变成标准正交基.

〖推论〗正交变换不改变向量的内积,也就是不改变向量的长度,从而不改变曲线或曲面的图形形状.一些试题据此来设计二次曲面的类型的识别问题.

【化作正交规范组的施密特方法】以三个向量为例

说明. 设 $\alpha_1,\alpha_2,\alpha_3$ 是线性无关向量组,首先产生与 $\alpha_1,\alpha_2,\alpha_3$ 等价的正交向量组

$$\beta_1=\alpha_1,\quad \beta_2=\alpha_2-\frac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}\beta_1,$$

$$\beta_3=\alpha_3-\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)}\beta_1-\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}\beta_2$$

再对 β_1,β_2,β_3 单位化即成正交规范组:

$$\gamma_1=\frac{1}{|\beta_1|}\beta_1,\quad \gamma_2=\frac{1}{|\beta_2|}\beta_2,\quad \gamma_3=\frac{1}{|\beta_3|}\beta_3.$$

〖点拨〗如果考虑的向量是实对称矩阵的分属不同的特征值的特征向量,则它们已经是正交的了,只需要单位化便可以.

常见题型·应对

针对问题的特点使用前面归纳的方法.

【判别向量组线性相关的问题】

例 判断向量组

$$\alpha_1=(1,1,0)^{\mathrm T},\quad \alpha_2=(0,-1,2)^{\mathrm T},$$
$$\alpha_3=(2,5,-6)^{\mathrm T}$$

是否线性相关.

解 求列向量矩阵的行阶梯形(箭头表示初等变换):

$$\boldsymbol{A}=(\alpha_1,\alpha_2,\alpha_3)$$
$$=\begin{pmatrix}1&0&2\\1&-1&5\\0&2&-6\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&-1&3\\0&0&0\end{pmatrix},$$

得到 $r(A)=2<3$,所以向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性相关.

〖点拨〗此题也可以考查行列式 $|A|$ 是否为零.

但是求矩阵秩的方法具有一般性且易于实现.

【判定向量组线性无关的问题】

例1　已知向量组$\alpha_1,\alpha_2,\alpha_3$线性无关,证明向量组$\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1$也线性无关.

解　考虑矩阵表示

$$\begin{aligned}\boldsymbol{A}&=(\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1)\\&=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix},\end{aligned}$$

等式右边两个矩阵的秩都是3,故$r(\boldsymbol{A})=3$,所证向量组线性无关.

例2　设$\boldsymbol{A}$是$n\times m$矩阵,$\boldsymbol{B}$是$m\times n$矩阵,且$n<m$.若$\boldsymbol{AB}=\boldsymbol{I}$,证明$\boldsymbol{B}$的列向量组线性无关.

证　欲证$r(\boldsymbol{B})=n$.一方面$r(\boldsymbol{B})\leqslant\min\{m,n\}=n$,另一方面$r(\boldsymbol{B})\geqslant r(\boldsymbol{AB})=r(\boldsymbol{I}_n)=n$,所以$r(\boldsymbol{B})=n$,即$\boldsymbol{B}$的列向量组的秩为$n$.而$\boldsymbol{B}$的列向量组恰有$n$个列向量,故$\boldsymbol{B}$的列向量组线性无关.

【确定向量的线性表示的问题】

例1　已知n维向量组$\alpha_1,\alpha_2,\alpha_3$线性无关,向量组$\alpha_2,\alpha_3,\alpha_4$线性相关.问:

(1)α_4能否由$\alpha_1,\alpha_2,\alpha_3$线性表示?为什么?

(2)α_1能否由$\alpha_2,\alpha_3,\alpha_4$线性表示?为什么?

解　(1)因$\alpha_1,\alpha_2,\alpha_3$线性无关,故$\alpha_2,\alpha_3$也线性无关;又$\alpha_2,\alpha_3,\alpha_4$线性相关,故$\alpha_4$可以由$\alpha_2,\alpha_3$线性表示,从而更可以由大组$\alpha_1,\alpha_2,\alpha_3$线性表示.

(2)假设α_1能由$\alpha_2,\alpha_3,\alpha_4$线性表示,由于$\alpha_4$可

以由α_2,α_3线性表示,故α_1能被α_2,α_3线性表示.这与已知$\alpha_1,\alpha_2,\alpha_3$线性无关相矛盾,故$\alpha_1$不能由向量组$\alpha_2,\alpha_3,\alpha_4$线性表示.

〖点拨〗注意认识反证法的特殊作用.

例2　设矩阵$\begin{pmatrix} 1 & & & \\ & 2 & & \\ & & t & 3 \\ & & 0 & 0 \end{pmatrix}$的列向量依次为$\alpha_1,\alpha_2,\alpha_3,\beta$,问:$t$为何值时,$\beta$不能由$\alpha_1,\alpha_2,\alpha_3$线性表示?

解　$t=0$时β不能由$\alpha_1,\alpha_2,\alpha_3$线性表示,因为

$$3=r(\alpha_1,\alpha_2,\alpha_3,\beta)\neq r(\alpha_1,\alpha_2,\alpha_3)=2.$$

【讨论向量组的秩】

例1　设n维向量组$\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$满足

子组(Ⅰ)$\alpha_1,\alpha_2,\alpha_3$的秩为$r(\text{Ⅰ})=3$;

子组(Ⅱ)$\alpha_1,\alpha_2,\alpha_3,\alpha_4$的秩为$r(\text{Ⅱ})=3$;

子组(Ⅲ)$\alpha_1,\alpha_2,\alpha_3,\alpha_5$的秩为$r(\text{Ⅲ})=4$.

求向量组(Ⅳ)$\alpha_1,\alpha_2,\alpha_3,\alpha_4+\alpha_5$的秩.

解1　基于分类和反证法.由于向量组(Ⅳ)中包含了线性无关组(Ⅰ),故$r(\text{Ⅳ})=3$或4.如果$r(\text{Ⅳ})=3$,则$\alpha_4+\alpha_5$必定可以由$\alpha_1,\alpha_2,\alpha_3$线性表示.同理,对比向量组(Ⅱ)知,$\alpha_4$可由$\alpha_1,\alpha_2,\alpha_3$线性表示,从而$\alpha_5$可由$\alpha_1,\alpha_2,\alpha_3$线性表示,这将与$r(\text{Ⅲ})=4$矛盾.故$r(\text{Ⅳ})=4$.

解2　基于初等变换不改变矩阵的秩.设

$$\boldsymbol{A}=(\alpha_1,\alpha_2,\alpha_3,\alpha_4+\alpha_5),$$

$$B=(\alpha_1,\alpha_2,\alpha_3,\alpha_5).$$

由已知得$r(B)=r(\alpha_1,\alpha_2,\alpha_3,\alpha_5)=4$. 由条件推知$\alpha_4$能由$\alpha_1,\alpha_2,\alpha_3$线性表示,故可设

$$\alpha_4=k_1\alpha_1+k_2\alpha_2+k_3\alpha_3,$$

从而可以通过初等变换将α_4除掉:

$$A=(\alpha_1,\alpha_2,\alpha_3,\alpha_4+\alpha_5)\to(\alpha_1,\alpha_2,\alpha_3,\alpha_5)=B.$$

$$-k_1\quad -k_2\quad -k_3$$

故

$$r(\text{Ⅳ})=r(\text{Ⅲ})=4.$$

例2　已知列向量组$\alpha_1,\alpha_2,\alpha_3$,以及向量组

$$\beta_1=\alpha_1,\ \beta_2=\alpha_1+2\alpha_2,\ \beta_3=\alpha_1+3\alpha_2,$$

证明

$$r(\alpha_1,\alpha_2,\alpha_3)=r(\beta_1,\beta_2,\beta_3).$$

证　由于　$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)A$,

其中

$$A=\begin{pmatrix}1&1&1\\0&2&0\\0&0&3\end{pmatrix}$$

是可逆矩阵,故

$$(\alpha_1,\alpha_2,\alpha_3)=(\beta_1,\beta_2,\beta_3)A^{-1},$$

说明两个向量组等价,从而秩相等.

【判定向量组的线性表示或等价问题】转换为矩阵的秩问题讨论.

第 12 章　线性方程组

大纲要求

会用克莱姆法则. 理解齐次线性方程组有非零解的充分必要条件及非齐次线性方程组有解的充分必要条件. 理解齐次线性方程组的基础解系、通解及解空间的概念，掌握齐次线性方程组基础解系和通解的求法. 理解非齐次线性方程组的解的结构以及通解的概念. 掌握用初等行变换求解线性方程组的方法.

注　数学二不要求解空间概念.

知识点·点拨

【克莱姆法则】 若线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

的系数行列式 $D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$，则该方程组有惟一的解

$$x_j = \frac{D_j}{D}, \quad j = 1, 2, \cdots, n.$$

其中，D_j 是将系数行列式 D 中的第 j 列替代为列向

量$(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$后所得的同阶行列式.

〖点拨〗应当注意的是,用克莱姆法则求解线性方程组必须满足条件$D\neq 0$.在其他情形,方程组或者无解,或者解不惟一.

【线性方程组的基本问题】

(1)线性方程组有解的充要条件是什么?

(2)当线性方程组有解时,它有多少个解?如何求解?

(3)当线性方程组解不惟一时,这些解之间有什么关系?

【线性方程组的行初等变换】

(1)互换两个方程的位置.

(2)用一个非零数乘某一个方程.

(3)用一数k乘某一方程后加到另一个方程上.

【高斯消元法·导出组】利用初等行变换把原方程组化为行阶梯形式的方程组(称为原方程组的导出组),再用回代方法解出$x_1,x_2,\cdots,x_n$.

【增广矩阵】方程组的系数矩阵和常数列写成的矩阵称为方程组的增广矩阵.

例　方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases}$$

用矩阵形式表示即为$\boldsymbol{AX}=\boldsymbol{b}$,其增广矩阵为

$$\widetilde{A}=(A\ \vdots\ b)=\left(\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

【方程组的解向量】满足方程 $AX=b$ 的列向量 X 称为方程组的一组解.

【高斯消元法求解线性方程组】如果能使用初等行变换法把增广矩阵 $\widetilde{A}$ 化为如下阶梯形:

$$\widetilde{A}\rightarrow\left(\begin{array}{cccccc:c} c_{11} & c_{12} & \cdots & c_{1r} & \cdots & c_{1n} & d_1 \\ & c_{22} & \cdots & c_{2r} & \cdots & c_{2n} & d_2 \\ & & \ddots & \vdots & & \vdots & \vdots \\ & & & c_{rr} & \cdots & c_{rn} & d_r \\ 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1} \\ \vdots & & \vdots & \vdots & & \vdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right),$$

(不失一般性,假定 $c_{11},c_{22},\cdots,c_{rr}$ 都不为零),则原方程组与以下简化的线性方程组

$$\begin{cases} c_{11}x_1+c_{12}x_2 \qquad +\cdots+c_{1n}x_n=d_1, \\ c_{22}x_2+c_{23}x_3 \qquad +\cdots+c_{2n}x_n=d_2, \\ \qquad\qquad\qquad\qquad\vdots \\ c_{rr}x_r+c_{r\,r+1}x_{r+1}+\cdots+c_{rn}x_n=d_r, \\ \qquad\qquad\qquad\qquad 0=d_{r+1} \end{cases}$$

同解,于是:

(1)当 $d_{r+1}\neq 0$ 时,原方程组无解,相当于 $r(A)\neq r(\widetilde{A})$;

(2)当$d_{r+1}=0$时,原方程组有解,相当于$r(A)=r(\widetilde{A})$,若未知数个数n大于$r(A)=r$,则方程组中有$n-r$个自由变量,可以求得无穷多组解;当$n=r$时,只有惟一解.

【非齐次线性方程组的解】设非齐次线性方程组$AX=\boldsymbol{b}$的系数矩阵$A_{m\times n}$的秩为r,$\widetilde{A}=(A \mid \boldsymbol{b})$为增广矩阵,则

(1)$r(A)\neq r(\widetilde{A})\Leftrightarrow$方程组无解;

(2)$r(A)=r(\widetilde{A})=n\Leftrightarrow$方程组有惟一解;

(3)$r(A)=r(\widetilde{A})=r<n\Leftrightarrow$方程组有无穷多组解.

〖点拨〗对于秩等于行数的矩阵$A_{m\times n}$,由于$r(A)=r(\widetilde{A})=m$,故对任何$\boldsymbol{b}$,$AX=\boldsymbol{b}$一定有解.

【齐次线性方程组的解】设齐次线性方程组$AX=\boldsymbol{0}$的系数矩阵$A_{m\times n}$的秩为r,则

(1)$r=n\Leftrightarrow$方程组有惟一零解;

(2)$r<n\Leftrightarrow$方程组有无穷多组解.

〖点拨〗齐次线性方程组至少有一个零解,通常关注的是它是否有非零解.

【$AX=\boldsymbol{0}$有非零解的判定】在以下情形下,$AX=\boldsymbol{0}$均有非零解:

(1)方程个数m少于未知量个数n;

(2)A为方阵时,系数行列式$|A|=0$;

(3)A 的列向量组线性相关.

【线性方程组的解的叠加性质】

(1)齐次线性方程组$AX=\mathbf{0}$的任意两个解X_1与X_2的线性组合$k_1X_1+k_2X_2$还是该方程组的解;

(2)非齐次线性方程组$AX=\boldsymbol{b}$的任意两个解X_1与X_2的差是相应的齐次方程组$AX=\mathbf{0}$的解;

(3)线性方程组$AX=\boldsymbol{b}_1$的解X_1与$AX=\boldsymbol{b}_2$的解X_2的和X_1+X_2是线性方程组$AX=\boldsymbol{b}_1+\boldsymbol{b}_2$的解.

【齐次线性方程组的通解结构】齐次线性方程组的解X都可表示为基础解系$\alpha_1,\alpha_2,\cdots,\alpha_s$的线性组合,即通解为$X=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$.

【非齐次线性方程组通解结构】设非齐次线性方程组$AX=\boldsymbol{b}$有解,则其任一解(通解)为

$$X=X_c+\eta,$$

其中,η是$AX=\boldsymbol{b}$的一个特解,X_c是对应的齐次线性方程组$AX=\mathbf{0}$的通解.

【解空间·基础解系】齐次线性方程组$AX=\mathbf{0}$解的全体构成一个向量空间,称为解空间,记为

$$N(A)=\{X\,|\,AX=\mathbf{0}\}.$$

称$N(A)$的基$\alpha_1,\alpha_2,\cdots,\alpha_s$为该方程组的基础解系.

〖点拨1〗方程组的基础解系是解空间的一个极大无关组,它的数目是确定的,但是其成员不是惟一的.

〖点拨2〗直接验证可知,非齐次线性方程组$AX=\boldsymbol{b}$的解向量全体不构成向量空间.

例　设$\alpha_1,\alpha_2,\alpha_3$是齐次方程$AX=\mathbf{0}$的基础解系,$\beta$是非齐次方程$AX=\boldsymbol{b}$的解,则直接由定义可以

证明 $\alpha_1,\alpha_2,\alpha_3,\beta$ 是线性无关的向量组，再由初等变换法可以看出 $\alpha_1+\beta,\alpha_2+\beta,\alpha_3+\beta,\beta$ 是非齐次方程 $\boldsymbol{AX}=\boldsymbol{b}$ 的线性无关的解向量组.

【齐次解空间的维数】设 $m\times n$ 矩阵 $\boldsymbol{A}$ 的秩 $r(\boldsymbol{A})=r$，则齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间 $N(\boldsymbol{A})$ 的维数 $\dim N(\boldsymbol{A})=n-r$.

常见题型·应对

【根据行初等变换法求解线性方程组】求给出具体数据的线性方程组的通解或相应的基础解系是基本题型，例如在对角化问题中的特征向量计算. 应对策略便是运用行初等变换法将增广矩阵化作行阶梯形后，变作较简单的方程组求解.

【利用解的结构计算线性方程组的通解】分别计算齐次方程组的基础解系，非齐次方程组的特解，然后依据通解结构来进行叠加.

例　已知方阵 $\boldsymbol{A}=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ 中的列向量 $\alpha_2,\alpha_3,\alpha_4$ 线性无关，$\alpha_1=2\alpha_2-\alpha_3$. 如果 $\beta=\alpha_1+\alpha_2+\alpha_3+\alpha_4$，求解线性方程组 $\boldsymbol{AX}=\beta$.

解　首先看出 $\boldsymbol{e}=(1,1,1,1)^{\mathrm{T}}$ 是所给非齐次方程组 $\boldsymbol{AX}=\beta$ 的一个解；而从条件 $\alpha_1=2\alpha_2-\alpha_3$ 知 $\alpha_1-2\alpha_2+\alpha_3+0\alpha_4=\boldsymbol{0}$，说明 $\boldsymbol{a}=(1,-2,1,0)^{\mathrm{T}}$ 是对应的齐次方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个非零解；再结合 $\alpha_2,\alpha_3,\alpha_4$ 线性无关知道，矩阵 $\boldsymbol{A}$ 的秩为 3，故齐次方程组解空间的维数是 1. 由解的结构可知，取 $\boldsymbol{a}$ 为基础解系，便得到所求方程的通解：$k\boldsymbol{a}+\boldsymbol{e}$.

【含参数的线性方程组讨论题】运用行初等变换法

讨论含有参数的非齐次线性方程组的解的存在问题和惟一问题,在有解时要求写出通解.求解时需要使用行初等变换法和矩阵的秩概念.

【含参数的线性方程组计算题】有时为了减少计算量,可以只要求从解的某一种状态中确定参数的值.这类题涉及行初等变换和解的结构知识,往往是单独的试题.解法同上.

例　设 B 是一个三阶非零矩阵,它的每一列是齐次方程组

$$\begin{cases} x_1 + 2x_2 - 2x_3 = 0, \\ 2x_1 - x_2 + \lambda x_3 = 0, \\ 3x_1 + x_2 - x_3 = 0 \end{cases}$$

的解,求 λ 的值和 $|\boldsymbol{B}|$.

解　因为非零矩阵 B 的列是上面齐次方程组的解,故该齐次方程组有非零解,从而系数行列式

$$|\boldsymbol{A}| = \begin{vmatrix} 1 & 2 & -2 \\ 2 & -1 & \lambda \\ 3 & 1 & -1 \end{vmatrix} = 0,$$

得 $5\lambda - 5 = 0$,即 $\lambda = 1$.

又因矩阵 A 的秩为 2,故 $\dim N(\boldsymbol{A}) = 3 - 2 = 1$,即基础解系只有一个解向量,因而 B 的三个列向量必线性相关,得 $|\boldsymbol{B}| = 0$.

【两个方程组的公共解计算题】给出方程组(Ⅰ)和方程组(Ⅱ),求它们的公共解.

〖解法 1〗首先求解一个方程组,将所得惟一解或通解代入到另一个方程组中,验证是否矛盾,不矛

盾或需要对参数加以限制时便得到公共解.

〖解法 2〗将两个方程组联立,然后求解.

【两个方程组的同解判定问题】若两个线性方程组同解,则有

(1)它们的导出组的解空间维数一样(在有解情形,这表现为系数矩阵的秩相同);

(2)解向量互相满足对方的方程组.

两个线性方程组同解的充分条件为以下条件之一:

(1)两个方程组的增广矩阵的行向量组是等价向量组;

(2)联立之后的大方程组与每一个方程组的解相同.

【几何学应用问题】将直线、平面关系用向量关系和方程组来表示的问题.

例　借助解空间维数可知,三个平面

$$a_1x+b_1y+c_1z=d_1,$$
$$a_2x+b_2y+c_2z=d_2,$$
$$a_3x+b_3y+c_3z=d_3$$

交于一条直线的充分必要条件是上述三个方程联立的非齐次方程组 $AX=\boldsymbol{b}$ 满足 $r(A)=r(\tilde{A})=2$.

【综合转换问题】结合矩阵的秩、解空间的维数、伴随矩阵性质、向量的线性相关、特征向量的定义的综合问题.应对的关键是掌握用不同的形式来表达相同的命题.

例如,方程 $AX=Y$ 有解等价于向量 Y 可以由矩

阵 A 的列向量组线性表示；$AX=0$ 有非零解 α，等价于非零向量 α 与矩阵 A 的每个行向量正交.

例　设 A 为 $m\times n$ 矩阵，B 为 $n\times k$ 矩阵，若 $AB=0$，证明 $r(A)+r(B)\leqslant n$.

证　将 B 按列分块为 $B=(B_1\ B_2\cdots\ B_k)$，由 $AB=0$，得

$$(AB_1\ AB_2\ \cdots\ AB_k)=(0\ 0\ \cdots\ 0),$$

即
$$AB_i=0,\quad i=1,2,\cdots,k.$$

上式表明，矩阵 B 的每一列是齐次线性方程组 $AX=0$ 的解，即 $B_1,B_2,\cdots,B_k\in N(A)$，从而

$$r(B_1\ B_2\ \cdots\ B_k)\leqslant \dim N(A)=n-r(A),$$

即
$$r(B)\leqslant n-r(A)\text{ 或 }r(A)+r(B)\leqslant n.$$

【由齐次方程组的基础解系求该方程组的系数矩阵问题】利用正交的概念建立方程.

例　求一个齐次线性方程组，使得

$$\beta_1=(1,2,-1,4)^{\mathrm{T}},\quad \beta_2=(1,0,-4,0)^{\mathrm{T}}$$

为其基础解系.

解　设所求方程为

$$a_1x_1+a_2x_2+a_3x_3+a_4x_4=0,$$

代入给定的解向量，得到两个方程

$$\begin{cases}a_1+2a_2-a_3+a_4=0,\\ a_1-4a_3=0,\end{cases}$$

解出的基础解系为

$$\alpha_1=(8,-3,2,0)^{\mathrm{T}},\quad \alpha_2=(0,2,0,-1)^{\mathrm{T}}.$$

于是所求齐次线性方程组如下：

$$\begin{cases}8x_1-3x_2+2x_3=0,\\ 2x_2-x_4=0.\end{cases}$$

第13章　矩阵的特征值和特征向量

大纲要求

理解矩阵的特征值和特征向量的概念及性质，会求矩阵的特征值和特征向量.理解相似矩阵的概念、性质及矩阵可相似对角化的充分必要条件.掌握将矩阵化为相似对角矩阵的方法.掌握实对称矩阵的特征值和特征向量的性质.

知识点·点拨

13.1　特征值和特征向量

【特征值·特征向量】设方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$**，若有数** λ **和** n **维非零向量** ξ**，使** $A\xi=\lambda\xi$ **成立，则称数** λ **为** A **的特征值，称向量** ξ **为矩阵** A **的对应于** λ **的特征向量.**

〖点拨〗对应于 λ **的特征向量是齐次线性方程** $(\lambda\boldsymbol{I}-A)\boldsymbol{X}=\boldsymbol{0}$ **的非零解，该方程有非零解的充分必要条件是** $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$.

【特征多项式·特征方程·特征值】称多项式

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$$

$$=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1\,n} \\ -a_{n1} & \cdots & -a_{n\,n-1} & \lambda-a_{nn} \end{vmatrix}$$

$$=\lambda^n+b_1\lambda^{n-1}+\cdots+b_{n-1}\lambda+b_n$$

为矩阵 A 的特征多项式. 称 $|\lambda \boldsymbol{I}-\boldsymbol{A}|=0$ 为特征方程. 特征方程的根 λ 便为 A 的特征值.

【特征多项式的系数】 设 n 阶方阵 $A=(a_{ij})_{n\times n}$ 的 n 个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,则有

(1) $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$;

(2) $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$.

〖推论〗方阵 A 可逆 $\Leftrightarrow A$ 的特征值都非零.

【特征向量的计算方法】 计算方法有两种:

(1) 依据特征向量的定义,使 $A\xi=\lambda\xi$;

(2) 求齐次线性方程组 $(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的非零解.

【特征值计算方法】 计算方法有三种:

(1) 依据特征值定义,若有非零向量 ξ,使 $A\xi=\lambda\xi$,则 λ 是 A 的一个特征值;

(2) 依据特征方程 $|\boldsymbol{A}-\lambda \boldsymbol{I}|=0$ 求解;

(3) 根据特征值的传递性质求解.

【特征值的传递性质】 设 λ 是 A 的特征值,α 是 A 的对应于特征值 λ 的特征向量. $h(\lambda)$ 是代数多项式,则 $\boldsymbol{A}^{\mathrm{T}},\boldsymbol{A}^{-1}$ 等的特征值及特征向量如下表:

矩　阵	特征值	特征向量
$\boldsymbol{A}$	λ	α
$\boldsymbol{A}^{\mathrm{T}}$	λ	α
$\boldsymbol{A}^{-1}$	λ^{-1}	α
$\boldsymbol{A}^{*}$	$\lvert\boldsymbol{A}\rvert\lambda^{-1}$	α
$\boldsymbol{A}^2-3\boldsymbol{I}$	λ^2-3	α
$h(\boldsymbol{A})$	$h(\lambda)$	α

【特殊矩阵的特征值】

(1)三角阵 $\begin{pmatrix} a_1 & * & \cdots & * \\ & a_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & a_n \end{pmatrix}$ 或对角矩阵的特征值是对角线上的元素：$a_1, a_2, \cdots, a_n$.

特别地，零矩阵的特征值只有0，数量矩阵 $k\boldsymbol{I}$ 的特征值只有 k.

(2)实对称阵($\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$ 且元素全是实数)的特征值是实数. 它的非零特征值的个数等于 $r(\boldsymbol{A})$.

(3)分块对角阵 $\begin{pmatrix} \boldsymbol{A} & \\ & \boldsymbol{B} \end{pmatrix}$ 的特征值由 $\boldsymbol{A}$ 及 $\boldsymbol{B}$ 的特征值构成.

(4)秩为1的矩阵必可写做(见矩阵分解定理)两个非零列向量 α, β 的乘积：$\boldsymbol{A}=\alpha\beta^{\mathrm{T}}$，由于 $\boldsymbol{A}\alpha=\alpha\beta^{\mathrm{T}}\alpha=(\beta^{\mathrm{T}}\alpha)\alpha$，故 $\beta^{\mathrm{T}}\alpha$ 是它的一个特征值.

(5)各行元素之和为 a 的矩阵有一个特征值 a，对应的特征向量是 $(1,1,\cdots,1)^{\mathrm{T}}$.

(6)正交矩阵的特征值或是1，或是−1.

(7)方阵 $\boldsymbol{A}, \boldsymbol{B}$ 的乘积 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 有相同的特征值.

(8)可逆阵的特征值全非零.

(9)如果矩阵 $\boldsymbol{A}$ 满足多项式矩阵方程 $h(\boldsymbol{A})=\boldsymbol{0}$，则矩阵 $\boldsymbol{A}$ 的特征值 λ 也满足代数方程 $h(x)=0$；但是要注意：$h(x)=0$ 的根不一定都是 $\boldsymbol{A}$ 的特征值.

【特征向量的叠加性质】矩阵 $\boldsymbol{A}$ 关于特征值 λ_i 的 m 个特征向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的任何非零线性组合 $\sum\limits_{i=1}^{m} k_i\alpha_i$

还是 A 关于 λ_i 的特征向量.

【特征向量的线性相关性】

(1)n 阶矩阵 A 至多有 n 个线性无关的特征向量,每个特征值至少对应一个非零的特征向量.

(2)n 阶矩阵 A 分属于不同特征值 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 的特征向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性无关的.

(3)n 阶矩阵 A 的不同特征值所属下的线性无关特征向量子组汇总在一起还是线性无关的.

(4)n 阶矩阵 A 的 t 重特征值对应的特征向量组的秩小于或等于 t.

〖推论〗若 A 有 n 个互异的特征值:$\lambda_1,\lambda_2,\cdots,\lambda_n$,则每个 λ_i 对应且仅对应一个线性无关的特征向量,从而 A 共有 n 个线性无关的特征向量.

13.2　矩阵相似对角化

【矩阵相似的概念】对 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,若有可逆的 n 阶方阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$,则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似,或 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$,记作 $\boldsymbol{A}\sim\boldsymbol{B}$,称矩阵 $\boldsymbol{P}$ 为相似变换矩阵.

【矩阵相似关系的性质】

(1)自发性　$\boldsymbol{A}\sim\boldsymbol{A}$;

(2)对称性　$\boldsymbol{A}\sim\boldsymbol{B}$,则 $\boldsymbol{B}\sim\boldsymbol{A}$;

(3)传递性　$\boldsymbol{A}\sim\boldsymbol{B}$ 且 $\boldsymbol{B}\sim\boldsymbol{C}$,则 $\boldsymbol{A}\sim\boldsymbol{C}$.

【相似矩阵的共性】设 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似,则有

(1)$r(\boldsymbol{A})=r(\boldsymbol{B})$;

(2)$|\lambda\boldsymbol{I}-\boldsymbol{A}|=|\lambda\boldsymbol{I}-\boldsymbol{B}|$;

(3)$|\boldsymbol{A}|=|\boldsymbol{B}|$,$\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{B})$;

(4)A 和 B 的特征值相同.

〖点拨〗两个实对称矩阵相似的充分必要条件是:它们有相同的特征值,包括重数.

【相似关系与等价关系的比较】

(1)相似关系是两个 n 阶方阵之间的关系,等价关系则是两个同型矩阵之间的关系.

(2)相似变换结果是 $P^{-1}AP=B$;等价变换结果则是 $QAP=B$. 都要求变换矩阵是可逆矩阵.

(3)两个矩阵相似⇒两个矩阵等价,反之不然.

(4)等价变换仅保持矩阵的秩不变,而相似变换还保持特征值不变等.

【n 阶矩阵 A 可以相似于对角矩阵的判定】

〖充要条件1〗A 有 n 个线性无关的特征向量.

〖充要条件2〗t_i 重特征值 λ_i 对应的矩阵 $\lambda_i\boldsymbol{I}-A$ 的秩为 $r(\lambda_i\boldsymbol{I}-A)=n-t_i$.

〖充分条件1〗A 有 n 个互异的特征值.

〖充分条件2〗A 为实数构成的对称矩阵.

【矩阵可对角化的传递性质】如果矩阵 A 可相似对角化,则 A^{T},A^{-1},A^* 以及 $h(A)$($h(x)$为多项式)也可以相似对角化.

〖点拨〗利用 $A=P^{-1}\Lambda P$,Λ 为对角矩阵来证明.

【相似对角化结果表述】设矩阵 A 可以相似于对角矩阵,则 A 有 n 个线性无关的列特征向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$,依次属于 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$(允许重复),则相似变换矩阵可取为 $P=(\alpha_1\ \alpha_2\cdots\ \alpha_n)$. 对角化结果表述为

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

【实对称矩阵正交相似对角化的结果表述】实对称矩阵一定可以使用正交矩阵化作对角矩阵. 即有正交矩阵C,使得$C^{-1}AC=C^{\mathrm{T}}AC$为对角矩阵$\Lambda$. 其中

$$\Lambda=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},\quad C=(\alpha_1\ \alpha_2\ \cdots\ \alpha_n),$$

$\alpha_1,\alpha_2,\cdots,\alpha_n$依次是$\lambda_1,\lambda_2,\cdots,\lambda_n$对应的彼此正交的标准特征向量.

【利用正交性计算三阶实对称矩阵的特征向量】根据不同特征值对应的特征向量必正交的性质,可以设计以下的算法. 设矩阵为三阶实对称阵,则有:

(1)若已求得的两个特征向量α_1,α_2来自不同的特征值,则使用向量代数中的叉乘可以得到第三个特征向量$\alpha_1\times\alpha_2$,这三个向量构成一个正交特征向量组;

(2)如果特征值是一个二重根和一个单根,则与单根的特征向量正交的不成比例的两个向量α_1,α_2必是另一个二重特征值的特征向量.

常见题型·应对

【计算具体矩阵的特征值、特征向量】直接依据定义,先求特征值,再通过齐次方程组的基础解系得到

对应的特征向量.

例　求$A=\begin{pmatrix}1&-3&3\\3&-5&3\\6&-6&4\end{pmatrix}$的特征值和特征向量.

解　由　$|\lambda I-A|=\begin{vmatrix}\lambda-1&3&-3\\-3&\lambda+5&-3\\-6&6&\lambda-4\end{vmatrix}=0$,

得特征值　　$\lambda_1=\lambda_2=-2,\quad \lambda_3=4.$

当 $\lambda_1=\lambda_2=-2$ 时,求解$(-2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$,即

$$\begin{pmatrix}-3&3&-3\\-3&3&-3\\-6&6&-6\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

得基础解系$(1,1,0)^{\mathrm{T}},(-1,0,1)^{\mathrm{T}}$ 为其对应的特征向量.

当 $\lambda_3=4$ 时,求解$(4\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$,即

$$\begin{pmatrix}3&3&-3\\-3&9&-3\\-6&6&0\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

得基础解系$(1,1,2)^{\mathrm{T}}$ 为其对应的特征向量.

【判定给定矩阵能否对角化】这是基本题型,依据可对角化的判别条件进行判定.

例 1　判断下列矩阵能否相似于对角阵,若能,求出相似变换矩阵 $\boldsymbol{P}$.

$$(1)\boldsymbol{A}_1=\begin{pmatrix}2&0&0\\1&1&0\\1&1&1\end{pmatrix};(2)\boldsymbol{A}_2=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}.$$

解　(1)由$|\lambda\boldsymbol{I}-\boldsymbol{A}_1|=(\lambda-2)(\lambda-1)^2=0$,得 λ_1

$=\lambda_2=1,\lambda_3=2$,其中 $\lambda_1=\lambda_2=1$ 为二重根.由于

$$r(1\cdot\boldsymbol{I}-\boldsymbol{A}_1)=r\left(\begin{pmatrix}-1&0&0\\-1&0&0\\-1&-1&0\end{pmatrix}\right)=2,$$

故二重特征值 $\lambda=1$ 只有一个线性无关的特征向量,故该矩阵不能相似于对角形.

(2)由 $|\lambda\boldsymbol{I}-\boldsymbol{A}_2|=(\lambda+1)^2(\lambda-5)=0$,得 $\lambda_1=\lambda_2=-1,\lambda_3=5$.因为 $\boldsymbol{A}_2$ 是实对称矩阵,故可对角化.$\lambda_1=\lambda_2=-1$ 对应的线性无关特征向量为

$$(-1,1,0)^{\mathrm{T}},(-1,0,1)^{\mathrm{T}}.$$

$\lambda_3=5$ 为单特征值,它对应的特征向量为 $(1,1,1)^{\mathrm{T}}$.所以

$$\boldsymbol{P}=\begin{pmatrix}-1&-1&1\\1&0&1\\0&1&1\end{pmatrix},\quad \boldsymbol{P}^{-1}\boldsymbol{A}_2\boldsymbol{P}=\begin{pmatrix}-1&&\\&-1&\\&&5\end{pmatrix}.$$

例2　确定参数 t 的值,使得 $\boldsymbol{A}=\begin{pmatrix}1&t&3\\&1&4\\&&2\end{pmatrix}$ 可以相似对角化.

解　关键是判定二重特征值 1 对应的矩阵 $A-\boldsymbol{I}$ 的秩能否为 1.由于

$$\boldsymbol{A}-\boldsymbol{I}=\begin{pmatrix}0&t&3\\&0&4\\&&1\end{pmatrix},$$

故当 $t=0$ 时,$r(\boldsymbol{A}-\boldsymbol{I})=1$,矩阵可以相似对角化,否则不行.

例3　满足矩阵方程$(\boldsymbol{A}-a\boldsymbol{I})(\boldsymbol{A}-b\boldsymbol{I})=\boldsymbol{0}$的$n$阶矩阵$A$可以相似对角化，其中$a$与$b$为不同的实数.

证　由于

$$r(\boldsymbol{A}-a\boldsymbol{I})+r(\boldsymbol{A}-b\boldsymbol{I})\leqslant n,$$

$$\begin{aligned}&r(\boldsymbol{A}-a\boldsymbol{I})+r(\boldsymbol{A}-b\boldsymbol{I})\\=&r(a\boldsymbol{I}-\boldsymbol{A})+r(\boldsymbol{A}-b\boldsymbol{I})\\\geqslant&r(a\boldsymbol{I}-\boldsymbol{A}+\boldsymbol{A}-b\boldsymbol{I})\\=&r((a-b)\boldsymbol{I})=n,\end{aligned}$$

故$r(\boldsymbol{A}-a\boldsymbol{I})+r(\boldsymbol{A}-b\boldsymbol{I})=n$. 于是，当$r(\boldsymbol{A}-a\boldsymbol{I})=r$时，特征值$a$对应的线性无关的特征向量有$n-r$个，此时，$r(\boldsymbol{A}-b\boldsymbol{I})=n-r$，特征值$b$对应的线性无关的特征向量有$r$个，加起来恰好是$n$个，故由可对角化充要条件知矩阵$A$可以相似对角化.

【计算矩阵或矩阵多项式的行列式】

例　设三阶方阵A的三个特征值分别为2，3，7，求行列式$|5\boldsymbol{A}+\boldsymbol{I}|$.

解　若λ_i是A的特征值，则$5\lambda_i+1$为$5\boldsymbol{A}+\boldsymbol{I}$的特征值，即$5\boldsymbol{A}+\boldsymbol{I}$有特征值

$$5\times2+1,\quad 5\times3+1,\quad 5\times7+1,$$

所以　$|5\boldsymbol{A}+\boldsymbol{I}|=11\times16\times36=6336.$

【含参数的矩阵特征值问题】给出含两个参数的矩阵$A(a,b)$的全部特征值$\lambda_1,\lambda_2,\lambda_3$，要求计算特征向量和变换矩阵.

例　已知矩阵$A=\begin{pmatrix}1&b&0\\-2&a&0\\0&0&-3\end{pmatrix}$的特征值为

3,3,0,计算参数 a,b 的值,并求变换矩阵 $\boldsymbol{P}$,使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}3&&\\&3&\\&&0\end{pmatrix}.$$

解 利用 $\operatorname{tr}(\boldsymbol{A})=\lambda_1+\lambda_2+\lambda_3$, $|\boldsymbol{A}|=\lambda_1\lambda_2\lambda_3$,便可求得 $a=2,b=-1$, $\boldsymbol{P}=\begin{pmatrix}1&0&1\\-2&0&1\\0&1&0\end{pmatrix}$(过程略).

〖点拨〗还可以利用 $|\lambda_i\boldsymbol{A}-\boldsymbol{I}|=0$ $(i=1,2,3)$,或者定义式 $A\alpha=\lambda\alpha$ 来列关于 a,b 的方程.

【其他含参数问题】理解题目的含义,列出基本的方程式,应当可以找到解题的思路.

例 已知 $\alpha=(1,k,1)^{\mathrm{T}}$ 是矩阵 $\boldsymbol{A}=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$ 的逆矩阵 A^{-1} 的特征向量.求 k 值.

解 设 $\boldsymbol{A}^{-1}\alpha=\lambda\alpha$,两边左乘 $\boldsymbol{A}$,得 $\lambda A\alpha=\alpha$,把 α, $\boldsymbol{A}$ 值代入,得

$$\lambda\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}\begin{pmatrix}1\\k\\1\end{pmatrix}=\begin{pmatrix}1\\k\\1\end{pmatrix},$$

展开消去 λ,得 $k^2+k-2=0$,从而得 $k_1=1,k_2=-2$.

【求矩阵 A 的 n 次幂问题】这类题的解法较多,此处主要是利用对角矩阵的 n 次幂还是对角矩阵的性质,将所给矩阵 A 对角化之后直接计算便是.

例 设 $\boldsymbol{A}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$,计算 $\boldsymbol{A}^n$.

解　可以求得特征值为 3 和 -1，对应的特征向量构成变换矩阵

$$\boldsymbol{P}=\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix},$$

于是

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}3 & \\ & -1\end{pmatrix},$$

从而

$$\boldsymbol{A}^n=\left(\boldsymbol{P}\begin{pmatrix}3 & \\ & -1\end{pmatrix}\boldsymbol{P}^{-1}\right)^n=\boldsymbol{P}\begin{pmatrix}3 & \\ & -1\end{pmatrix}^n\boldsymbol{P}^{-1}$$

$$=\boldsymbol{P}\begin{pmatrix}3^n & \\ & (-1)^n\end{pmatrix}\boldsymbol{P}^{-1}$$

$$=\frac{1}{2}\begin{pmatrix}3^n+(-1)^n & 3^n-(-1)^n\\3^n-(-1)^n & 3^n+(-1)^n\end{pmatrix}.$$

第14章　二　次　型

大纲要求

掌握二次型及其矩阵表示，了解二次型的秩的概念，了解合同变换与合同矩阵的概念，了解二次型的标准形、规范形的概念以及惯性定理．掌握用正交变换化二次型为标准形的方法，会用配方法化二次型为标准形．理解正定二次型、正定矩阵的概念，并掌握其判别法．

注　数学二不要求本章内容．

知识点·点拨

14.1　二次型的标准形

【二次型】称 n 个实变量的二次齐次多项式

$$
\begin{aligned}
&f(x_1,x_2,\cdots,x_n)\\
&=a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n\\
&\quad+a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n\\
&\quad+\cdots+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2
\end{aligned}
$$

为 n 元实二次型，简称 n 元二次型．

【二次型的矩阵表示】设 $a_{ij}=a_{ji}\quad(1\leqslant i,j\leqslant n)$，取向量 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，记矩阵

$$
\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}
$$

则上述二次型 f 可以表示为

$$f(x_1,x_2,\cdots,x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X},$$

称之为二次型的矩阵形式，其中，$\boldsymbol{A}$ 称为二次型 $f(x_1,x_2,\cdots,x_n)$ 所对应的矩阵.

〖点拨〗二次型的矩阵 $\boldsymbol{A}$ 是实对称矩阵.

例　二次型 $3x_1^2+2x_2^2-x_3^2+2x_1x_2-4x_2x_3$ 可以表示为 $(x_1,x_2,x_3)\begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & -2 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

【二次型的秩】指二次型的矩阵的秩.

【二次型的标准形】只含平方项的二次型

$$f(x_1,x_2,\cdots,x_n) = d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2$$

称为二次型的标准形.

显然，标准形的矩阵为对角矩阵：

$$\boldsymbol{\Lambda} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

〖点拨〗标准形中系数的大小不惟一，但是正负系数的个数(即正负惯性指数)是确定的.

【二次型的规范形】只含平方项的二次型且形如

$$f(x_1,x_2,\cdots,x_n)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2$$

的二次型称为二次型的规范形.

规范形的矩阵如下所示：

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}.$$

【化标准形问题】二次型的基本问题是求一个可逆的线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，即

$$\begin{cases} x_1 = p_{11}y_1 + p_{12}y_2 + \cdots + p_{1n}y_n, \\ x_2 = p_{21}y_1 + p_{22}y_2 + \cdots + p_{2n}y_n, \\ \quad \vdots \\ x_n = p_{n1}y_1 + p_{n2}y_2 + \cdots + p_{nn}y_n, \end{cases}$$

将二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 化为标准形 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{Y}$.

使用矩阵语言来说便是：对一个实对称矩阵 $\boldsymbol{A}_{n\times n}$，求一个可逆矩阵 $\boldsymbol{P}_{n\times n}$，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

14.2　矩阵的合同

【矩阵的合同】设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵，若存在 n 阶可逆

矩阵 P,使 $P^{\mathrm{T}}AP=B$,则称 A 合同于 B,记为 $A\simeq B$.

【合同关系的性质】

(1)自反性:$A\simeq A$;

(2)对称性:$A\simeq B$,则 $B\simeq A$;

(3)传递性:$A\simeq B$,且 $B\simeq C$,则 $A\simeq C$.

【合同关系的不变性】如果 $A\simeq B$,则

(1)$r(A)=r(B)$;

(2)A,B 的惯性指数相同;

(3)若 A 正定,则 B 也正定.

〖点拨〗两个合同矩阵的特征值可以不同.

【合同与相似的比较】

(1)合同推不出相似:例如,取 $P=\begin{pmatrix}1&\\&2\end{pmatrix}$ 可以使得 $\begin{pmatrix}1&\\&1\end{pmatrix}\simeq\begin{pmatrix}1&\\&4\end{pmatrix}$,但是 $\begin{pmatrix}1&\\&1\end{pmatrix}$ 与 $\begin{pmatrix}1&\\&4\end{pmatrix}$ 却不相似,因为特征值不同.

(2)若合同变换矩阵 P 是正交矩阵,则(由于 $P^{-1}=P^{\mathrm{T}}$)合同关系 $P^{\mathrm{T}}AP=B$ 也是相似关系 $A\sim B$.

【实对称矩阵的合同与相似】

(1)两个实对称矩阵相似$\Leftrightarrow$它们的特征值相同.

(2)两个实对称矩阵合同$\Leftrightarrow$它们的正负惯性指数相同.

(3)两个实对称矩阵相似$\Rightarrow$两个实对称矩阵合同.

例　判定以下两个矩阵是否合同:

$$A=\begin{pmatrix}1&2&0\\2&1&0\\0&0&3\end{pmatrix},\quad B=\begin{pmatrix}1&&\\&-2&\\&&3\end{pmatrix}.$$

解　先将矩阵 A 的第一行的 -2 倍加到第二行,然后将所得结果的第一列的 -2 倍加到第二列,使之成为对角矩阵,其惯性指数与矩阵 B 一样,故两个矩阵合同.

【分块矩阵的合同】$\boldsymbol{P}_1\simeq\boldsymbol{Q}_1,\boldsymbol{P}_2\simeq\boldsymbol{Q}_2$ 的充分必要条件为

$$\begin{pmatrix} \boldsymbol{P}_1 & \\ & \boldsymbol{P}_2 \end{pmatrix} \simeq \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix}.$$

【标准形定理】任何一个实对称矩阵 A 都可以合同于某个对角矩阵,即存在可逆矩阵 P,使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix},$$

从而任何一个实二次型 $X^{\mathrm{T}}AX$ 都可用可逆线性变换 $X=PY$ 化为标准形.

〖点拨〗进一步的结果:对任何实对称矩阵,都可以求得正交矩阵 C,使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

其中 $\lambda_i(1\leqslant i\leqslant n)$是矩阵 A 的特征值. 正交矩阵 C 是对角化问题中的特征向量矩阵经过施密特方法改造之后的矩阵.

【三种化二次型为标准形的方法】

(1)拉格朗日配方法;

(2)对矩阵的行列同步的初等变换法;

(3)对矩阵的正交变换法.

14.3 二次型的正定性

【二次型的正定性问题】任取 $X\neq 0$,考虑实二次型 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$ 是否恒取正值的问题.

〖点拨〗列向量 $X\neq 0$ 是指它至少有一个分量 x_i 不为零,不是所有的分量不为零.

【正(负)定二次型】对 n 元实二次型 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$,如果任取 $X\neq 0$,恒有 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX>0(f<0)$,则称 $f(x_1,x_2,\cdots,x_n)$ 为正(负)定二次型,并称二次型矩阵 A 为正(负)定矩阵.

【惯性定理】实二次型 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$ 经可逆线性变换化为标准形时,其标准形中正、负平方项的个数是惟一确定的,它们的和等于矩阵 A 的秩.

〖推论〗任何一个实二次型 $f(x_1,x_2,\cdots,x_n)$ 经过适当的可逆线性变换 $X=PY$ 可以化为规范形

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2,$$

该规范形是惟一的.

【惯性指数】在实二次型 $f(x_1,x_2,\cdots,x_n)$ 的标准形或规范形中,正平方项的个数 p 称为二次型 $f(x_1,x_2,\cdots,x_n)$ 的正惯性指数,负平方项个数 q 称为二次型 $f(x_1,x_2,\cdots,x_n)$ 的负惯性指数,其差 $(p-q)$ 称为二次型的符号差.

〖点拨〗注意到 $p+q=r(A)=r$,所以二次型

的符号差：

$$p-q=p-(r-q)=2p-r.$$

【顺序主子阵】设 $A=(a_{ij})_{n\times n}$ 为 n 阶实对称矩阵，则顺序取 A 的前 k 行与前 k 列构成的矩阵

$$\boldsymbol{A}_k=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix} \quad (k=1,2,\cdots,n)$$

称为 A 的 k 阶顺序主子阵，它的行列式 $|A_k|$ 称为 A 的 k 阶顺序主子式. A 共有 n 个顺序主子阵，它们仍然是实对称矩阵.

【判定二次型为正定的方法归纳】判定 n 元实二次型 $X^{\mathrm{T}}AX$ 为正定二次型或判定二次型矩阵 A 为正定矩阵的常见方法如下.

〖充要条件〗下列条件之一：

(1) $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 的正惯性指数等于 n；

(2) $\boldsymbol{A}$ 的全部特征值都是正数；

(3) 存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\boldsymbol{I}$；

(4) 存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$；

(5) $\boldsymbol{A}$ 的所有顺序主子式都大于 0.

〖必要条件〗下列条件之一：

(1) 矩阵 A 中主对角线上的元素都为正数；

(2) 矩阵 A 的行列式大于 0；

(3) 矩阵 A 可以分解为两个正定矩阵的乘积.

【判定二次型负定的充要条件】n 元实二次型 $X^{\mathrm{T}}AX$ 为负定二次型的充要条件是下列条件之一：

(1)$X^{\mathrm{T}}AX$ 的负惯性指数为 n；

(2)存在可逆矩阵 P，使 $P^{\mathrm{T}}AP=-I_n$；

(3)A 的奇数阶顺序主子式都小于零，A 的偶数阶顺序主子式都大于零.

【正定矩阵的重要性质】设 A 和 B 是同阶正定矩阵，则有：

(1)A 与 B 的和 $A+B$ 是正定矩阵；

(2)如果 $AB=BA$，则 AB 是正定矩阵；

(3)A^{-1}，A^{T}，A^{*}，$h(A)$ 也是正定矩阵(多项式 $h(x)>0$，$x>0$).

〖点拨〗正定矩阵必须首先是对称矩阵.

【特殊矩阵的正定性】设 $A_{n\times n}$ 为实对称矩阵，则有：

(1)若多项式方程 $h(x)=0$ 只有正根，矩阵 A 满足 $h(A)=\mathbf{0}$，则矩阵 A 为正定矩阵；

(2)若矩阵 A 可逆，则 AA^{T} 正定；

(3)若 $r(A_{m\times n})=m$，则 AA^{T} 正定；

(4)若矩阵 A 正定，B 反对称，则 $A-B^2$ 正定；

(5)正定且正交的矩阵是单位矩阵.

例　如果实矩阵 $A_{3\times 4}$ 的秩为 3，证明 AA^{T} 为正定矩阵.

证　显然，AA^{T} 是对称矩阵. 任取 $X\neq\mathbf{0}$，只需证

$$X^{\mathrm{T}}AA^{\mathrm{T}}X=(A^{\mathrm{T}}X)^{\mathrm{T}}(A^{\mathrm{T}}X)>0.$$

这只要说明向量 $A^{\mathrm{T}}X\neq\mathbf{0}$. 由于 A^{T} 的秩为 3，故齐次方程 $A^{\mathrm{T}}X=\mathbf{0}$ 的解空间的维数是 0，它没有非零解. 由于 $X\neq\mathbf{0}$，因此 $A^{\mathrm{T}}X\neq\mathbf{0}$.

常见题型·应对

【使用配方法化二次型为标准形】拉格朗日配方法是利用代数公式将二次型配成完全平方式的方法. 下面举例说明该方法.

例 1　用配方法化二次型

$$f(x_1,x_2,x_3)$$
$$= x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$$

为标准形.

解

$$f= x_1^2+2x_1x_2+2x_1x_3+2x_2^2+6x_2x_3+5x_3^2$$
$$=(x_1+x_2+x_3)^2+x_2^2+4x_3^2+4x_2x_3$$
$$=(x_1+x_2+x_3)^2+(x_2+2x_3)^2,$$

于是可以令可逆变换为

$$\begin{cases} y_1=x_1+x_2+x_3, \\ y_2=x_2+2x_3, \\ y_3=x_3, \end{cases}$$

即

$$\begin{cases} x_1=y_1-y_2+y_3, \\ x_2=y_2-2y_3, \\ x_3=y_3, \end{cases}$$

或

$$\boldsymbol{X}=\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\boldsymbol{Y},$$

便可以把题给二次型化为标准形

$$f(x_1,x_2,x_3)=y_1^2+y_2^2.$$

例 2　用配方法化二次型

$$f(x_1,x_2,x_3)=x_1x_2+4x_1x_3+x_2x_3$$

为标准形.

解　由于$f(x_1,x_2,x_3)$不含平方项,故先用下列变换让$f(x_1,x_2,x_3)$产生平方项:

令$\begin{cases}x_1=y_1+y_2,\\x_2=y_1-y_2,\\x_3=y_3,\end{cases}$　或　$\begin{cases}y_1=\dfrac{1}{2}x_1+\dfrac{1}{2}x_2,\\y_2=\dfrac{1}{2}x_1-\dfrac{1}{2}x_2,\\y_3=x_3.\end{cases}$

得$f=y_1^2-y_2^2+4(y_1+y_2)y_3+(y_1-y_2)y_3$

$$=y_1^2-y_2^2+5y_1y_3+3y_2y_3\quad(\text{配方})$$

$$=\left(y_1+\frac{5}{2}y_3\right)^2-\left(y_2-\frac{3}{2}y_3\right)^2-(2y_3)^2$$

$$=z_1^2-z_2^2-z_3^2.$$

其中可逆线性变换为

$$\begin{cases}z_1=y_1+\dfrac{5}{2}y_3,\\z_2=y_2-\dfrac{3}{2}y_3,\\z_3=2y_3.\end{cases}$$

从而

$$\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\begin{pmatrix}1&0&\frac{5}{2}\\0&1&-\frac{3}{2}\\0&0&2\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

$$=\begin{pmatrix}1&0&\frac{5}{2}\\0&1&-\frac{3}{2}\\0&0&2\end{pmatrix}\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&0\\\frac{1}{2}&-\frac{1}{2}&0\\0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},$$

令 $P=\left(\begin{pmatrix}1 & 0 & \frac{5}{2}\\ 0 & 1 & -\frac{3}{2}\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & -\frac{1}{2} & 0\\ 0 & 0 & 1\end{pmatrix}\right)^{-1}$，则可逆线性变换 $X=PY$ 将二次型化为标准形

$$f(x_1,x_2,x_3)=z_1^2-z_2^2-z_3^2.$$

【使用合同变换化二次型为标准形】 在使用成对的初等列变换和初等行变换产生对角矩阵时，记录所作的变换，产生可逆变换矩阵 P.

例 用合同变换把二次型

$$f=x_1^2+2x_2^2+5x_3^2+2x_1x_2+6x_2x_3+2x_1x_3$$

化为标准形.

解

$$\begin{pmatrix}A\\ \hdashline I\end{pmatrix}=\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 3\\ 1 & 3 & 5\\ \hdashline 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\xrightarrow{①}\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 2\\ 1 & 2 & 4\\ \hdashline 1 & -1 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{②}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 2\\ 0 & 2 & 4\\ \hdashline 1 & -1 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\xrightarrow{③}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 2 & 0\\ \hdashline 1 & -1 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{④}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ \hdashline 1 & -1 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Lambda}\\ \hdashline \boldsymbol{P}\end{pmatrix},$$

对角化成功. 其中,①把第一列乘(−1)分别加到第二列和第三列上,②把第一行乘(−1)分别加到第二行和第三行上,③把第二列乘(−2)加到第三列,④把第二行乘(−2)加到第三行上. 于是

$$\boldsymbol{P}=\begin{pmatrix}1 & -1 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix},\quad \boldsymbol{D}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

故在线性变换 $\boldsymbol{X}=\boldsymbol{PY}$ 下所求标准形为 $y_1^2+y_2^2$.

【使用正交变换化二次型为标准形】实二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 必可由正交变换 $\boldsymbol{X}=\boldsymbol{CY}$ 化为以特征值为系数的标准形:

$$f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}\overset{\boldsymbol{X}=\boldsymbol{CY}}{=\!=\!=\!=}\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2,$$

其中,$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值.

〖主要步骤〗

(1)求 $\boldsymbol{A}$ 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$;

(2)对 λ_i,求$(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系;

(3)对 $t(t>1)$重特征值 λ_j,用施密特正交化方法,将其 t 个线性无关的特征向量正交化. 注意不同特征值对应的特征向量之间已经彼此正交.

(4)将 A 的 n 个正交的特征向量单位化,再以它们为列向量构成正交矩阵 C,并写出相应的正交变换式 $X=CY$ 和二次型的标准形.

【含参数标准形问题】已知含参数的二次型使用正交变换后所变成的标准形,求其中的参数.

例　已知二次型 $ax_1^2+3x_2^2+3x_3^2+4x_2x_3$,通过正交变换 $X=CY$ 化为标准形 $y_1^2+2y_1y_2^2+5y_3^2$,求参数 a 及所用的正交变换矩阵 C.

解　由于正交合同关系也是相似关系,因此,二次型的矩阵

$$\boldsymbol{A}=\begin{pmatrix} a & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

与标准形的矩阵 $\begin{pmatrix} 1 & & \\ & 2 & \\ & & 5 \end{pmatrix}$ 是相似关系,相似关系能保持特征值不变,故 A 的特征值为 1,2,5,且由迹 $\operatorname{tr}(\boldsymbol{A})=a+3+3=1+2+5$,得 $a=2$. 对 $\lambda_1=1,\lambda_2=2,\lambda_3=5$ 分别求得特征向量

$\alpha_1=(0,1,-1)^{\mathrm{T}},\alpha_2=(1,0,0)^{\mathrm{T}},\alpha_3=(0,1,1)^{\mathrm{T}}$,

由于特征值互不相同,故 $\alpha_1,\alpha_2,\alpha_3$ 是相互正交的,将它们单位化,便得所求的正交矩阵为

$$\boldsymbol{C}=\begin{pmatrix} 0 & 1 & 0 \\ \dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \\ \dfrac{-1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \end{pmatrix}.$$

【判定二次型的正定性问题】直接按照判别条件和主要性质操作.

【含参数的二次型问题】已知含参数的二次型的正定性,求其中的参数变化范围.

例　求t的取值范围,使下面二次型为正定二次型:

$$f = x_1^2 + x_2^2 + 5x_3^2 + 2tx_1x_2 - 2x_1x_3 + 4x_2x_3.$$

解　二次型矩阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & t & -1 \\ t & 1 & 2 \\ -1 & 2 & 5 \end{pmatrix},$$

取t使A的顺序主子式大于零.

$$|\boldsymbol{A}_1| = 1 > 0, \quad |\boldsymbol{A}_2| = \begin{vmatrix} 1 & t \\ t & 1 \end{vmatrix} = 1 - t^2 > 0,$$

$$|\boldsymbol{A}_3| = |\boldsymbol{A}| = -5t^2 - 4t > 0,$$

解联立不等式

$$\begin{cases} t^2 - 1 < 0, \\ t(5t + 4) < 0, \end{cases}$$

即

$$-\frac{4}{5} < t < 0,$$

故当t取值满足$-\frac{4}{5}<t<0$时,f是正定二次型.

第3篇　概率论与数理统计

第15章　随机事件和概率

大纲要求

了解样本空间(基本事件空间)的概念,理解随机事件的概念,掌握事件间的关系及运算.理解概率、条件概率的概念、掌握概率的基本性质,会计算古典型概率和几何型概率.掌握计算概率的加法公式、减法公式、乘法公式、全概率公式,以及贝叶斯公式.理解事件独立性的概念,掌握用事件独立性进行概率计算;理解独立重复试验的概念,掌握计算有关事件概率的方法.

知识点·点拨

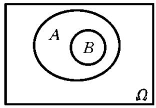
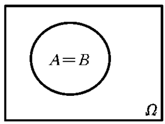
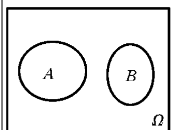
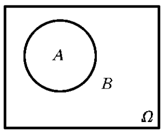
15.1　随机事件与样本空间

【样本空间】随机试验的最简单结果称为基本事件.全体基本事件构成的集合称为样本空间(基本事件空间),常记作Ω.样本空间中的元素(基本事件)也称为样本点.

【随机事件】若干基本事件(样本点)构成的集合称

为随机事件，简称为事件. 它表现为样本空间的子集，常记作 $A,B,C,\cdots$. 特别地，空集称为不可能事件，记作 $\varnothing$. 样本空间 Ω 也称为必然事件.

【事件的关系】

名称及记号	定　义	集合解释	图　形
A 包含 B，记作 $A\supseteq B$ 或 $B\subseteq A$	若 B 发生，则 A 也发生	B 中的样本点也含在 A 中	
A 与 B 等价记作 $A=B$	$A\supseteq B$ 且 $B\supseteq A$ 即 A 与 B 同时发生或同时不发生	B 中的样本点含在 A 中，A 中的样本点也含在 B 中	
A 与 B 互斥或 A 与 B 互不相容，记作 $AB=\varnothing$	A 与 B 不能同时发生	A 与 B 没有共同的样本点，即 A 与 B 不相交	
A 与 B 为对立事件，或 A 与 B 为互逆事件，记作 $A=\overline{B}$ 或 $B=\overline{A}$	A 与 B 互斥且 A 与 B 必有一个发生	不在 A 中的样本点含在 B 中，且 A 中的样本点不在 B 中，即 A 与 B 互补	

〖事件的不相容拆分〗将一事件分解成若干互斥事件的和是计算概率的基本技巧.

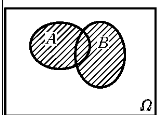
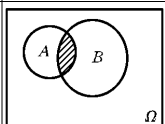
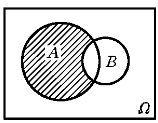
例　将任意 n 件事件 $A_1,A_2,\cdots,A_n$ 的和事件写成 n 个互斥事件的和.

解　设 $B_1=A_1,B_2=\overline{A}_1A_2,B_3=\overline{A}_1\overline{A}_2A_3,\cdots,B_n=\overline{A}_1\overline{A}_2\cdots\overline{A}_{n-1}A_n$,则

$$\bigcup_{i=1}^{n}A_i=\sum_{i=1}^{n}B_i$$

〖逆事件的关键词〗描述逆事件或对立事件的关键词是“不发生”. 将逆事件看作对事件的运算,它与交、并运算不可交换运算次序. 例如,$\overline{A}\ \overline{B}$表示 A 和 B 都不发生,而$\overline{AB}$则表示 A 与 B 不都发生.

【事件的运算】

名称及记号	定　义	集合解释	图　形
A 与 B 的并(和)事件,记作 $C=A\cup B$. 若 $AB=\varnothing$,可记作 $C=A+B$	A 与 B 至少有一个发生,即或者 A 发生,或者 B 发生	$C=A\cup B$ 中的样本点或者在 A 中或者在 B 中	
A 与 B 的交(积)事件,记作 $C=A\cap B$ 或 $C=AB$	A 与 B 同时发生,即 A 发生且 B 也发生	$C=AB$ 中的样本点即在 A 中也在 B 中	
A 与 B 的差事件,记作 $C=A-B$	A 发生且 B 不发生	$C=A-B$ 的样本点在 A 中,但不在 B 中	

〖交、并运算的关键词〗事件并(和)运算的关键词是"至少"、"或者",事件交(积)运算的关键词是"同时"、"且". 例如,$\bigcup_{i=1}^{n} A_i$ 表示 $A_1, A_2, \cdots, A_n$ 中至少有一个发生,$\bigcap_{i=1}^{n} A_i$ 表示 $A_1, A_2, \cdots, A_n$ 同时发生.

【事件运算法则】

交换律　$A\cup B=B\cup A, AB=BA$.

结合律　$(A\cup B)\cup C=A\cup(B\cup C)$,

$(AB)C=A(BC)$.

分配律　$(A\cup B)C=AC\cup BC$,

$(AB)\cup C=(A\cup C)(B\cup C)$.

德摩根(De Morgan)律

$$\overline{\bigcup_i A_i}=\bigcap_i \overline{A_i},\quad \overline{\bigcap_i A_i}=\bigcup_i \overline{A_i}$$

〖交、并运算的转换〗德摩根律(也称为对偶律)表明:通过取事件的逆事件(对立事件)可实现事件间交运算与并运算的转换. 例如,A 与 B 都不发生 $\overline{A}\,\overline{B}$,也就是 A 与 B 不能至少有一个发生 $\overline{A\cup B}$;A 与 B 不同时发生 $\overline{AB}$,也就是 A 与 B 至少有一个不发生 $\overline{A}\cup\overline{B}$.

15.2　概率的定义、性质及计算

【事件的概率】设 Ω 为一个随机试验的样本空间. 对 Ω 上的任一事件 A,规定一个实数 $P(A)$ 与之对应,使其满足:

非负性　$P(A)\geqslant 0$,

规范性　$P(\Omega)=1$,

可列可加性 若 $A_iA_j=\varnothing, i\neq j$，则

$$P\left(\sum_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i),$$

那么，称 $P(A)$ 为事件 A 发生的概率.

【概率的性质】

(1) **不可能事件的概率**：$P(\varnothing)=0$.

(2) **有限可加性**，若 $A_iA_j=\varnothing, i\neq j$，则

$$P\left(\sum_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i).$$

(3) **对立事件的概率**：$P(\overline{A})=1-P(A)$.

(4) **加法公式**：

$$P(A\cup B)=P(A)+P(B)-P(AB).$$

n 个事件的加法公式：

$$\begin{aligned}P\left(\bigcup_{i=1}^{n}A_i\right)=&\sum_{i=1}^{n}P(A_i)-\sum_{i\neq j}P(A_iA_j)\\&+\sum_{i\neq j\neq k}P(A_iA_jA_k)+\cdots\\&+(-1)^{n-1}P(A_1A_2\cdots A_n)\end{aligned}$$

(5) **减法公式**：若 $A\supseteq B$，则 $P(A-B)=P(A)-P(B)$，从而 $P(A)\geqslant P(B)$.

〖**概率的序关系**〗由概率的加法公式和减法公式易见，对任何事件 A 和 B，有

$$\begin{aligned}0&\leqslant P(AB)\leqslant \min\{P(A),P(B)\}\\&\leqslant P(A)\leqslant \max\{P(A),P(B)\}\\&\leqslant P(A\cup B)\leqslant P(A)+P(B).\end{aligned}$$

【概率的计算】

(1) **古典型概率**：设样本空间 Ω 由 n 个样本点构

成，且每个样本点作为一个基本事件有相同的概率. 若事件 A 由 Ω 中的 k 个样本点构成，则其概率为

$$P(A)=\frac{k}{n}.$$

（2）几何型概率：设样本空间 Ω 具有有限的测量值 $\mu(\Omega)$（长度、面积、体积等），且每个样本点出现的可能性相同. 若事件 A 作为 Ω 的子集具有测量值 $\mu(A)$，则其概率为 $P(A)=\dfrac{\mu(A)}{\mu(\Omega)}$.

〖加法原理与乘法原理〗设事件 A 出现共有 k 类方法，第 i 类又有 m_i 种方法，则 A 出现的方法共有

$$n=m_1+m_2+\cdots+m_k$$

种. 此为加法原理. 设事件 A 的发生可分为 k 个步骤，第 i 个步骤有 m_i 种可能的方法，则 A 发生共有

$$n=m_1\times m_2\times\cdots\times m_k$$

种方法. 此为乘法原理.

〖排列与组合〗从 n 个不同的元素中任取 m（$m\leqslant n$）个元素按取出的顺序排成一列，所有可能的结果为排列数

$$P_n^m=n(n-1)\cdots(n-m+1)=\frac{n!}{n-m!}.$$

从 n 个不同元素中一次取出 m 个元素，所有可能的结果为组合数

$$C_n^m=\frac{P_n^m}{m!}=\frac{n!}{m!(n-m)!}.$$

〖抽样模型与放球模型〗古典型概率的计算问题实际上是所有可能情况的计数问题. 它们都可解释为从一批不同的产品中任取一部分产品时所有可

能情况的计数,即抽样模型.也可以解释为将一批球任意放入若干盒子中时所有可能情况的计数,即放球模型.在抽样模型中,强调抽取次序时用排列数计算,不强调抽取次序时用组合数计算.在放球模型中,球为可辨(强调那些球放在某个盒子中)时用排列数计算,球不可辨(仅强调某个盒子放几个球)时用组合数计算.

<table>
<tr><th>模　型</th><th colspan="2">试验方式</th><th>所有情况计数</th></tr>
<tr><td rowspan="4">抽样模型:从 n 件产品中任取 m 件</td><td rowspan="2">有放回抽样</td><td>强调次序</td><td>n^m</td></tr>
<tr><td>不强调次序</td><td>C_{n+m-1}^m</td></tr>
<tr><td rowspan="2">不放回抽样</td><td>强调次序</td><td>P_n^m</td></tr>
<tr><td>不强调次序</td><td>C_n^m</td></tr>
<tr><td rowspan="4">放球模型:将 m 个球任意放入 n 个盒子中</td><td rowspan="2">每盒可放任意多球</td><td>球可辨</td><td>n^m</td></tr>
<tr><td>球不可辨</td><td>C_{n+m-1}^m</td></tr>
<tr><td rowspan="2">每盒最多放一个球</td><td>球可辨</td><td>P_n^m</td></tr>
<tr><td>球不可辨</td><td>C_n^m</td></tr>
</table>

用此法计算古典型概率时,分母和分子的计数一定要在同一试验方式下计算.

例　设袋中有 a 只黑球和 b 只白球,逐一进行不放回地随机抽取,直到剩余的球颜色相同为止.求剩余的球为黑球的概率.

分析　本题从试验方法上看应属抽样模型,但由于它是由未取出的球的颜色确定抽取次数.讨论已取出球的情况就很复杂,不便计算.因此可将此试验“翻译”成放球模型:将 $a+b$ 只球逐一随机放入 a

$+b$ 个有序号的盒子中. 每个盒子只能放一只球,第 i 个盒子中的球就是第 i 次取到的球. 于是在第 k 次抽取后停止试验意着后面的 $a+b-k$ 个盒子放的是同颜色的球.

解　将试验解释成放球模型后,记 $A=\{$后面的盒子中放的全是黑球$\}$,$B=\{$最后一个盒子中放的是黑球$\}$. 注意到,如果后面的盒子放的全是黑球,最后一个盒子放的一定是黑球. 即 $A\subseteq B$;反之,如果最后一个盒子放的是黑球. 那么,后面的盒子放的也全是黑球,即 $B\subseteq A$,所以,$A=B$. 于是所求概率为

$$P(A)=P(B)=\frac{C_{a+b-1}^{a}}{C_{a+b}^{a}}=\frac{a}{a+b}.$$

15.3　条件概率与独立性

【条件概率】设 A,B 为两事件,且 $P(B)>0$,称

$$P(A/B)=\frac{P(AB)}{P(B)}$$

为在事件 B 已发生的条件下,事件 A 发生的条件概率.

〖点拨 1〗条件概率 $P(A|B)$是将 B 视为样本空间 时的概率,从而概率 $P(A)$是当 $B=\Omega$ 时的条件概率的特例. 因此在计算 $P(A|B)$或讨论相应的性质时,可不考虑 B 以外的样本点.

〖点拨 2〗条件概率也是概率,即 $P(A|B)$满足. 概率定义的三条公理,从而概率的所有性质对条件概率也成立,如 $P(\overline{A}|B)=1-P(A|B)$.

例　已知 $0<P(B)<1$ 且 $P[(A_1+A_2)|B]=$

$P(A_1|B)+P(A_2|B)$，则下列选项成立的是：

(A) $P[(A_1+A_2)|\overline{B}]=P(A_1|\overline{B})+P(A_2|\overline{B})$

(B) $P(A_1B+A_2B)=P(A_1B)+P(A_2B)$

(C) $P(A_1+A_2)=P(A_1|B)+P(A_2|B)$

(D) $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)$

解　题设条件表明在 B 的“内部”，A_1 与 A_2“不相交”，而选项(B)也有同样的“含义”. 严格推导如下：由条件概率的加法公式

$$P(A_1A_2|B)$$
$$=P(A_1|B)+P(A_2|B)-P[(A_1+A_2)|B]=0$$
$$P(A_1A_2B)=P(B)P(A_1A_2|B)=0$$
$$P(A_1B+A_2B)=P(A_1B)+P(A_2B)-P(A_1A_2B)$$
$$=P(A_1B)+P(A_2B),$$

即选项(B)成立.

【乘法公式】$P(AB)=P(A)P(B|A)$

$$=P(B)P(A|B)$$

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

〖同时与依次〗$P(AB)$是事件 A、B 同时发生的概率. 在多数情况下，可将试验过程看作两个步骤：第一步 A 可能发生，第二步 B 可能发生. 在计算 $P(AB)$时，可以依次先计算 A 发生的概率 $P(A)$，再计算当 A 发生时 B 发生的条件概率 $P(B|A)$. 两者的乘积即为 $P(AB)$. 对两个以上事件同时发生的概率也可类似依次计算.

【完备事件组】设一组事件 $A_1,A_2,\cdots,A_n$ 满足

$$A_iA_j=\varnothing(i\neq j)\text{ 且 }\sum_{i=1}^{n}A_i=\Omega,$$

则称 $A_1,A_2,\cdots,A_n$ 为一个完备事件组.

【全概率公式】设 $A_1,A_2,\cdots,A_n$ 为一完备事件组,则对任何事件 B 有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i).$$

【贝叶斯公式】设 $A_1,A_2,\cdots,A_n$ 为一完备事件组,事件 B 满足 $P(B)>0$,则

$$P(A_j|B)=\frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)},\quad j=1,2,\cdots,n$$

〖先验概率与后验概率〗全概率公式和贝叶斯公式体现了一种对随机问题求解的思想.公式中的 A_i 为"原因",B 为"结果".全概率公式是由"原因"求"结果";贝叶斯公式则是已知"结果"去查"原因".公式中 $P(A_i)$称为先验概率,$P(A_i|B)$称为后验概率.

〖实用的完备条件组〗全概率公式和贝叶斯公式中的完备事件组,从本质上看是对事件 B 的一个划分:

$$B=A_1B+A_2B+\cdots+A_nB,$$

因此在应用这对公式时可作如下灵活处理:

(1) n 可取为无穷大,即 $B=\sum_{i=1}^{\infty}A_iB$;

(2) 只要 $\sum_{i}A_i\supseteq B$ 就不必要求 $\sum_{i}A_i=\Omega$.

〖多层试验模型〗全概率公式中的"原因"事件可以推广到多层.以两层为例:第一层可能的"原因"

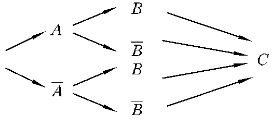
为 $A_1,A_2,\cdots,A_m$,第二层可能的"原因"为:$B_1,B_2,\cdots,B_n$."结果"事件为 C,则这种情况下的全概率公式为

$$\begin{aligned}P(C)&=\sum_{k=1}^{n}P(B_k)P(C\mid B_k)\\&=\sum_{k=1}^{n}\Big[\sum_{i=1}^{m}P(A_i)P(B_k\mid A_i)\Big]P(C\mid B_k)\\&=\sum_{k=1}^{n}\sum_{i=1}^{m}P(A_i)P(B_k\mid A_i)P(C\mid B_k).\end{aligned}$$

全概率公式的"全"可以理解为考虑了导致结果 C 发生的全部途径:$\to A_i\to B_k\to C$ $(i=1,\cdots,m;k=1,\cdots,n)$.

例　某数字信号发射站分别以概率 0.4 和 0.6 发出信号 0 和 1,经过中继站转发后,由接收站接收.设中继站正确转发信号 0 和 1 的概率分别为 0.9 和 0.95,对中继站转发的信号,接收站正确接收的概率为 0.99.试求接收站正确接收到发射站的一个信号的概率.

解　记 $A=\{$发射站发出信号 0$\}$,$B=\{$中继站转发信号 0$\}$,$C=\{$接收站正确的接收到发射站的一个信号$\}$,则 C 发生的全部途径为



相应的概率为

$$P(A)=0.4,\quad \begin{matrix} P(B|A)=0.9, & P(C|AB)=0.99, \\ P(\overline{B}|A)=0.1, & P(C|A\overline{B})=0.01; \end{matrix}$$

$$P(\overline{A})=0.6,\quad \begin{matrix} P(B|\overline{A})=0.05, & P(C|\overline{A}B)=0.01, \\ P(\overline{B}|\overline{A})=0.95, & P(C|\overline{A}\ \overline{B})=0.99. \end{matrix}$$

于是所求概率为

$$\begin{aligned} P(C)&=0.4\times0.9\times0.99+0.4\times0.1\times0.01 \\ &\quad+0.6\times0.05\times0.01+0.6\times0.95\times0.99 \\ &=0.9214. \end{aligned}$$

【事件的独立性】若两事件 A,B **满足**

$$P(AB)=P(A)P(B),$$

则称 A **与** B **相互独立. 若** n **个事件** $A_1,A_2,\cdots,A_n$ **满足如下** 2^n-n-1 **个等式**

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}),$$

$$1\leqslant i_1<i_2<\cdots<i_k\leqslant n,\quad 2\leqslant k\leqslant n,$$

成立, 则称 $A_1,A_2,\cdots,A_n$ **相互独立.**

【独立重复试验】将某试验重复 n **次:** $E_1,E_2,\cdots,E_n$, **若试验** E_i **的任一事件** A_i**. 都有** $A_1,A_2,\cdots,A_n$ **相互独立, 则称** $E_1,E_2,\cdots,E_n$ **为独立重复试验.**

〖**互不相容与相互独立**〗**事件互不相容与相互独立是两个不同的重要概念, 它们的区别和关系如下:**

(1) A **与** B **互不相容由事件是否发生定义,** $AB=\varnothing$; A **与** B **相互独立则是由事件的概率来定义,** $P(AB)=P(A)P(B)$.

(2)**从集合的观点看,** A **与** B **互不相容表示** A **与**

B 没有公共的样本点;A 与 B 独立则表示 A 在 B 中的“比例”与 A 在 $\overline{B}$ 中的“比例”相同. 这说明不论 B 是否发生,A 发生的概率不会改变.

(3)A 与 B 互不相容使加法公式简化为 $P(A+B)=P(A)+(B)$;A 与 B 相互独立使乘法公式简化为 $P(AB)=P(A)P(B)$.

(4)n 个事件两两互不相容,则 n 个事件不相容(不能同时发生),反之不然;n 个事件相互独立,则 n 个事件两两独立,反之不然.

(5)若 A 与 B 的概率均不为零,即 $P(A)P(B)>0$,那么 A 与 B 互不相容则必不独立. 因此,若 A 与 B 既互不相容又相互独立,那么 A 与 B 中至少有一个的概率为零.

例　对任意二事件 A 和 B,

(A) 若 $AB\neq\varnothing$,则 A,B 一定独立

(B) 若 $AB\neq\varnothing$,则 A,B 有可能独立

(C) 若 $AB=\varnothing$,则 A,B 一定独立

(D) 若 $AB=\varnothing$,则 A,B 一定不独立

解　如果没有附加条件,互不相容与相互独立不能互相导出,故只有(B)正确.

常见题型·应对

【事件概率的计算问题】计算简单事件的概率是概率论的基本问题. 对此类问题应注意以下几点:

(1)用排列组合计算概率时,首先应确认它是否古典概型,即该试验的样本空间是否有限集(有限性),每个基本事件(样本点)发生的概率是否相等

(等可能性).

(2)如果对古典概率问题的文字描述难以把握.可将其“翻译”成抽样模型或放球模型来计算概率.

(3)对几何概率的计算,通过作图可以收到事半功倍的效果.

(4)对于较为复杂的古典型概率,应避免直接计数法计算.而尽量利用概率的性质将问题简化分解后再计数.

例　从5双不同的手套中任取4只,求此4只手套中至少有2只配成一双的概率.

分析　此题的试验方式和所求结果都未强调抽取次序,故适合用组合数来计数,总数显然是C_{10}^4.记A为取到的4只手套中至少有2只配成一双.$P(A)$的计算方法有多种.

直接计数法:所取的4只中必有2只成双.它们来自5双中的一双,有C_5^1种情况.另外2只可在余下的8只中任取,有C_8^2种情况,但这样将4只恰为两双的C_5^2种情况进行了重复计数,故概率为

$$P(A)=\frac{C_5^1C_8^2-C_5^2}{C_{10}^4}=\frac{13}{21}.$$

分解法:记A_1为4只手套中恰有2只配成一双,A_2为4只手套恰配成两双,则

$$\begin{aligned}P(A)&=P(A_1+A_2)=P(A_1)+P(A_2)\\&=\frac{C_5^1C_4^2C_2^1C_2^1}{C_{10}^4}+\frac{C_5^2}{C_{10}^4}=\frac{13}{21}.\end{aligned}$$

对立事件法:事件的描述中有“至少”一词时应用概率的对立事件性质往往可收到事半功倍的效

果. 对于本题

$$P(A)=1-P(\overline{A})=1-\frac{C_5^4C_2^1C_2^1C_2^1C_2^1}{C_{10}^4}=\frac{13}{21}.$$

加法公式法:将 5 双手套编号,记 A_i 为所取的 4 只手套中有第 i 双手套($i=1,2,\cdots,5$). 则

$$\begin{aligned}P(A)&=P(\bigcup_{i=1}^{5}A_i)\\&=\sum_{i=1}^{5}P(A_i)-\sum_{i\neq j}P(A_iA_j)+\sum_{i\neq j\neq k}P(A_iA_jA_k)\\&\quad-\sum_{i\neq j\neq k\neq l}P(A_iA_jA_kA_l)+P(A_1A_2A_3A_4A_5)\\&=5\times\frac{C_8^2}{C_{10}^4}-C_5^2\times\frac{1}{C_{10}^4}+C_5^3\times 0-C_5^4\times 0+0\\&=\frac{13}{21}.\end{aligned}$$

注　本题的计数也可用排列数计算,不过分子、分母都要按强调抽取次序来计数,如用对立事件法,计算结果为

$$P(A)=1-\frac{P_5^4P_2^1P_2^1P_2^1}{P_{10}^4}=\frac{13}{21}.$$

事实上,在上述所有概率的计算公式中,分子分母同乘 4! 就是强调抽取次序的计数法.

【全概率公式与贝叶斯公式的应用问题】运用全概率公式和贝叶斯公式解决实际问题是概率论的典型问题. 解决此类问题的步骤如下:

(1)当随机试验分两步(或多步)进行时,或问题中的随机事件涉及两个(或多个)层次时,可确定用全概率公式和贝叶斯公式计算概率.

(2)确定“原因”事件(即完备事件组):A_1,A_2,

$\cdots, A_n$ 以及(先验)概率 $P(A_i)$，$i=1,2,\cdots,n$.

(3)确定“结果”事件 B 以及由“原因”导致“结果”的条件概率 $P(B|A_i), i=1,2,\cdots,n$.

(4)用全概率公式求 $P(B)$，用贝叶斯公式求 $P(A_i|B), i=1,2,\cdots,n$.

例　一架长机带两架僚机执行空中轰炸任务，只有长机知道具体目标. 在飞行途中要经过敌高炮防空区，这时任一架飞机被击落的概率为 0.2. 到达目标上空后，各机将独立地进行轰炸，炸毁目标的概率都是 0.3. 试求目标被炸毁的概率. 如果目标已被炸毁，求三架飞机到达目标上空的概率.

解　记 A_i 为长机与 i 架僚机到达目标上空，$i=0,1,2$，B 为目标被炸毁，则

$$P(A_0)=0.8\times 0.2^2=0.032,$$

$$P(B|A_0)=0.3,$$

$$P(A_1)=C_2^1 0.8^2\times 0.2=0.256,$$

$$P(B|A_1)=1-0.7^2=0.51,$$

$$P(A_2)=0.8^3=0.512,$$

$$P(B|A_2)=1-0.7^3=0.657,$$

故目标被炸毁的概率为

$$P(B)=\sum_{i=0}^{2}P(A_i)P(B|A_i)=0.476\ 5$$

若目标被炸毁，三架飞机均到达目标上空的概率为

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{2}P(A_i)P(B|A_i)}$$

$$= \frac{0.512 \times 0.657}{0.4765} = 0.706.$$

【独立事件的判断与应用问题】关于事件独立性的问题一般表现为两个方面：

(1)**对事件独立性的判断和证明.证明** A **与** B **独立应严格按事件独立的定义进行,即证明** $P(AB)=P(A)P(B)$.**判断** A **与** B **是否独立则可结合独立性的含义:无论** B **是否发生,**A **发生的概率不变,即**

$$P(A|B) = P(A|\overline{B}) = P(A).$$

(2)**利用独立性计算概率.若** $A_1,A_2,\cdots,A_n$ **相互独立,则**

$$P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i),$$

$$P(\bigcup_{i=1}^{n} A_i) = P(\overline{\bigcap_{i=1}^{n} \overline{A}_i}) = 1 - \prod_{i=1}^{n} [1 - P(A_i)].$$

事件的独立性对于取逆、交、并、差运算保持不变,即若 A,B,C **相互独立,则**

$$P(A\overline{B}) = P(A)P(\overline{B}),\quad P(\overline{A}\,\overline{B}) = P(\overline{A})P(\overline{B}),$$

$$P(A(BC)) = P(A)P(BC),$$

$$P(A(B\cup C)) = P(A)P(B\cup C),$$

$$P(A(B-C)) = P(A)P(B-C).$$

例　**设** A,B **为两事件,**$0<P(A)<1$,**证明** A **与** B **相互独立的充分必要条件是** $P(B|A)=P(B|\overline{A})$.

证　**由** $0<P(A)<1$ **知** $P(B|A)$ **和** $P(B|\overline{A})$ **都存在.**

(1)**必要性**　**由** A **与** B **独立知** $\overline{A}$ **与** B **也独立,因此**

$$P(B|A)=P(B)=P(B|\overline{A}).$$

(2)**充分性**　**由**$P(B|A)=P(B|\overline{A})$,**得**

$$\frac{P(AB)}{P(A)}=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(B)-P(AB)}{1-P(A)},$$

$$P(AB)[1-P(A)]=P(A)P(B)-P(A)P(AB),$$

$$P(AB)=P(A)P(B),$$

即A**与**B**独立**.

第16章　随机变量及其概率分布

大纲要求

理解随机变量的概念；理解分布函数

$$F(x)=P\{X\leqslant x\}(-\infty<x<+\infty)$$

的概念及性质；会计算与随机变量相联系的事件的概率. 理解离散型随机变量及其概率分布的概念，掌握 0-1 分布、二项分布、超几何分布、泊松分布及其应用. 了解泊松定理的结论和应用条件，会用泊松分布近似表示二项分布. 理解连续型随机变量及其概率密度的概念，掌握均匀分布、正态分布 $N(\mu,\sigma^2)$、指数分布及其应用，其中参数为 $\lambda(\lambda>0)$ 的指数分布的密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

会求随机变量函数的分布.

知识点·点拨

16.1　随机变量及其分布函数

【随机变量】设 Ω 为随机试验的样本空间，X 为定义在 Ω 上的实函数，即对每一样本点 $\omega\in\Omega$，$X(\omega)$ 为一实数，则称 X 为一个随机变量.

【分布函数】设 X 为一个随机变量，则称实函数

$$F(x)=P(X\leqslant x),\ -\infty<x<+\infty$$

为 X 的分布函数.

【分布函数的性质】

1. **有界性**　$0\leqslant F(x)\leqslant 1$，**且**

$$F(-\infty)=0,\quad F(+\infty)=1.$$

2. **单调不减性**　**若** $x_1<x_2$，**则**

$$F(x_1)\leqslant F(x_2).$$

3. **右连续性**　$F(x+0)=\lim\limits_{t\to x^+}F(t)=F(x).$

〖分布函数的鉴别〗分布函数的有界性，单调不减性和右连续性是一个实函数成为分布函数的充分必要条件，即此三条可作为判断一个实函数是否为分布函数的试金石.

例　在下述函数中，可以作为某个随机变量的分布函数是(　　).

(A) $F(x)=\dfrac{1}{x^2}$

(B) $F(x)=\dfrac{1}{\pi}\arctan x+\dfrac{1}{2}$

(C) $F(x)=\begin{cases}\dfrac{1}{2}(1-\mathrm{e}^{-x}), & x>0,\\ 0, & x\leqslant 0\end{cases}$

(D) $F(x)=\displaystyle\int_{-\infty}^{x}f(t)\mathrm{d}t$，**其中** $\displaystyle\int_{-\infty}^{+\infty}f(t)\mathrm{d}t=1$

分析　(A)**不是单调函数；由于** $f(x)$ **不一定非负，**(D)**也不是单调函数；**(C)**不满足** $F(+\infty)=1$；**只有**(B)**满足分布函数的三条性质. 故本题应选**(B).

〖分布函数与概率〗由分布函数的定义知

$$P(a<X\leqslant b)=F(b)-F(a)\quad(a<b),$$

$$P(X=a)=F(a)-F(a-0).$$

因此，由 X **的分布函数可以确定** X **在任何区间中取**

值的概率,由此还可以看出分布函数 $F(x)$ 是用其函数值上升幅度的大小来描述随机变量 X 的概率分布的.

例 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ \dfrac{1}{8}, & x=-1,\\ ax+b, & -1<x<1,\\ 1, & x\geqslant 1,\end{cases}$$

且已知 $P\{X=1\}=\dfrac{1}{4}$,则 $a=$ ________,$b=$ ________.

解 由 $F(x)$ 在右连续性及已知条件,有

$$\begin{cases}F(-1+0)=-a+b=\dfrac{1}{8},\\ F(1)-F(1-0)=1-(a+b)=\dfrac{1}{4},\end{cases}$$

解此方程组得

$$a=\frac{5}{16},\quad b=\frac{7}{16}.$$

16.2 离散型随机变量

【离散型随机变量】 若随机变量 X 所有可能的取值为有限个或可列无穷个,则称 X 为离散型随机变量.

【分布列(律)】 $P(X=x_i)=p_i,\quad i=1,2,\cdots.$

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

〖分布列与分布函数的关系〗 分布函数 $F(x)$ 与

分布列 p_i 都描述了离散型随机变量 X 的概率分布. 它们的区别在于：分布列描述 X 在各点上取值的概率，分布函数 $F(x)$ 则描述 X 在 x 左边取值的累积概率. 两者的关系为 $F(x)=\sum\limits_{x_i\leqslant x}p_i$.

【分布列的性质】

1. $p_i\geqslant 0, i=1,2,\cdots$；　2. $\sum\limits_{i}p_i=1$.

〖点拨〗$p_i\geqslant 0(i=1,2,\cdots)$ 和 $\sum\limits_{i}p_i=1$ 是一数列 $\{p_i\}$ 成为分布列的充分必要条件. 它常用来确定分布中的待定常数的条件.

【常用分布】

名　称	记　号	分布列(律)
0-1 分布	$B(1,p)$	$P(X=1)=p$, $P(X=0)=1-p$
二项分布	$B(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k}$, $k=0,1,\cdots,n, 0\leqslant p\leqslant 1$
超几何分布	$H(N,M,n)$	$P(X=k)=\dfrac{C_N^k C_M^{n-k}}{C_{N+M}^n}, k\leqslant N$, $n-k\leqslant M$ 为非负整数
泊松分布	$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,2,\cdots,\lambda>0$

〖实际背景〗对常用分布应结合随机变量产生的背景来掌握.

(1)若 X 为 n 重贝努利试验中事件 A 发生的次数,则 $X\sim B(n,p)$,其中 p 为每次试验中 A 发生的概率. 0-1 分布是二项分布 $n=1$ 时的特例.

(2)设袋中有 N 个红球和 M 个白球,从中任取 n 个球,X 为取到的红球个数,则 $X\sim H(N,M,n)$.

(3)设 X 为一段时间内,事件 A 发生的次数,则 $X\sim P(\lambda)$.

例　一射手对同一目标独立地进行 4 次射击. 若他至少命中一次的概率为 80/81,则他至多命中一次的概率为________.

解　设 X 为该射手命中的次数,则 $X\sim B(4,p)$,其中 p 为命中率. 由已知条件,他至少命中的一次的概率为

$$P(X\geqslant 1)=1-(1-p)^4=\frac{80}{81},$$

因而解得 $p=\dfrac{2}{3}$,故他至多命中一次的概率为

$$P(X\leqslant 1)=(1-p)^4+C_4^1p(1-p)^3=\frac{1}{9}.$$

【泊松定理及其应用】若数列 $\{p_n\}$ 满足

$$\lim_{n\to\infty}np_n=\lambda>0,$$

则对任何非负整数 k,有

$$\lim_{n\to\infty}C_n^kp_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

由此结论,对 $X\sim B(n,p)$,当 $n>10,p<0.1$ 时,可近似计算

$$P(X=k)\approx\frac{(np)^k}{k!}e^{-np},\quad k=0,1,2,\cdots.$$

16.3　连续型随机变量

【连续型随机变量及其概率密度】若存在非负实函数 $f(x)$ 使随机变量 X 的分布函数

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}x,\quad -\infty<x<+\infty,$$

则称 X 为连续型随机变量，且称 $f(x)$ 为 X 的概率密度函数.

〖密度函数与分布函数〗连续型随机变量 X 的概率密度函数 $f(x)$ 描述了 X 在 x 附近取值的概率：

$$P(x-\varepsilon<X<x+\varepsilon)=\int_{x-\varepsilon}^{x+\varepsilon} f(t)\mathrm{d}t=f(\xi)\cdot 2\varepsilon,$$

$$x-\varepsilon<\xi<x+\varepsilon.$$

而其分布函数描述了 X 在 $(-\infty,x]$ 取值的累积概率.

【性质】连续型随机变量 X 的分布函数 $F(x)$ 和概率密度函数 $f(x)$ 有如下性质：

(1) $F(x)$ 为连续函数；

(2) 对任何实数 a，$P(X=a)=0$；

(3) $P(X\in(a,b))=\int_{a}^{b} f(x)\mathrm{d}x$；

(4) 对 $f(x)$ 的连续点 x，有 $F'(x)=f(x)$；

(5) $f(x)\geqslant 0$，$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$.

〖点拨〗$f(x)\geqslant 0$，$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$ 是实函数 $f(x)$ 成为概率密度函数的充分必要条件，它常用来确定概率密度的待定常数.

【常用分布】

名　称	记　号	概率密度函数
均匀分布	$U[a,b]$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其他},\end{cases}$ $-\infty<a<b<+\infty$
正态分布	$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\,\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty<\mu<+\infty,\sigma>0$
指数分布	$E(\lambda)$	$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$ $\lambda>0$

〖几何形态与实际背景〗对常用分布可结合其分布形态来掌握.

(1)均匀分布的“均匀”意味着其概率分布在$[a,b]$内无高低起伏,因而其概率密度函数图像表现为水平线. X 在$[c,d]$内取值的概率表现为仅与区间长度$(d-c)$有关而与区间的位置无关$(a\leqslant c<d\leqslant b)$.

(2)正态分布的“正态”描述了概率的中间大两边小的正常分布状态,其概率密度函数图像表现为以 $x=\mu$ 为中心的对称曲线. 因而其数学期望 μ 又称为位置参数. 方差σ^2 反映了该曲线的平坦程度,因而也称 σ^2 为形状参数. 性质 $\Phi(-x)=1-\Phi(x)$也是其对称性的体现.

(3)指数分布因其概率密度简洁的指数函数形

式而得名,它通常用来描述寿命,即正常工作的概率随时间的增加而呈负指数函数衰减的规律. 指数分布还具有“无记忆”性质:

$$P(X>t_1+t_2|X>t_1)=P(X>t_2),\quad t_1,t_2>0.$$

例　已知随机变量 X 的概率密度函数

$$f(x)=\begin{cases}Ae^{-x}, & x>\lambda,\\ 0, & x\leqslant\lambda\end{cases}(\lambda>0,A\text{ 为常数}),$$

则概率 $P(\lambda<X<\lambda+a)(a>0)$ 的值

(A) 与 a 无关,随 λ 的增大而增大

(B) 与 a 无关,随 λ 的增大而减小

(C) 与 λ 无关,随 a 的增大而增大

(D) 与 λ 无关,随 a 的增大而减小

解　$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{\lambda}^{+\infty}Ae^{-x}\mathrm{d}x=Ae^{-\lambda}=1$,求得 $A=e^{\lambda}$. 因此概率

$$P(\lambda<X<\lambda+a)=\int_{\lambda}^{\lambda+a}e^{\lambda-x}\mathrm{d}x=-e^{-\lambda-x}\Big|_{\lambda}^{\lambda+a}$$
$$=1-e^{-a}$$

与 λ 无关,且随 a 的增大而增大. 故选(C).

本题还可以结合指数分布来分析. 令 $Y=X-\lambda$,则 Y 的概率密度函数为

$$f_Y(y)=\begin{cases}Ce^{-y}, & y>0,\\ 0, & y\leqslant 0\end{cases}\quad(C=Ae^{-\lambda}).$$

显然 Y 服从参数为 1 的指数分布,即 $C=1$,从而概率

$$P(\lambda<X<\lambda+a)=P(Y<a)=F_y(a)$$

与 λ 无关,且为 a 的单调增函数.

【正态分布的性质】设 $X\sim N(0,1)$,称 X 服从标准正

态分布,其概率密度函数和分布函数为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\quad \Phi(x)=\int_{-\infty}^{x}\varphi(t)\mathrm{d}t.$$

若 $X\sim N(\mu,\sigma^2)$,则

(1) $P(a<X<b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$.

(2) $P(|X-\mu|<a\sigma)=2\Phi(a)-1$,

$P(|X-\mu|>a\sigma)=2[1-\Phi(a)]$.

(3) $Y=aX+b\sim N(a\mu+b,a^2\sigma^2)(a\neq 0)$.

16.4　随机变量函数的分布

【随机变量函数的分布】设 X 是一个随机变量,$y=g(x)$是一个连续函数,则 $Y=g(x)$也是一个随机变量,称 Y 的分布为 X 的函数的分布.

【离散型随机变量函数的分布】设 X 的分布列为

$$P(X=x_i)=p_i,\quad i=1,2,\cdots,$$

则 $Y=g(x)$为离散型随机变量,记

$$y_i=g(x_i)\quad (i=1,2,\cdots).$$

(1)若 $y_i=g(x_i)$的值互不相等,则 Y 的分布列为

$$P(Y=y_i)=p_i\quad (i=1,2,\cdots).$$

(2)若各 y_i 的值有相等的情形,则应将那些相等的值分别合并,并将相等 y_i 值对应的概率相加,就得到 Y 的分布列.

【连续型随机变量函数的分布】设 X 的概率密度函数为 $f_X(x)$,$y=g(x)$为连续函数,则 $Y=g(X)$也是连续型随机变量.求 Y 的概率密度函数有两种方法.

(1)公式法.若 $y=g(x)$严格单调,且其反函数

$x=g^{-1}(y)$可导，则$Y=g(X)$的概率密度函数为

$$f_Y(y)=\begin{cases} f_X[g^{-1}(y)]\,|g^{-1\prime}(y)|, & \alpha<y<\beta, \\ 0, & 其他. \end{cases}$$

其中　　$\alpha=\min\{g(-\infty),g(+\infty)\}$，

$\beta=\max\{g(-\infty),g(+\infty)\}$.

(2)定义法. 若$y=g(x)$不是严格单调函数，则先计算Y的分布函数$F_Y(y)=P(g(X)\leqslant y)$. 然后由$f_Y(y)=F'_Y(y)$求出$Y$的概率密度函数.

〖点拨〗连续型随机变量函数的分布的情形(2)与离散型随机变量的分布的情形(2)本质上是一样的，即函数$y=f(x)$有多个自变量的值对应相同因变量的值. 因此所计算的分布函数往往是多个概率的和

$$F_Y(y)=P(g(X)\leqslant y)=\sum_i P(X\in B_i),\quad B_i\subseteq \mathbf{R}.$$

当定义法用于$y=g(x)$严格单调的情形时，就是对公式法的推导.

例　设随机变量$X\sim N(0,1)$，求$Y=2X^2+1$的概率密度函数.

解　此题可用两种方法求解.

(1)定义法. Y的分布函数为

$$F_Y(y)=P(2X^2+1\leqslant y)=P\left(X^2\leqslant\frac{y-1}{2}\right),$$

当$y<1$时，$F_Y(y)=0$，当$y\geqslant 1$时，

$$F_Y(y)=P\left(-\sqrt{\frac{y-1}{2}}\leqslant X\leqslant\sqrt{\frac{y-1}{2}}\right)$$

$$=\int_{-\sqrt{\frac{y-1}{2}}}^{\sqrt{\frac{y-1}{2}}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x$$

$$=2\int_0^{\sqrt{\frac{y-1}{2}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

故Y的概率密度函数为

$$f_Y(y)=F'_Y(t)=\begin{cases}0, & y<1,\\ \dfrac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}, & y\geqslant 1.\end{cases}$$

(2)公式法. 由于$y=2x^2+1$在$(-\infty,+\infty)$不是严格单调函数,故将整个区间分为$(-\infty,0)$和$(0,+\infty)$两个单调区间,则在$(-\infty,0)$内有反函数$x=-\sqrt{\frac{y-1}{2}}(y\geqslant1)$,在$(0,+\infty)$内有反函数$x=\sqrt{\frac{y-1}{2}}(y\geqslant1)$,且均可导. 当$y\geqslant1$时,由公式得$Y$的概率密度函数为

$$\begin{aligned}f_Y(y)&=\varphi\left(-\sqrt{\frac{y-1}{2}}\right)\left|\left(-\sqrt{\frac{y-1}{2}}\right)'\right|\\&\quad+\varphi\left(\sqrt{\frac{y-1}{2}}\right)\left|\left(\sqrt{\frac{y-1}{2}}\right)'\right|\\&=2\frac{1}{\sqrt{2\pi}}e^{-\frac{y-1}{4}}\times\frac{1}{2\sqrt{2}\sqrt{y-1}},\end{aligned}$$

其中,$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$为$X$的概率密度函数.

由于$Y=2X^2+1\geqslant1$,所以当$y<1$时$f_Y(y)=0$,即

$$f_Y(y)=\begin{cases}0, & y<1,\\ \dfrac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}, & y\geqslant 1.\end{cases}$$

常见题型·应对

【分布常数的确定问题】确定分布中的待定常数有如下原则.

(1)**若题目给出的是分布函数** $F(x)$,**则可由** $F(-\infty)=0, F(+\infty)=1$ **和** $F(x+0)=F(x)$ **来确定待定常数. 必要时,可由** $0\leqslant F(x)\leqslant 1$ **和** $F(x)$ **单调不减来约束待定常数的取值范围.**

(2)**若题目给出的是分布列** $P(X=x_i)=p_i(i=1,2,\cdots)$,**则由** $\sum\limits_i p_i=1$ **来确定待定常数.**

(3)**若题目给出的是概率密度函数** $f(x)$,**则由** $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$ **来确定待定常数.**

例　**设离散型随机变量** X **的分布列为**

$$P(X=k)=cp^k,\ k=0,1,2,\cdots,$$

已知 $P(X=0)=0.1$,**则** $p=$________.

解　**因为** $P(X=0)=c=0.1$,**由**

$$\sum_{k=0}^{\infty} cp^k=\frac{c}{1-p}=\frac{0.1}{1-p}=1$$

得

$$p=0.9.$$

【正态分布的参数与概率】正态分布是应用最广泛的分布,熟练掌握其分布形式和其参数的作用,可以大大提高解题的速度.

(1)**如果连续型随机变量** X **的概率密度函数有形式**

$$f(x)=k\mathrm{e}^{-ax^2+bx+c},$$

其中 k,a,b,c **为常数,且** $a>0$,**则** X **服从正态分布.**

例 1　设连续型随机变量的概率密度函数为 $f(x)=ae^{-\frac{1}{2}x^2+bx}$，且 $EX=3$，则 $a=$________，$b=$________.

解　将 $f(x)$ 化为正态密度的规范形式

$$f(x)=ae^{-\frac{(x-b)^2}{2}+\frac{b^2}{2}}=ae^{\frac{b^2}{2}}e^{-\frac{(x-b)^2}{2\times 1}}.$$

由题设条件知 $b=EX=3$，$ae^{\frac{9}{2}}=\dfrac{1}{\sqrt{2\pi}\times 1}$，从而

$$a=\frac{1}{\sqrt{2\pi}}e^{-\frac{9}{2}}.$$

(2)若 $X\sim N(\mu,\sigma^2)$，则标准化随机变量 $Y=\dfrac{X-\mu}{\sigma}\sim N(0,1)$. 这意味着 Y 在某确定区域 D 内取值的概率与 μ 和 σ^2 的取值无关.

例 2　设随机变量 $X\sim N(\mu,1)$，则概率 $P(X>\mu+3)$ 随着 μ 增加而(　　).

(A) 增加　　(B) 减小

(C) 不变　　(D) 不能确定

分析　因为

$$P(X>\mu+3)=P(X-\mu>3)=1-\Phi(3)$$

与 μ 无关，故应选(C).

【随机变量函数的综合应用】此类问题通常由大段文字描述，解题时应注意下面几个环节：

(1)设置合理的随机变量是解决实际问题的关键. 随机变量 X 是一个随机的数，不能把它设成某个事件发生. 事实上，X 的不同取值可以表示一切可能的事件.

(2)此类问题中所涉随机变量大多数服从常用分布.因此,熟练掌握这些分布的表达式和实际背景是解题的关键,特别是二项分布和正态分布.

(3)随机变量函数的分布可将 $y=g(x)$ 推广到非连续的情形.比如,连续型随机变量 X 的概率密度函数为 $f(x)$,分段函数

$$g(x)=\begin{cases} y_0, & x\leqslant a_1, \\ y_i, & a_{i-1}<x\leqslant a_i, \quad i=2,3,\cdots,k-1, \\ y_k, & x>a_{k-1}, \end{cases}$$

$$-\infty=a_0<a_1<a_2<\cdots<a_{k-1}<a_k=+\infty,$$

则 $Y=g(X)$ 为离散型随机变量,其分布列为

$$\begin{aligned} P(Y=y_i) &= P(a_{i-1}<X\leqslant a_i) \\ &= \int_{a_{i-1}}^{a_i} f(x)\mathrm{d}x, \quad i=1,\cdots,k. \end{aligned}$$

例 假设一部机器一天内发生故障的概率为0.2,机器发生故障时全天停止工作.若一周5个工作日内无故障,可产生利润10万元;发生一次故障可产生利润5万元;发生两次故障产生零利润;发生三次或三次以上故障产生负利润2万元,求一周内产生利润的分布律.

解 设 X 为一周5天内机器发生故障的天数,则 $X\sim B(5,0,2)$.因此

$$P(X=0)=0.8^5=0.328,$$

$$P(X=1)=C_5^1\times 0.2\times 0.8^4=0.410,$$

$$P(X=2)=C_5^2\times 0.2^2\times 0.8^3=0.205,$$

$$P(X\geqslant 3)=1-P(X=0)-P(X=1)$$

$$-P(X=2)=0.057.$$

再设 Y 为一周产生的利润(单位:万元),则根据题设,有

$$Y=g(X)=\begin{cases}10, & X=0,\\ 5, & X=1,\\ 0, & X=2,\\ -2, & X\geqslant 3,\end{cases}$$

故 Y 的分布律为

Y	10	5	0	-2
P	0.328	0.410	0.205	0.057

第 17 章　多维随机变量及其分布

大纲要求

理解多维随机变量的概念，理解多维随机变量的分布的概念和性质. 理解二维离散型随机变量的概率分布，边缘分布和条件分布；理解二维连续性随机变量的概率密度、边缘密度和条件密度. 会求与二维随机变量相关事件的概率. 理解随机变量的独立性及不相关性的概念，掌握随机变量相互独立的条件. 掌握二维均匀分布，了解二维正态分布的概率密度，理解其中参数的概率意义. 会求多个相互独立随机变量的简单函数的分布.

知识点 · 点拨

17.1　二维随机变量的联合分布

【二维随机变量】在同一样本空间 Ω 上定义的两个随机变量 X、Y 构成的向量 (X,Y) 称为二维随机变量.

【联合分布函数】称二元实函数

$$F(x,y)=P(X\leqslant x,Y\leqslant y)$$

为二维随机变量 (X,Y) 的联合分布函数.

〖联合分布函数的求法〗求联合分布函数是一项细致的工作，在二维情况下是对一切 $(x,y)\in\mathbf{R}^2$ 计算事件 $\{X\leqslant x,Y\leqslant y\}$ 的概率值. 通常是让 x 和 y 从 $-\infty$ 向 $+\infty$ 变化，将区域 $\{X\leqslant x,Y\leqslant y\}$ 内的概率

求和(离散型)或积分(连续型).

【离散型随机向量的联合概率分布】若二维随机变量(X,Y)仅取有限个或可列无限个值,则称(X,Y)为二维离散型随机变量.其联合概率分布为

$$P(X=x_i,Y=y_i)=p_{ij},\quad i,j=1,2,\cdots$$

或

$X \backslash Y$	y_1	$\cdots$	y_j	$\cdots$
x_1	p_{11}	$\cdots$	p_{1j}	$\cdots$
$\vdots$	$\vdots$		$\vdots$	
x_i	p_{i1}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$		$\vdots$	

p_{ij}具有性质:$p_{ij}\geqslant 0$, $\sum\limits_{i,j} p_{ij}=1$.

【连续型随机向量的联合概率密度】若有二元非负实函数$f(x,y)$,使二维随机变量(X,Y)的联合分布函数

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\mathrm{d}v,$$

则称(X,Y)为二维连续型随机变量,称$f(x,y)$为其联合概率密度函数.它具有如下基本性质:

(1) $f(x,y)\geqslant 0$, $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y=1$;

(2) $P((X,Y)\in G)=\iint\limits_{G} f(x,y)\mathrm{d}x\mathrm{d}y$;

(3) 对$f(x,y)$的连续点有

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}.$$

〖二维分布的判定法〗同一维分布类似，二维分布的基本性质是判断二元函数能否成为概率分布和确定分布中待定常数的根据.

判断对象	充分必要条件
联合分布函数 $F(x,y)$	1° $0\leqslant F(x,y)\leqslant 1, F(-\infty,y)=F(x,-\infty)=0$, $F(+\infty,+\infty)=1$; 2° 若 $x_1<x_2, y_1<y_2$ 则 $F(x_1,y)\leqslant F(x_2,y), F(x,y_1)\leqslant F(x,y_2)$; 3° $\lim\limits_{x\to a^+} F(x,y)=F(a,y)$, $\lim\limits_{y\to b^+} F(x,y)=F(x,b)$; 4° 若 $x_1<x_2, y_1<y_2$，则 $F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\geqslant 0$
联合分布列 p_{ij}	$p_{ij}\geqslant 0, \sum\limits_{i,j} p_{ij}=1$
联合密度函数 $f(x,y)$	$f(x,y)\geqslant 0, \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y=1$

例　设 $f_1(x,y)$ 与 $f_2(x,y)$ 都是联合概率密度函数，要使 $f(x,y)=af_1(x,y)+bf_2(x,y)$ 也是联合概率密度函数，则当且仅当 a,b 满足条件

(A) $a+b\leqslant 1$　　(B) $0\leqslant a\leqslant 1, 0\leqslant b\leqslant 1$

(C) $a>0$ 且 $b>0$　　(D) $a\geqslant 0, b\geqslant 0$ 且 $a+b=1$

解　因(A)不能保证 $f(x,y)\geqslant 0$，(B)和(C)不能保证 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y=a+b=1$，而(D)同时满足 $f(x,y)$ 的这两条充要条件，故应选(D).

【二维均匀分布】若二维随机变量 (X,Y) 的联合概

率密度为

$$f(x,y)=\begin{cases}1/S_G, & (x,y)\in G,\\ 0, & \text{其他}.\end{cases}$$

其中,G 为平面上某一区域,S_G 为其面积,则称 (X,Y) 服从区域 G 上的均匀分布.

【二维正态分布】若二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{q(x,y)}{2(1-\rho^2)}\right\},$$

其中 $$q(x,y)=\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2.$$

则称 (X,Y) 服从参数为 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 的二维正态分布,记作 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 其中 μ_1 和 μ_2 分别为 X 和 Y 的数学期望,σ_1^2 和 σ_2^2 分别为 X 和 Y 的方差,ρ 为 X 与 Y 的相关系数.

17.2　边缘分布和条件分布

【边缘分布】二维随机变量 (X,Y) 中某一个随机变量的分布称为边缘分布.

设 $F(x,y)$ 为 (X,Y) 的联合分布函数,则

$$F_X(x)=\lim_{y\to+\infty}F(x,y),\quad F_Y(y)=\lim_{x\to+\infty}F(x,y)$$

分别为 X 和 Y 的边缘分布函数. 设二维离散型随机变量 (X,Y) 的联合分布为

$$P(X=x_i,Y=y_i)=p_{ij}(i,j=1,2,\cdots),$$

则 $$p_{i.}=P(X=x_i)=\sum_j p_{ij}\quad(i=1,2,\cdots),$$

$$p_{.j}=P(Y=y_j)=\sum_i p_{ij}\quad (j=1,2,\cdots)$$

分别为 X 和 Y 的边缘分布列. 设二维连续型随机变量 (X,Y) 的联合概率密度函数为 $f(x,y)$，则

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$$

分别为 X 和 Y 的边缘概率密度函数.

【条件分布】在二维随机变量 (X,Y) 中，已知一个随机变量的取值时，另一个随机变量的分布称为条件分布.

设二维随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，X 和 Y 的边缘分布为 $F_X(x)$ 和 $F_Y(y)$，则

$$F_{X|Y}(x|y)=\lim_{\Delta y\to 0}\frac{F(x,y+\Delta y)-F(x,y-\Delta y)}{F_Y(y+\Delta y)-F_Y(y-\Delta y)}$$

为已知 $Y=y$ 时 X 的条件分布函数. 类似地有已知 $X=x$ 时 Y 的条件分布函数

$$F_{Y|X}(y|x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x,y)-F(x-\Delta x,y)}{F_X(x+\Delta x)-F_X(x-\Delta x)}.$$

设二维离散型随机变量 (X,Y) 的联合分布为 $P(X=x_i,Y=y_j)=p_{ij}(i,j=1,2,\cdots)$. X 和 Y 的边缘分布分别为 $P(X=x_i)=p_{i.}\ (i=1,2,\cdots)$ 和 $P(Y=y_j)=p_{.j}\ (j=1,2,\cdots)$，则

$$P(X=x_i|Y=y_j)=\frac{p_{ij}}{p_{.j}}\quad (i=1,2,\cdots),$$

$$P(Y=y_j|X=x_i)=\frac{p_{ij}}{p_{i.}}\quad (j=1,2,\cdots)$$

分别为已知 $Y=y_j$ 时 X 的条件分布列和已知 $X=x_i$ 时 Y 的条件分布列.

设二维连续型随机变量 (X,Y) 的联合概率密度函数为 $f(x,y)$，X 和 Y 的边缘概率密度函数分别为 $f_X(x)$ 和 $f_Y(y)$，则

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)},$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$

分别为已知 $Y=y$ 时 X 的条件概率密度函数和已知 $X=x$ 时 Y 的条件概率密度函数.

〖点拨〗边缘分布和条件分布都是(一维)随机变量的概率分布，同样具有概率分布的基本性质：

$$p_{i\cdot}\geqslant 0,\quad \sum_i p_{i\cdot}=1;$$

$$(p_{ij}/p_{\cdot j})\geqslant 0,\quad \sum_i (p_{ij}/p_{\cdot j})=1;$$

$$f_X(x)\geqslant 0,\quad \int_{-\infty}^{+\infty} f_X(x)\mathrm{d}x=1;$$

$$f_{X|Y}(x|y)\geqslant 0,\quad \int_{-\infty}^{+\infty} f_{X|Y}(x|y)\mathrm{d}x=1.$$

它们可用来确定分布中的待定值和检验计算结果的正确性.

〖联合分布、边缘分布、条件分布三者的关系〗由 (X,Y) 的联合分布可以求出 X 和 Y 的边缘分布，反之，由 X 和 Y 的边缘分布一般不能确定 (X,Y) 的联合分布. 如果已知一个随机变量的边缘分布和另一个随机变量的条件分布，则可求得 (X,Y) 的联合

分布，事实上

$$p_{ij}=p_{i\cdot}(p_{ij}/p_{i\cdot})=p_{\cdot j}(p_{ij}/p_{\cdot j}),$$

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y),$$

这是概率的乘法公式在随机变量概率分布上的表现.

〖边缘分布和条件分布的几何表示〗从几何上看，边缘分布是将平面上的概率分布累积到一个坐标轴上，条件分布则表现为平面上概率分布在某直线上的截面. 从而不难理解下面结论.

(1)二维正态分布的边缘分布和条件分布仍为正态分布，即若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

$$X\sim N(\mu_1,\sigma_1^2),$$

$$(X|Y=y)\sim N\left(\mu_1+\frac{\sigma_1}{\sigma_2}\rho(y-\mu_2),(1-\rho^2)\sigma_1^2\right)$$

$$Y\sim N(\mu_2,\sigma_2^2),$$

$$(Y|X=x)\sim N\left(\mu_2+\frac{\sigma_2}{\sigma_1}\rho(x-\mu_1),(1-\rho^2)\sigma_2^2\right).$$

(2)区域 G 上的二维均匀分布，边缘分布一般不是均匀分布(除非 G 为各边平行于坐标轴的矩形)，而条件分布仍为均匀分布.

17.3　独　立　性

【独立性的定义】设 $X_1,X_2,\cdots,X_n$ 为 n 个随机变量，若对任意实数集合 $A_1,A_2,\cdots,A_n$，有

$$P(X_1\in A_1,X_2\in A_2,\cdots X_n\in A_n)$$
$$=P(X_1\in A_1)P(X_2\in A_2)\cdots P(X_n\in A_n),$$

则称 $X_1,X_2,\cdots,X_n$ 相互独立.

〖对随机变量独立性的理解〗X 与 Y 独立的概念有下面两种理解:

(1)X 所能描述的任一事件与 Y 所能描述的任一事件相互独立.

(2)无论 Y 取何值,X 的条件分布不会改变,同样 Y 的条件分布也不随 X 的取值不同而改变.事实上,这时的条件分布就是边缘分布.

【独立性的判定条件】按 17.1 节和 17.2 节的记号,有

类　型	$X_1,X_2,\cdots,X_n$ 相互独立的条件	X 与 Y 相互独立的条件
一般型	$F(x_1,\cdots,x_n)$ $=F_{X_1}(x_1)\cdots F_{X_n}(x_n)$	$F(x,y)$ $=F_X(x)F_Y(y)$
离散型	$P(X_1=x_1,\cdots,X_n=x_n)$ $=P(X_1=x_1)\cdots P(X_2=x_2)$	$p_{ij}=p_{i.}\,p_{.j}$
连续型	$f(x_1,\cdots,x_n)$ $=f_{X_1}(x_1)\cdots f_{X_n}(x_n)$	$f(x,y)$ $=f_X(x)f_Y(y)$

【相关性】若 X 与 Y 的协方差

$$\operatorname{cov}(X,Y)=E[(X-EX)(Y-EY)]=0,$$

则称 X 与 Y 不相关.

〖对随机变量相关性的理解〗X 与 Y 的协方差 $\operatorname{cov}(X,Y)$和相关系数 ρ_{XY}反映了 X 与 Y 具有线性关系的程度.所谓线性关系可以这样理解:

(1)若 $\rho_{XY}>0$,当 X 值偏大时,Y 值偏大的概率增大;当 X 值偏小时,Y 值偏小的概率增大.这时称

X 与 Y 正相关.

(2)若 $\rho_{XY}<0$,当 X 值偏大时,Y 值偏小的概率增大;当 X 值偏小时,Y 值偏大的概率增大. 这时称 X 与 Y 负相关.

(3)若 $\rho_{XY}=0$,则没有上述变化趋势. 这时称 X 与 Y 不相关.

【不相关与独立的关系】若 X 与 Y 独立,则 X 与 Y 不相关;反之,若 X 与 Y 不相关,则 X 与 Y 不一定独立.

〖两个特例〗一般情况下,X 与 Y 不相关不能导出 X 与 Y 独立,但在下列特殊情况下,由不相关性可以导出独立性.

(1)若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,则 X 与 Y 相互独立的充分必要条件是 X 与 Y 不相关,即 $\rho=0$.

(2)若 X 和 Y 均为两点分布,则 X 与 Y 相互独立的充分必要条件是 X 与 Y 不相关.(见下例)

例　设 $X\sim B(1,p_1)$,$Y\sim B(1,p_2)$,X 与 Y 不相关,试证明 X 与 Y 相互独立.

证　记 (X,Y) 的联合分布为

$X \backslash Y$	0	1	$p_{i\cdot}$
0	p_{00}	p_{01}	$1-p_1$
1	p_{10}	p_{11}	p_1
$p_{\cdot j}$	$1-p_2$	p_2	1

则
$$EX=p_1,\quad EY=p_2,$$
$$E(XY)=1\times p_{11}+0\times(1-p_{11})=p_{11},$$

由条件

$\text{cov}(X,Y)=E(XY)-(EX)(EY)=p_{11}-p_1p_2=0$,

即
$$p_{11}=p_1p_2, \tag{①}$$

注意到
$$P(X=1)=p_{10}+p_{11}=p_1,$$
$$P(Y=1)=p_{01}+p_{11}=p_2.$$

由式①导出
$$p_{10}=p_1-p_1p_2=p_1(1-p_2), \tag{②}$$
$$p_{01}=p_2-p_1p_2=(1-p_1)p_2, \tag{③}$$
$$\begin{aligned}p_{00}&=1-p_{11}-p_{10}-p_{01}\\&=1-p_1p_2-p_1(1-p_2)-(1-p_1)p_2\\&=(1-p_1)(1-p_2).\end{aligned} \tag{④}$$

式①～④说明X与Y相互独立.

17.4 二维随机变量函数的分布

【一般函数的分布】设二元函数$z=g(x,y)$连续,则对二维离散型随机变量(X,Y),$Z=g(X,Y)$为一维离散型随机变量,其概率分布为

$$P(Z=z_k)=\sum_{g(x_i,y_j)=z_k}p_{ij},\quad k=1,2,\cdots$$

其中$p_{ij}=P(X=x_i,Y=y_i)$为(X,Y)的联合分布.对二维连续型随机变量(X,Y),$Z=g(X,Y)$为一维连续型随机变量,其概率密度为

$$f_Z(z)=F'_Z(z)=\frac{\mathrm{d}}{\mathrm{d}z}\left[\iint_{g(x,y)\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y\right],$$

其中$f(x,y)$为(X,Y)的联合概率密度.

【和的概率分布】设X和Y均为取整数值的随机变量,则$Z=X+Y$的概率分布为

$$P(Z=k)=\sum_i P(X=i,Y=k-i)$$
$$=\sum_j P(X=k-j,Y=j),$$

k 为整数.

设二维连续型随机变量(X,Y)的联合概率密度为 $f(x,y)$,则$Z=X+Y$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x$$
$$=\int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y.$$

这个公式称为卷积公式.

〖可加性〗若两个独立同类分布的随机变量之和仍为该类分布,则称该类分布具有可加性. 下面重要分布具有可加性:

(1)设$X\sim B(n_1,p)$,$Y\sim B(n_2,p)$,且X 与Y 相互独立,则 $X+Y\sim B(n_1+n_2,p)$.

(2)设 $X\sim P(\lambda_1)$,$Y\sim P(\lambda_2)$,且 X 与 Y 相互独立,则 $X+Y\sim P(\lambda_1+\lambda_2)$.

(3)设$X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$,且$X$ 与Y 相互独立,则 $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

(4)设 $X\sim \chi^2(n_1)$,$Y\sim \chi^2(n_2)$,且 X 与 Y 相互独立,则 $X+Y\sim \chi^2(n_1+n_2)$.

【最大最小值的分布】设 X_i 的分布函数为 $F_i(x)$,$i=1,2,\cdots,n$ 且相互独立,则$U=\max\limits_{1\leqslant i\leqslant n} X_i$ 的分布函数为

$$F_{\max}(x)=\prod_{i=1}^{n} F_i(x),$$

$V=\min\limits_{1\leqslant i\leqslant n} X_i$ 的分布函数为

$$F_{\min}(x)=1-\prod_{i=1}^{n}[1-F_i(x)].$$

特别地,若 $X_1,X_2,\cdots,X_n$ 为独立同分布的连续型随机变量,概率密度为 $f(x)$,分布函数为 $F(x)$,则 U 和 V 的概率密度分别为

$$f_{\max}(x)=n[F(x)]^{n-1}f(x),$$

$$f_{\min}(x)=n[1-F(x)]^{n-1}f(x).$$

常见题型·应对

【求边缘分布和条件分布】从本质上讲,求边缘分布就是将联合分布对某个变量所有可能取值的概率进行累加.求条件分布就是在其中一个变量取值一定的情况下,求另一个变量取各个值的条件概率.理解了这个本质就能应对各种复杂的问题.

〖离散型边缘分布的求法〗二维离散型随机变量 (X,Y) 的联合分布通常表现为表格形式:

X \ Y	y_1	$\cdots$	y_n
x_1	p_{11}	$\cdots$	p_{1n}
$\vdots$	$\vdots$		$\vdots$
x_m	p_{m1}	$\cdots$	p_{mn}

这时,表中各行元素的和就是 X 的边缘分布列,表中各列元素的和就是 Y 的边缘分布列.

〖连续型边缘分布的求法〗二维连续型随机变量 (X,Y) 的联合概率密度 $f(x,y)$ 通常为分段函数,求边缘分布时,应画出 $f(x,y)$ 的非零区域 $D=\{(x,y)\mid f(x,y)>0\}$.这样,当 x 在 D 内时,边缘密度

$f_X(x)$为$f(x,y)$从D的下边界到上边界对y积分，当x在D外时，$f_X(x)=0$. 同样，当y在D内时，边缘密度$f_Y(y)$为$f(x,y)$从D的左边界到右边界对x积分，当y在D外时，$f_Y(y)=0$.

例　设二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 2,\max\{0,x-1\}\leqslant y\leqslant\min\{1,x\},\\ 0, & \text{其他},\end{cases}$$

求条件概率密度$f_{Y|Z}(y|x)$.

解　首先画出$f(x,y)$的非零区域，如图17-4-1所示.

$$D=\{(x,y)\mid 0\leqslant x\leqslant 2,$$
$$\max\{0,x-1\}\leqslant y\leqslant\min\{1,x\}\}.$$

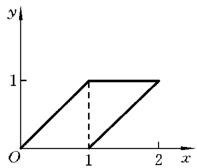


图 17-4-1

其次求X的边缘密度

$$f_X(x)=\begin{cases}\displaystyle\int_0^x 1\mathrm{d}y=x, & 0<x\leqslant 1,\\ \displaystyle\int_{x-1}^1 1\mathrm{d}y=2-x, & 1<x<2,\\ 0, & \text{其他}.\end{cases}$$

最后得到条件密度为,当 $0<x\leqslant 1$ 时,

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{x}, & 0\leqslant y\leqslant x,\\ 0, & \text{其他};\end{cases}$$

当 $1<x<2$ 时,

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{2-x}, & x-1\leqslant y\leqslant 1,\\ 0, & \text{其他};\end{cases}$$

当 $x\leqslant 0$ 或 $x\geqslant 2$ 时,Y 无条件密度.

【独立性的判断】要证明独立性,必须对每一对 (x,y) 验证独立性条件,而要否定独立性,只需对某一对 (x,y) 指出独立性条件不满足即可.

〖可分离变量判别法〗若二维连续型随机变量 (X,Y) 的联合概率密度函数有如下形式:

$$f(x,y)=ch(x)g(y)\quad(c\text{ 为常数}),$$

则 X 与 Y 相互独立.

值得注意的是,当 $f(x,y)$ 为分段函数时,其非零区域 $D=\{(x,y)|a\leqslant x\leqslant b,c\leqslant y\leqslant d$,其中 a,b,c,d 为常数$\}$时方可分离变量. 例如,若 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases}cy(2-x), & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

则 X 与 Y 相互独立. 但若 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases}cy(2-x), & 0\leqslant x\leqslant y,0\leqslant y\leqslant x,\\ 0, & \text{其他},\end{cases}$$

则 X 与 Y 不相互独立.

【独立性的应用】当题设条件给出随机变量相互独立时,可以作出如下推断.

(1)**相互独立的随机变量的边缘分布可惟一地确定它们的联合分布.其实在一般情况下,X和Y的边缘分布不能确定(X,Y)的联合分布.其原因在于边缘分布不能反映X与Y的关系,而X与Y独立正好弥补了这一信息.当然,X与Y的关系也能用其他条件来表现.**(**见下例**)

(2)**若X与Y相互独立,可方便地使用如下公式:**

$$f_{X+Y}(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x,$$

$$f_{\max}(z)=F_X(z)f_Y(z)+f_X(z)F_Y(z),$$

$$f_{\min}(z)=[1-F_X(z)]f_Y(z)+f_X(z)[1-F_Y(y)],$$

$$E(XY)=(EX)(EY),$$

$$D(X\pm Y)=DX+DY.$$

(3)**在独立的条件下可应用正态分布、二项分布和泊松分布的可加性.**

(4)**若$(X_1,X_2,\cdots,X_m)$与$(Y_1,Y_2,\cdots,Y_n)$相互独立,则任一$X_i(i=1,2,\cdots,m)$与任一$Y_j(j=1,2,\cdots,n)$相互独立,这两组随机变量的连续函数$g(X_1,X_2,\cdots,X_m)$与$h(Y_1,Y_2,\cdots,Y_n)$也相互独立.**

例　已知随机变量X_1和X_2的概率分布

$$X_1\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{pmatrix},\quad X_2\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},$$

且$P(X_1X_2=0)=1$.

(1)求 X_1 和 X_2 的联合分布；

(2)问 X_1 和 X_2 是否独立,为什么？

解　设(X_1,X_2)的联合分布为

$X_1 \backslash X_2$	0	1	$p_{i\cdot}$
-1	p_{11}	p_{12}	$\frac{1}{4}$
0	p_{21}	p_{22}	$\frac{1}{2}$
1	p_{31}	p_{32}	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{2}$	

(1)由条件 $P(X_1X_2=0)=1$ 知 $P(X_1X_2\neq 0)=p_{12}+p_{32}=0$,即

$$p_{12}=p_{32}=0,$$

从而

$$p_{11}=\frac{1}{4}-p_{12}=\frac{1}{4},$$

$$p_{31}=\frac{1}{4}-p_{32}=\frac{1}{4},$$

$$p_{21}=\frac{1}{2}-p_{11}-p_{31}=0,$$

$$p_{22}=\frac{1}{2}-p_{21}=\frac{1}{2}.$$

(2)因为 $p_{11}=\frac{1}{4}\neq\frac{1}{4}\cdot\frac{1}{2}=p_{1\cdot}p_{\cdot 1}$,所以 X 与 Y 不独立.

【求两个随机变量函数的分布】解此类题应在理解一般函数的分布(见相应条目)的基础上.合理地运

用公式.

(1)应用随机变量函数的公式时务必注意题设条件.例如,若X与Y不独立就不能用卷积公式的特殊形式

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x$$
$$=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y.$$

(2)连续型随机变量的函数可以是离散型随机变量.这类问题的一般表现形式和解法如下.

已知(X,Y)的联合概率密度$f(x,y)$,定义随机变量Z:

当$(X,Y)\in D_i$时,$Z=z_i,i=1,2,\cdots$,其中D_1,$D_2,\cdots$,是对平面$\mathbf{R}^2$的一个划分,即$\bigcup\limits_i D_i=\mathbf{R}^2$,且$D_i\cap D_j=\varnothing(i\neq j)$.则$Z$的分布律为

$$P(Z=z_i)=\iint\limits_{D_i}f(x,y)\mathrm{d}x\mathrm{d}y,\ i=1,2,\cdots.$$

例　设随机变量(X,Y)服从区域

$$D=\{(x,y)\mid y>0,x^2+y^2\leqslant 1\}$$

上的均匀分布.定义随机变量U,V如下:

$$U=\begin{cases}0, & X<0,\\ 1, & 0\leqslant X<Y,\\ 2, & X\geqslant Y,\end{cases}$$

$$V=\begin{cases}0, & X\geqslant\sqrt{3}Y,\\ 1, & X<\sqrt{3}Y.\end{cases}$$

求(U,V)的联合分布,并计算$P(UV\neq 0)$.

解　由题设知(X,Y)的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{2}{\pi}, & (x,y)\in D,\\ 0, & (x,y)\notin D.\end{cases}$$

(U,V)所有可能的取值有$3\times2=6$个:$(0,0)$,$(0,1)$,$(1,0)$,$(1,1)$,$(2,0)$,$(2,1)$.分别计算概率如下:

$$P(U=0,V=0)=P(\varnothing)=0,$$

$$P(U=1,V=0)=P(\varnothing)=0,$$

$$P(U=1,V=1)=P(0\leqslant X\leqslant Y,X<\sqrt{3}Y)$$

$$=P(0\leqslant X<Y)=\iint\limits_{0\leqslant x<y}f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{S_{扇COD}}{S_{ADE}}=\frac{1}{4}.$$

其中,$S_{扇COD}$和S_{ADE}分别表示图17-4-2中扇形COD和半圆ADE的面积.因为(X,Y)服从二维均匀分布.所以可以由面积比计算概率.

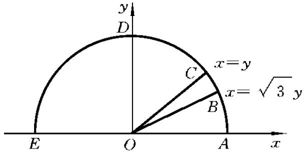


图17-4-2

同理有

$$P(U=0,V=1)=P(X<0,X<\sqrt{3}Y)$$

$$= P(X<0) = \frac{S_{\text{扇}DOE}}{S_{ADE}} = \frac{1}{2},$$

$$P(U=2,V=0) = P(X \geqslant Y, X \geqslant \sqrt{3}Y)$$

$$= P(X \geqslant \sqrt{3}Y) = \frac{S_{\text{扇}AOB}}{S_{ADE}}$$

$$= \frac{1}{6},$$

$$P(U=2,V=0) = P(X \geqslant Y, X < \sqrt{3}Y)$$

$$= P(Y \leqslant X < \sqrt{3}Y) = \frac{S_{\text{扇}BOC}}{S_{ADE}}$$

$$= \frac{1}{12}.$$

所以(U,V)的联合分布列为

U \ V	0	1
0	0	$\frac{1}{2}$
1	0	$\frac{1}{4}$
2	$\frac{1}{6}$	$\frac{1}{12}$

从而

$$P(UV \neq 0) = P(U=1,V=1) + P(U=2,V=1)$$

$$= \frac{1}{4} + \frac{1}{12} = \frac{1}{3}.$$

(3)**求随机变量函数$Z=g(X,Y)$的分布，其关键是求概率$F_Z(z)=P(g(X,Y)\leqslant z)$. 对最值函数，常用方法是将事件$\{\max\limits_{1\leqslant i\leqslant n} X_i \leqslant x\}$理解为**

$$\{X_i \leqslant x\} \quad (i=1,2,\cdots,n)$$

都发生,将事件$\{\min\limits_{1\leqslant i\leqslant n} X_i \leqslant x\}$理解为至少有一个$\{X_i \leqslant x\}(i=1,2,\cdots,n)$发生,利用$X_1,X_2,\cdots,X_n$的独立性,就可计算出有关$\max\limits_{1\leqslant i\leqslant n} X_i$和$\min\limits_{1\leqslant i\leqslant n} X_i$取值的概率.

例　设总体X的分布函数为$F(x)$,$X_1,X_2,\cdots,X_n$是取自X的一个简单随机样本,试求顺序统计量$X_1^* = \min\limits_{1\leqslant i\leqslant n} X_i$和$X_n^* = \max\limits_{1\leqslant i\leqslant n} X_i$的联合分布函数.

解　由于

$$\{X_n^* \leqslant y\} = \{X_1^* \leqslant x, X_n^* \leqslant y\} + \{X_1^* > x, X_n^* \leqslant y\},$$

(X_1^*, X_n^*)的联合分布函数可表示为

$$\begin{aligned} F(x,y) &= P(X_1^* \leqslant x, X_n^* \leqslant y) \\ &= P(X_n^* \leqslant y) - P(X_1^* > x, X_n^* \leqslant y), \end{aligned}$$

而样本$X_1,X_2,\cdots,X_n$是独立同分布的随机变量,因此有

$$P(X_n^* \leqslant y) = \prod_{i=1}^{n} P(X_i \leqslant y) = [F(y)]^n,$$

当$x<y$时,

$$\begin{aligned} P(X_1^* > x, X_n^* \leqslant y) &= P(x < X_i \leqslant y, i=1,2,\cdots,n) \\ &= \prod_{i=1}^{n} P(x < X_i \leqslant y) \\ &= [F(y)-F(x)]^n; \end{aligned}$$

当$x \geqslant y$时,

$$P(X_1^* > x, X_n^* \leqslant y) = 0.$$

综上计算得

$$F(x,y) = \begin{cases} [F(y)]^n - [F(y)-F(x)]^n, & x<y, \\ [F(y)]^n, & x \geqslant y. \end{cases}$$

第18章　随机变量的数字特征

大纲要求

理解随机变量数字特征(数学期望、方差、标准差、矩、协方差、相关系数)的概念,会运用数字特征的基本性质,并掌握常用分布的数字特征.会求随机变量函数的数学期望.

知识点·点拨

18.1　数学期望、方差及其性质

【数学期望】

(1)设离散型随机变量 X 的分布列为$P(X=x_i)=p_i(i=1,2,\cdots)$,若$\sum\limits_i |x_i| p_i$ 收敛,则称

$$EX=\sum_i x_i p_i$$

为 X 的数学期望或均值.

(2)设连续型随机变量 X 的概率密度函数为$f(x)$,若积分$\int_{-\infty}^{+\infty}|x|f(x)$收敛,则称

$$EX=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$$

为 X 的数学期望或均值.

〖位置参数〗数学期望刻画了随机变量取值的平均位置通常称为位置参数,它具有与随机变量相同的量纲.

【方差与标准差】若 $E(X^2)$ 存在，则称

$$DX=E(X-EX)^2$$

为 X 的方差，并称 $\sqrt{DX}$ 为 X 的标准差. 方差常用下式计算：

$$DX=E(X^2)-(EX)^2,$$

若离散型随机变量 X 的分布列为

$$P(X=x_i)=p_i \quad (i=1,2,\cdots),$$

则　$DX=\sum\limits_i (x_i-EX)^2 p_i=\sum\limits_i x_i^2 p_i-(EX)^2.$

若连续型随机变量 X 的概率密度函数为 $f(x)$，则

$$\begin{aligned} DX &= \int_{-\infty}^{+\infty}(x-EX)^2 f(x)\mathrm{d}x \\ &= \int_{-\infty}^{+\infty}x^2 f(x)\mathrm{d}x-(EX)^2. \end{aligned}$$

〖形状参数〗方差和标准差刻画了随机变量的不稳定程度，在分布的图形上，它决定了其形状的“平坦”程度，通常称为形状参数. 方差的量纲是随机变量量纲的平方，标准差具有同随机变量一样的量纲.

【基本性质】

(1)常量性质：对任何常数 a，有

$$E(a)=a,\quad D(a)=0.$$

(2)系数性质：对任何常系数 a，有

$$E(aX)=aEX,\quad D(aX)=a^2DX.$$

(3)和性质：设 X 和 Y 是任意两个随机变量，则

$$E(X\pm Y)=EX\pm EY,$$

$$D(X\pm Y)=DX+DY\pm 2\mathrm{cov}(X,Y).$$

一般地,对 n 个随机变量 $X_1,X_2,\cdots,X_n$,有

$$E\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} EX_i,$$

$$D\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} DX_i+2\sum_{i<j}^{n}\operatorname{cov}(X_i,X_j).$$

〖随机变量的标准化〗若随机变量 X 的期望 EX 和方差 DX 均存在,则称

$$X^*=\frac{X-EX}{\sqrt{DX}}$$

为 X 的标准化变量,它满足 $EX^*=0,DX^*=1$,是解决概率统计问题的常用手段. 如:

若 $$X\sim N(\mu,\sigma^2),$$

则 $$P(X\leqslant a)=\Phi\left(\frac{a-\mu}{\sigma}\right);$$

若 $$X\sim B(n,p),$$

则对充分大的 n 有

$$P(a<X\leqslant b)=\Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right).$$

它们都使用了标准化处理.

【常用分布的数字特征】

名称及记号	数学期望 EX	方差 DX
0-1 分布 $B(1,p)$	p	$p(1-p)$
二项分布 $B(n,p)$	np	$np(1-p)$
泊松分布 $P(\lambda)$	λ	λ
超几何分布 $H(N,M,n)$	$n\dfrac{N}{N+M}$	$n\dfrac{NM}{(N+M)^2}\dfrac{N+M-n}{N+M-1}$

续表

名称及记号	数学期望EX	方差DX
均匀分布$U[a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布$N(\mu,\sigma^2)$	μ	σ^2
指数分布$E(\lambda)$	$1/\lambda$	$1/\lambda^2$

对于常用分布应用结合其分布的背景掌握它们的数学期望和方差(见16.2节和16.3节的【常用分布】条目).

【随机变量函数的数学期望】设随机变量X的分布已知(离散型时为分布列p_i,连续型时为概率密度$f(x)$),若$Y=g(X)$有数学期望存在,则

$$EY=E[g(X)]=\begin{cases}\sum\limits_i g(x_i)p_i, & X\text{ 为离散型},\\ \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x, & X\text{ 为连续型}.\end{cases}$$

设二维随机变量(X,Y)的联合分布已知(离散型时为联合分布列p_{ij},连续型时为联合概率密度$f(x,y)$),若$Z=g(X,Y)$有数学期望存在,则

$$EZ=E[g(X,Y)]=\begin{cases}\sum\limits_i\sum\limits_j g(x_i,y_j)p_{ij}, & (X,Y)\text{为离散型},\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y, & (X,Y)\text{为连续型}.\end{cases}$$

18.2 协方差、相关系数和矩

【协方差】对二维随机变量(X,Y),若

$$E|(X-EX)(Y-EY)|$$

存在，则称

$$\operatorname{cov}(X,Y)=E[(X-EX)(Y-EY)]$$

为 X **与** Y **的相关系数.**

【相关系数】对二维随机变量(X,Y)，**若**$\operatorname{cov}(X,Y)$，DX **和** DY **存在，则称**

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\,\sqrt{DY}}$$

为 X **与** Y **的相关系数.**

〖**无量纲化**〗**协方差**$\operatorname{cov}(X,Y)$**和相关系数**ρ_{XY}**都是描述**X **与** Y **相关程度的数字特征，它们的区别在于**$\operatorname{cov}(X,Y)$**与** X **和** Y **的量纲有关，而** ρ_{XY}**与量纲无关. 当**X^* **和**Y^* **分别为**X **和**Y **的标准化随机变量时有**

$$\operatorname{cov}(X^*,Y^*)=\rho_{X^*Y^*}.$$

【不相关及其等价条件】若 $\rho_{XY}=0$，**则称** X **与** Y **不相关；否则称**X **与**Y **相关.** X **与**Y **不相关有如下充分必要条件：**

(1)$\operatorname{cov}(X,Y)=0$；

(2)$E(XY)=(EX)(EY)$；

(3)$D(X+Y)=DX+DY$.

【基本性质】

(1)$\operatorname{cov}(X,Y)=E(XY)-(EX)(EY)$；

(2)$\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$；

(3)$\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$($a,b$ **为常数**)；

(4)$\operatorname{cov}(X_1+X_2,Y)=\operatorname{cov}(X_1,Y)+\operatorname{cov}(X_2,Y)$；

(5)$-1\leqslant\rho_{XY}\leqslant 1$ **且** $|\rho_{XY}|=1\Leftrightarrow$**存在** a,b，**使**

$$P(Y=aX+b)=1.$$

〖**线性性质的一般形式与特例**〗**利用线性性质求线性函数** $Z=aX+bY$ **的数字特征是计算数字特征的典型问题. 基本性质**(3)、(4)**的一般形式为**

$$\begin{aligned}&\operatorname{cov}(aX+bY,cX+dY)\\&=acDX+(ad+bc)\operatorname{cov}(X,Y)+bdDY.\end{aligned}$$

当 $a=c=1,b=d=\pm1$ **时,有**

$$D(X+Y)=DX\pm2\operatorname{cov}(X,Y)+DY,$$

当 $c=1,d=0$ **时,有**

$$\operatorname{cov}(aX+bY,X)=aDX+b\operatorname{cov}(X,Y).$$

例　**设** $DX=DY=4,\operatorname{cov}(X,Y)=-1,Z=X+cY$,**若** $\operatorname{cov}(X,Z)=\operatorname{cov}(Y,Z)$,**则** $\rho_{XZ}=$________.

解　$\operatorname{cov}(X,Z)=DX+c\operatorname{cov}(X,Y)=4-c$,

$\operatorname{cov}(Y,Z)=\operatorname{cov}(X,Y)+cDY=-1+4c$,

由 $\operatorname{cov}(X,Z)=\operatorname{cov}(Y,Z)$ **得** $c=1$. **且**

$$DZ=D(X+Y)=DX+2\operatorname{cov}(X,Y)+DY=6.$$

故 $\rho_{XZ}=\dfrac{\operatorname{cov}(X,Z)}{\sqrt{DX}\sqrt{DZ}}=\dfrac{3}{\sqrt{4}\ \sqrt{6}}=\dfrac{\sqrt{6}}{4}.$

〖**相关的含义**〗**性质** $|\rho_{XY}|=1\Leftrightarrow P(Y=aX+b)=1$ **表明相关系数** ρ_{XY} **是反映** X **与** Y **存在线性关系的程度的数字特征.** X **与** Y **不相关**($\rho_{XY}=0$)**仅说明** X **与** Y **不存在线性关系而不能说明** X **与** Y **不存在其他关系,因而一般不能导出** X **与** Y **相互独立**(**见** 17.3 **节条目【不相关与独立的关系】**).

【随机变量的原点矩和中心矩】若下面所涉及的数学期望存在,则称

$v_k=E(X^k)$ **为** X **的** k **阶原点矩**　$(k=1,2,\cdots)$;

$\mu_k=E[(E-EX)^k]$为 X 的 k 阶中心矩　$(k=1,2,\cdots)$；

$v_{kl}=E(X^kY^l)$为 X 与 Y 的 $k+l$ 阶混合原点矩 $(k,l=1,2,\cdots)$；

$\mu_{kl}=E[(X-EX)^k(Y-EY)^l]$为 X 与 Y 的 $k+l$ 阶混合中心矩　$(k,l=1,2,\cdots)$.

〖矩的特殊形式〗随机变量的数字特征一般都表现为特殊的矩或矩函数. 如数学期望 EX 是一阶原点矩 v_1，方差 DX 是二阶中心矩阵 μ_2，相关系数 $\rho_{XY}=\dfrac{\mu_{11}}{\sqrt{\mu_{X2}}\sqrt{\mu_{Y2}}}$等等.

常见题型·应对

【警示】随机变量的数字特征是反映随机变量分布特点的参数，它是常数，不是函数，更不是随机变量.

【数字特征的典型求法】求随机变量的各种数字特征，究其本质都是求数学期望，其方法有以下几种.

〖定义法〗先确定随机变量 X 的概率分布，再按定义求 EX.

〖性质法〗已知简单随机变量的数字特征或常用分布的数字特征. 应用基本性质求复杂随机变量的数字特征.

例　设 Y 服从参数为 λ 的指数分布，随机变量

$$X_k=\begin{cases}0, & Y\leqslant k/\lambda,\\ 1, & Y>k/\lambda\end{cases}\quad (k=1,2),$$

求 $E(X_1+X_2)$和 $D(X_1+X_2)$.

解　注意到 X_1 和 X_2 均服从两点分布，且

$$P(X_1=1)=P\left(Y>\frac{1}{\lambda}\right)=\int_{\frac{1}{\lambda}}^{+\infty}\lambda e^{-\lambda x}dx=e^{-1},$$

$$P(X_2=1)=P\left(Y>\frac{2}{\lambda}\right)=\int_{\frac{2}{\lambda}}^{+\infty}\lambda e^{-\lambda x}dx=e^{-2},$$

即　$X_1\sim B(1,e^{-1}),\quad X_2\sim B(1,e^{-2}).$

因此　$E(X_1)=e^{-1},D(X_1)=e^{-1}(1-e^{-1}),$

$E(X_2)=e^{-2},D(X_2)=e^{-2}(1-e^{-2}).$

同样 $Z=X_1X_2$ **也服从两点分布**

$$\begin{aligned}E(X_1X_2)&=P(X_1X_2=1)=P(X_1=1,X_2=1)\\&=P\left(Y<\frac{1}{\lambda},Y>\frac{2}{\lambda}\right)\\&=P\left(Y>\frac{2}{\lambda}\right)=e^{-2}.\end{aligned}$$

于是

$$\begin{aligned}\operatorname{cov}(X_1,X_2)&=E(X_1X_2)-(EX_1)(EX_2)\\&=e^{-2}(1-e^{-1}),\end{aligned}$$

$$E(X_1+X_2)=E(X_1)+E(X_2)=e^{-1}+e^{-2},$$

$$\begin{aligned}D(X_1+X_2)&=D(X_1)+D(X_2)+2\operatorname{cov}(X_1,X_2)\\&=e^{-1}(1-e^{-1})+e^{-2}(1-e^{-2})\\&\quad+2e^{-2}(1-e^{-1})\\&=e^{-1}+2e^{-2}-2e^{-3}-e^{-4}.\end{aligned}$$

〖**函数法**〗**应用定理(见条目【随机变量函数的数学期望】),由** X **的分布求其函数** $Y=g(X)$ **的数学期望,或由** (X,Y) **的联合分布求其函数** $Z=g(X,Y)$ **的数学期望.**

例　**设随机变量** X **与** Y **相互独立,且都服从区间** $[0,1]$ **上的均匀分布,求** $E[\max\{X^2,Y\}]$.

解　由题设知(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$$

故
$$\begin{aligned}E[\max\{X^2,Y\}]&=\int_0^1\int_0^1\max\{x^2,y\}\mathrm{d}x\mathrm{d}y\\&=\int_0^1\left(\int_0^{x^2}x^2\mathrm{d}y+\int_{x^2}^1 y\mathrm{d}y\right)\mathrm{d}x\\&=\int_0^1\left[x^4+\frac{1}{2}(1-x^4)\right]\mathrm{d}x\\&=\frac{3}{5}.\end{aligned}$$

〖分解变量法〗将一个较复杂的随机变量分解成若干个简单随机变量的和，然后利用和性质求数学期望和方差是一个典型的解题技巧.

例　**一台设备有三个独立工作的部件，已知这三个部件需要调整的概率分别为**0.1，0.2，0.3. **试求同时需要调整部件个数的数学期望和方差.**

解1（直接法）　**设**X**为需要调整的部件个数，则**

$$P(X=0)=0.9\times 0.8\times 0.7=0.504,$$
$$\begin{aligned}P(X=1)&=0.1\times 0.8\times 0.7+0.9\times 0.2\times 0.7\\&\quad+0.9\times 0.8\times 0.3=0.398,\end{aligned}$$
$$\begin{aligned}P(X=2)&=0.1\times 0.2\times 0.7+0.1\times 0.8\times 0.3\\&\quad+0.9\times 0.2\times 0.3=0.092,\end{aligned}$$
$$P(X=3)=0.1\times 0.2\times 0.3=0.006.$$

故
$$\begin{aligned}EX&=\sum_{k=0}^{3}kP(X=k)\\&=1\times 0.398+2\times 0.092+3\times 0.006=0.6,\end{aligned}$$

$$\begin{aligned} DX &= E(X^2) - (EX)^2 \\ &= 1 \times 0.398 + 4 \times 0.092 + 9 \times 0.006 - 0.6^2 \\ &= 0.46. \end{aligned}$$

解 2(分解法) 设 X 为需要调整的部件个数,

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 个部分需要调整}, \\ 0, & \text{否则}, \end{cases} \quad i = 1,2,3,$$

由题意知 X_1, X_2, X_3 相互独立,则

$$EX_1 = 0.1, DX_1 = 0.1 \times 0.9 = 0.09,$$

$$EX_2 = 0.2, DX_2 = 0.2 \times 0.8 = 0.16,$$

$$EX_3 = 0.3, DX_3 = 0.3 \times 0.7 = 0.21.$$

故
$$EX = \sum_{i=1}^{3} EX_i = 0.6,$$

$$DX = \sum_{i=1}^{3} DX_i = 0.46.$$

【不相关与独立的判断】 随机变量 X 与 Y 是否不相关由协方差 $\operatorname{cov}(X,Y)$ 或相关系数 ρ_{XY} 是否为零来判断;X 与 Y 是否相互独立用联合概率密度函数 $f(x,y)$ 是否等于边缘概率密度函数的乘积 $f_X(x)f_Y(y)$ 来判断. 这类题的判断结果往往是,X 与 Y 不相关,但 X 与 Y 不独立.

例 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y) = \frac{1}{2}[\varphi_1(x,y) + \varphi_2(x,y)],$$

其中 $\varphi_1(x,y)$ 和 $\varphi_2(x,y)$ 都是二维正态密度函数,且它们对应的二维随机变量的相关系数分别为 $\frac{1}{3}$ 和 $-\frac{1}{3}$,它们的边缘密度函数对应的随机变量的数学

期望均为 0,方差均为 1.

(1)求随机变量 X 和 Y 的密度函数 $f_1(x)$ 和 $f_2(y)$,及 X 和 Y 的相关系数 ρ(可以直接利用二维正态密度的性质);

(2)X 和 Y 是否独立?为什么?

解　(1)由于二维正态密度函数的两个边缘密度都是正态密度函数,因此,$\varphi_1(x,y)$ 和 $\varphi_2(x,y)$ 的两个边缘密度为标准正态密度函数. 故

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y \\ &= \frac{1}{2}\left[\int_{-\infty}^{+\infty}\varphi_1(x,y)\mathrm{d}y + \int_{-\infty}^{+\infty}\varphi_2(x,y)\mathrm{d}y\right] \\ &= \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} + \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\right) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}, \end{aligned}$$

同理
$$f_2(y) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}.$$

由于 $X \sim N(0,1)$,$Y \sim N(0,1)$,可见 $E(X) = E(Y) = 0$,$D(X) = D(Y) = 1$,故 X 与 Y 的相关系数为

$$\begin{aligned} \rho &= \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = E(XY) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\mathrm{d}x\mathrm{d}y \\ &= \frac{1}{2}\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\varphi_1(x,y)\mathrm{d}x\mathrm{d}y \right. \\ &\qquad \left. + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\varphi_2(x,y)\mathrm{d}x\mathrm{d}y\right] \\ &= \frac{1}{2}\left(\frac{1}{3} - \frac{1}{3}\right) = 0. \end{aligned}$$

(2)由题设

$$f(x,y)=\frac{1}{2}[\varphi_1(x,y)+\varphi_2(x,y)]$$
$$=\frac{3}{8\pi\sqrt{2}}\left[e^{-\frac{9}{16}(x^2-\frac{2}{3}xy+y^2)}+e^{-\frac{9}{16}(x^2+\frac{2}{3}xy+y^2)}\right],$$

而 $f_1(x)f_2(y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$,故

$$f(x,y)\neq f_1(x)f_2(y),$$

所以,X 与 Y 不独立.

第19章　大数定律和中心极限定理

大纲要求

了解切比雪夫不等式. 了解切比雪夫大数定律、贝努利大数定律和辛钦大数定律(独立同分布随机变量的大数定律). 了解棣莫佛-拉普拉斯定理(二项分布以正态分布为极限分布)和列维-林德伯格定理(独立分布的中心极限定理).

知识点·点拨

19.1　大数定律

【切比雪夫不等式】若随机变量 X 的数学期望 $\mu=EX$ 和方差 $\sigma^2=DX$ 存在,则对任何 $\varepsilon>0$,有

$$P(|X-\mu|\geqslant\varepsilon)\leqslant\frac{\sigma^2}{\varepsilon^2}$$

或

$$P(|X-\mu|<\varepsilon)\geqslant 1-\frac{\sigma^2}{\varepsilon^2}.$$

〖利用期望和方差估算概率〗切比雪夫不等式说明无需已知 X 的分布,仅由其数字特征 EX 和 DX 就可估算 X 在 EX 附近取值的概率,通常是 X 在 EX 附近区域取值的概率较大. 而这种"较大"的程度取决于 DX.

例　设随机变量 X 和 Y 的数学期望都是 2,方差分别为 1 和 4,而相关系数为 0.5,则根据切比雪夫不等式 $P(|X-Y|\geqslant 6)\leqslant$________.

解　设 $Z=X-Y$,由期望和方差的和性质有

$$EZ=EX-EY=0,$$

$$DZ = DX + DY - 2\mathrm{cov}(X,Y)$$
$$= 1 + 4 - 2 \times 0.5 \times \sqrt{1} \times \sqrt{4} = 3.$$

根据切比雪夫不等式

$$P(|X-Y| \geqslant 6) = P(|Z-EZ| \geqslant 6)$$
$$\leqslant \frac{3}{6^2} = \frac{1}{12}.$$

【大数定律】若随机变量序列 $X_1, X_2, \cdots$ **满足:对任何** $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(\left| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i \right| < \varepsilon \right) = 1,$$

则称随机变量序列 $\{X_n\}$ **服从大数定律.它揭示了随机现象由不确定性向确定性演化的规律.**

【三个重要的大数定律】

名　称	条　件	结　论
切比雪夫大数定律	$X_1, X_2, \cdots$ **相互独立,方差存在且有界**:$DX_i < C$ $(i=1,2,\cdots)$	**对任何** $\varepsilon > 0$,**有** $\lim\limits_{n\to\infty} P(\lvert\overline{X} - \mu_n\rvert < \varepsilon) = 1$, **其中** $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, $\mu_n = E(\overline{X})$
贝努利大数定律	n_A **为** n **重贝努利试验中** A **发生的次数**,p **为每次试验中** A **发生的概率**	**对任何** $\varepsilon > 0$,**有** $\lim\limits_{n\to\infty} P\left(\left\lvert\frac{n_A}{n} - p\right\rvert < \varepsilon\right) = 1$
辛钦大数定律	$X_1, X_2, \cdots$ **独立同分布,且** $EX_i = \mu (i=1,2,\cdots)$	**对任何** $\varepsilon > 0$,**有** $\lim\limits_{n\to\infty} P(\lvert\overline{X} - \mu\rvert < \varepsilon) = 1$, **其中** $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$

〖大数定律的应用〗大数定律在本课程中有如下应用：

(1)大数定律是参数估计中矩估计方法的理论依据.

(2)大数定律是验证估计量一致性(相容性)的主要工具.

(3)贝努利大数定律是切比雪夫大数定律和辛钦大数定律的特殊情形，它揭示了事件发生的频率 n_A/n 趋于事件发生的概率 p 的统计规律.

19.2　中心极限定理

【列维-林德伯格定理(独立同分布的中心极限定理)】设 $X_1, X_2, \cdots$，为独立同分布的随机变量序列，且 $EX_i=\mu, DX_i=\sigma^2(i=1,2,\cdots)$，则随机变量

$$Y_n=\frac{\sum_{i=1}^{n} X_i-n\mu}{\sqrt{n\sigma^2}}$$

的分布函数 $F_n(x)$ 满足

$$\lim_{n\to\infty} F_n(x)=\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t=\Phi(x).$$

〖正态分布适用的依据〗中心极限定理揭示了大量随机变量的和在一定条件下趋于正态分布的统计规律，它是正态分布得以广泛应用的重要原因.

【棣莫佛-拉普拉斯定理】设 X_n 为 n 重贝努利试验中事件 A 发生的次数. $p=P(A)$，即 $X_n \sim B(n,p)$，则对任何实数 x 都有

$$\lim_{n\to\infty}P\left(\frac{X_n-np}{\sqrt{np(1-p)}}\leqslant x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt.$$

〖定理的应用〗棣莫佛-拉普拉斯中心极限定理是列维-林德伯格中心极限定理的特殊情况.它可直接用来近似计算二项分布的概率:设$X\sim B(n,p)$,若n充分大,则可认为$\frac{X-np}{\sqrt{np(1-p)}}\sim N(0,1)$近似成立.

常见题型·应对

【解题要点】应用列维-林德伯格中心极限定理的要点如下:

(1)将问题中的随机变量表示为n个独立同分布随机变量的和:$X=\sum_{i=1}^{n}X_i$,通常$n\geqslant 30$.

(2)将和进行标准化:$Y=\frac{X-EX}{\sqrt{DX}}$.

(3)视Y为标准正态分布$N(0,1)$计算概率.

(4)如果题设随机变量$X\sim B(n,p)$且n充分大,则可近似计算概率,即

$$P(a<X\leqslant b)\approx\Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right).$$

例　计算机做加法运算时,要对每个加数取整(即取最接近它的整数),设所有的取整误差是相互独立的,且它们都服从均匀分布$U[-0.5,0.5]$,如果将1 500个数相加,求误差总和的绝对值超过15

的概率.

解　设 X_i 为第 i 个加数的取整误差,由题设条件知

$$X_i \sim U[-0.5, 0.5], \quad i = 1, 2, \cdots, 1\,500,$$

相互独立,则误差总和 $X = \sum_{i=1}^{1\,500} X_i$,且

$$EX = 1\,500 \times EX_i = 0,$$

$$DX = 1\,500 DX_i = 1\,500 \times \frac{1}{12} = 125.$$

由列维-林德伯格中心极限定理 $Y = \frac{X-0}{\sqrt{125}}$ 近似服从 $N(0,1)$,故所求概率

$$\begin{aligned} P(|X| > 15) &= P\left(\left| \frac{X}{\sqrt{125}} \right| > \frac{15}{\sqrt{125}} \right) \\ &\approx 2\left[1 - \Phi\left(\frac{15}{\sqrt{125}} \right) \right] = 0.179\,9. \end{aligned}$$

第 20 章　数理统计的基本概念

大纲要求

理解总体、简单随机样本、统计量、样本均值、样本方差及样本矩的概念，其中样本方差定义为

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

了解 χ^2 分布、t 分布和 F 分布的概念及性质，了解分位数的概念并会查表计算. 了解正态总体的常用抽样分布.

知识点・点拨

20.1　总体、样本与统计量

【总体】在数理统计中，将研究对象的数量指标的集合称为总体或母体. 它表现为一个随机变量 X 或 X 的分布函数 $F(x)$.

【简单随机样本】从总体中抽取的部分个体称为样本. 它们表现为 n 个相互独立且与总体同分布的随机变量 $X_1,X_2,\cdots,X_n$，也称为简单随机样本. 抽取完成后得到的数值 $x_1,x_2,\cdots,x_n$ 称为样本观察值或样本值. n 称为样本容量.

〖样本与其观察值的区别〗样本 $X_1,X_2,\cdots,X_n$ 在试验前是 n 个独立同分布的随机变量，在试验后则是 n 个实数，即样本观察值. 同样统计量 T 作为样本的函数在试验前也是随机变量，其分布可用来分析它的

统计作用和性质，试验后作为样本观察值的函数是一个实数，它是一次参数估计或假设检验的结果.

【统计量】 设 $T(X_1, X_2, \cdots, X_n)$ 为样本的连续函数，若 $T(X_1, X_2, \cdots, X_n)$ 不含有总体的任何未知参数，则称之为一个统计量.

〘统计量的判断〙 统计量是对样本的一次“加工”，“加工”的目的在于提取样本中有关总体分布的重要信息. 这种“加工”在数学上就表现为函数. 要求此函数连续是为了保证它仍然是一个随机变量，这对于常用的函数都满足要求. 因此判断样本的函数是否为统计量可以仅看其中是否含有未知参数.

【常用统计量】

样本均值：$\overline{X} = \dfrac{1}{n} \sum\limits_{i=1}^{n} X_i$；

样本方差：$S^2 = \dfrac{1}{n-1} \sum\limits_{i=1}^{n} (X_i - \overline{X})^2$；

样本标准差：$S = \sqrt{\dfrac{1}{n-1} \sum\limits_{i=1}^{n} (X_i - X)^2}$；

样本 k 阶原点矩：$A_k = \dfrac{1}{n} \sum\limits_{i=1}^{n} X_i^k$；

样本 k 阶中心矩：$B_k = \dfrac{1}{n} \sum\limits_{i=1}^{n} (X_i - \overline{X})^k$；

顺序统计量：$X_{(1)} = \min\limits_{1 \leqslant i \leqslant n} \{X_i\}$，$X_{(n)} = \max\limits_{1 \leqslant i \leqslant n} \{X_i\}$.

〘总体均值、总体方差与样本均值、样本方差〙 总体 X 的均值（期望）EX 和方差 DX 是 X 分布的数字特征，是客观存在的实数，在数理统计中常常是总体分布的未知参数. 而样本均值 $\overline{X}$ 和样本方差 S^2 则

是统计量,也是随机变量,在不同的试验中会有不同的取值,$\overline{X}$ 和 S^2 常常用来估计 EX 和 DX.

20.2 抽样分布

【χ^2 分布】设 $X_1, X_2, \cdots, X_n$ 相互独立,均服从 $N(0,1)$,则称 $Y=\sum\limits_{i=1}^{n} X_i^2$ 服从自由度为 n 的 χ^2 分布,记作 $Y\sim\chi^2(n)$.

【t 分布】设 $X\sim N(0,1)$,$Y\sim\chi^2(n)$,且 X 与 Y 相互独立,则称 $T=\dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布,记作 $T\sim t(n)$.

【F 分布】设 $X\sim\chi^2(m)$,$Y\sim\chi^2(n)$,且 X 与 Y 相互独立,则称 $F=\dfrac{X/m}{Y/n}$ 服从自由度为 m,n 的 F 分布,记作 $F\sim F(m,n)$.

〖三个统计分布的关系及性质〗对 χ^2 分布、t 分布、F 分布主要应掌握它的构造性定义,为此可如下记忆:

$$
\begin{array}{ccc}
 & N(0,1) & \\
\swarrow & & \searrow \\
\sum\limits_{i=1}^{n}[N(0,1)]_i^2\sim\chi^2(n) & \rightarrow & \dfrac{N(0,1)}{\sqrt{\chi^2(n)/n}}\sim t(n) \\
\downarrow & & \\
\dfrac{\chi^2(m)/m}{\chi^2(n)/n}\sim F(m,n) & &
\end{array}
$$

(表达式中各量相互独立.)

在此基础上就不难理解这三个分布的如下性

质：

(1)χ^2 分布的可加性：设 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，且 X 与 Y 相互独立，且 $X+Y\sim\chi^2(n_1+n_2)$.

(2)设 $X\sim\chi^2(n)$，则 $EX=n$，$DX=2n$.

(3)设 $T\sim t(n)$，则 $ET=0$，且 n 充分大时，T 近似服从 $N(0,1)$.

(4)设 $X\sim F(m,n)$，则 $\frac{1}{X}\sim F(n,m)$.

【(上侧)分位数】若 $P(X>\lambda_\alpha)=\alpha$，则称 λ_α 为 X 分布的 α(上侧)分位数或(上侧)分位点. 特别记

μ_α 为正态分布 $N(0,1)$ 的 α 分位数；

$\chi^2_\alpha(n)$ 为 χ^2 分布 $\chi^2(n)$ 的 α 分位数；

$t_\alpha(n)$ 为 t 分布 $t(n)$ 的 α 分位数；

$F_\alpha(m,n)$ 为 F 分布 $F(m,n)$ 的 α 分位数.

〖分位点的转换〗分位点是计算机求得的统计分布边界点与概率的关系，由于书籍篇幅所限，所列出的值不能完全满足统计分析的要求，可用以下方法求出所需分位点.

(1)线性插值. 将 λ_α 近似为 α 的线性函数，若已知 $\lambda_1=\lambda_{\alpha_1}$，$\lambda_2=\lambda_{\alpha_2}$，$\alpha_1<\alpha_0<\alpha_2$ 则

$$\lambda_0=\lambda_{\alpha_0}\approx\lambda_1+\frac{\lambda_2-\lambda_1}{\alpha_2-\alpha_1}(\alpha_0-\alpha_1).$$

(2)当 $n>45$ 时，

$$t_\alpha(n)\approx u_\alpha,\quad \chi^2_\alpha(n)\approx\frac{1}{2}(\mu_\alpha+\sqrt{2n-1})^2.$$

(3)$F_{1-\alpha}(m,n)=\dfrac{1}{F_\alpha(n,m)}$.

【抽样定理】设 $X_1,X_2,\cdots,X_n$ 为取自正态总体

$N(\mu_1,\sigma^2)$的样本,则

(1)$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$;

(2)$\dfrac{n-1}{\sigma^2}S^2=\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)$;

(3)$\overline{X}$ 与 S^2 相互独立.

〖正态总体的 $\overline{X}$ 和 S^2〗样本均值 $\overline{X}$ 和样本方差 S^2 是构造统计量的基本要素,而抽样定理的三个结论描述了正态总体的 $\overline{X}$ 和 S^2 的基本性质,为导出统计量的分布提供理论依据.

〖推论 1〗设 $X_1,X_2,\cdots,X_n$ 为取自正态总体 $N(\mu,\sigma^2)$的样本,样本均值和样本方差为 $\overline{X}$ 和 S^2,则

$$\frac{\overline{X}-\mu}{S}\sqrt{n}\sim t(n-1).$$

〖推论 2〗设 $X_1,X_2,\cdots,X_{n1}$ 为取自正态总体 $N(\mu_1,\sigma^2)$的样本,样本均值和样本方差为 $\overline{X}$ 和 S_1^2, $Y_1,Y_2,\cdots,Y_{n2}$ 为取自正态总体 $N(\mu_2,\sigma^2)$的样本,样本均值和样本方差为 $\overline{Y}$ 和 S_2^2,则

$$T=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2),$$

其中　$$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

〖推论 3〗设 S_i^2 为取自总体 $N(\mu_i,\sigma_i^2)$的样本方差,样本容量为 $n_i(i=1,2)$,则

$$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1).$$

常见题型 · 应对

【解题模式】本章考题的典型模式如下：

题设条件：$X_1,X_2,\cdots,X_n$ 为取自正态总体的样本，或 $X_1,X_2,\cdots,X_n$ 独立同分布于 $N(\mu,\sigma^2)$.

推导依据：抽样定理和 χ^2 分布、t 分布、F 分布的构造性定义.

问题结论：所涉统计量（通常为 $\overline{X}$ 和 S^2 的函数）服从 χ^2 分布或 t 分布或 F 分布.

抽样定理的三个推论就是上述模式下的结果.

例　设 $X_1,X_2,\cdots,X_9$ 是来自正态总体 X 的简单随机样本.

$$Y_1=\frac{1}{6}(X_1+X_2+\cdots+X_6),$$

$$Y_2=\frac{1}{3}(X_7+X_8+X_9),$$

$$S^2=\frac{1}{2}\sum_{i=7}^{9}(X_i-Y_2)^2,$$

$$Z=\frac{\sqrt{2}(Y_1-Y_2)}{S},$$

求统计量 Z 的分布.

解　视 $X_1,\cdots,X_6$ 和 $X_7,\cdots,X_9$ 为来自正态总体 $N(\mu,\sigma^2)$ 的两个样本，则 Y_1 和 Y_2 为两者的样本均值，S^2 为后者的样本方差，由抽样定理知

$$Y_1\sim N\left(\mu,\frac{\sigma^2}{6}\right),\quad Y_2\sim N\left(\mu,\frac{\sigma^2}{3}\right),$$

$$\frac{2}{\sigma^2}S^2\sim\chi^2(2).$$

且三者相互独立,从而

$$U=\frac{Y_1-Y_2}{\sqrt{\frac{\sigma^2}{6}+\frac{\sigma^2}{3}}}=\sqrt{\frac{2}{\sigma^2}}(Y_1-Y_2)\sim N(0,1),$$

$$V=\frac{2}{\sigma^2}S^2\sim\chi^2(2),$$

且U **与** V **相互独立,故**

$$Z=\frac{U}{\sqrt{V/2}}\sim t(2).$$

第21章 参 数 估 计

大纲要求

理解参数的点估计、估计量与估计值的概念. 掌握矩估计法(一阶、二阶矩)和最大似然估计法. 了解估计量的无偏性、有效性(最小方差性)和一致性(相合性)的概念,并会验证估计量的无偏性. 理解区间估计的概念,会求单个正态总体的均值和方差的置信区间,会求两个正态总体的均值差和方差比的置信区间.

知识点 · 点拨

21.1 点估计方法

【点估计】用一个量来估计一个未知参数的方法称为点估计.

【估计量与估计值】设总体分布为$F(x,\theta)$,θ为未知参数. 由样本$X_1,X_2,\cdots,X_n$构造一个统计量$\hat{\theta}(X_1,X_2,\cdots,X_n)$用来估计$\theta$,则称$\hat{\theta}(X_1,X_2,\cdots,X_n)$为$\theta$的估计量,其观察值$\hat{\theta}(x_1,x_2,\cdots,x_n)$称为$\theta$的估计值.

【矩估计法】它是依据大数定律建立的一种估计方法,其步骤如下(限于两个未知参数):

(1)通过求总体的一、二阶原点矩v_1,v_2,将未知参数θ_1,θ_2表示成v_1,v_2的函数,即

$$\text{求}\begin{cases}EX=v_1(\theta_1,\theta_2),\\E(X^2)=v_2(\theta_1,\theta_2),\end{cases}\text{解出}\begin{cases}\theta_1=h_1(v_1,v_2),\\\theta_2=h_2(v_1,v_2).\end{cases}$$

(2)**用样本原点矩** $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k(k=1,2)$ **代替总体原点矩得到参数的矩估计**:

$$\begin{cases}\hat{\theta}_1=h_1(A_1,A_2),\\\hat{\theta}_2=h_2(A_1,A_2).\end{cases}$$

〖**常用估计量**〗**设总体的数学期望** $\mu=E(X)$ **和方差** $\sigma^2=D(X)$ **均存在,则其矩估计量分别为**

$$\hat{\mu}=\overline{X},\quad \hat{\sigma}^2=B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

〖**矩估计的函数**〗(**见** 21.2 **节相关条目**)

【**最大似然估计(极大似然估计)**】**它是依据"已发生事件的概率应该最大"的原则建立的一种估计方法,其步骤为**:

(1)**对样本值** $x_1,x_2,\cdots,x_n$ **写出似然函数**

$$L(\theta;x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}p_\theta(x_i),$$

其中,对离散型总体,$p_\theta(x_i)$ **为概率** $P_\theta(X=x_i)$;**对连续型总体**,$p_\theta(x_i)$ **为概率密度值** $f(x_i;\theta)$.

(2)**视似然函数为未知参数** θ **的函数** $L(\theta)$,**求其最大值点** $\hat{\theta}$,**即为** θ **的最大似然估计,即**

$$L(\hat{\theta})=\max_\theta L(\theta).$$

〖**常用估计量**〗**对于常用分布的总体,其参数的最大似然估计如下**:

(1)$X\sim N(\mu,\sigma^2)$,$\hat{\mu}=\overline{X}$.

当 μ **未知时**,$\hat{\sigma}^2=B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$;

当 μ 已知时，$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$.

(2) $X\sim B(1,p)$，$\hat{p}=\overline{X}$.

(3) $X\sim P(\lambda)$，$\hat{\lambda}=\overline{X}$.

(4) $X\sim U[a,b]$，

$$\hat{a}=X_{(1)}=\min_{1\leqslant i\leqslant n}\{X_i\},\quad \hat{b}=X_{(n)}=\max_{1\leqslant i\leqslant n}\{X_i\}.$$

〖最大似然估计的函数〗(见 21.2 节相关条目)

例 设总体 $X\sim N(\mu,1)$，由样本 $(X_1,X_2,\cdots,X_n)$ 求参数 $\theta=P(X>1)$ 的最大似然估计.

解 因为

$$\theta=P(X>1)=1-\Phi\left(\frac{1-\mu}{1}\right)=\Phi(\mu-1)$$

是 μ 的单调增函数，而 μ 的最大似然估计为 $\hat{u}=\overline{X}$，所以 θ 的最大似然估计为

$$\hat{\theta}=\Phi(\hat{\mu}-1)=\Phi(\overline{X}-1).$$

21.2 估计量的评选标准

【无偏性】设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为未知参数 θ 的估计量. 若 $E(\hat{\theta})=\theta$，则称 $\hat{\theta}$ 为 θ 的无偏估计量.

〖不惟一性〗对一个未知参数 θ 可以有多个(甚至无穷多个)无偏估计量，比如 $\frac{1}{2}X_1+\frac{1}{2}X_2$ 和 $\frac{1}{3}X_1+\frac{2}{3}X_2$ 都是总体期望的无偏估计. 因此有必要比较它们的有效性.

【有效性】设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(X_1,

$X_2,\cdots,X_n)$为未知参数θ的两个无偏估计量,若$D(\hat{\theta}_1)<D(\hat{\theta}_2)$,则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

〖讨论的前提〗有效性是在估计量已具有无偏性的前提下的优良性质,失去了无偏性讨论有效性是无意义的.

【相合性(一致性)】设$\hat{\theta}_n=\hat{\theta}(X_1,X_2,\cdots,X_n)$是$\theta$的一个估计量序列,如果对任何正数$\varepsilon$,都有

$$\lim_{n\to\infty}P(|\hat{\theta}_n-\theta|<\varepsilon)=1,$$

则称$\hat{\theta}_n$为θ的相合估计(一致估计).

〖常用估计量的优良性〗关于估计量的优良性有如下结论:

(1)若总体的期望(均值)EX和方差DX均存在,则

$$E(\overline{X})=EX,E(S^2)=DX,$$

即样本均值和样本方差分别是总体均值和总体方差的无偏估计.

(2)若总体均值和方差均存在,则

$$D(\overline{X})\leqslant D\left(\sum_{i=1}^{n}a_iX_i\right),\quad \sum_{i<1}^{n}a_i=1,$$

即样本均值$\overline{X}$在所有总体均值μ的线性无偏估计$\hat{\mu}=\sum\limits_{i=1}^{n}a_iX_i(\sum\limits_{i=1}^{n}a_i=1)$中方差最小.

(3)对于正态总体$N(\mu,\sigma^2)$,$\overline{X}$是μ的有效估计.当μ已知时,$S_0^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$是$\sigma^2$的有效估计(也是极大似然估计).这里有效估计是指方差达到

最小的无偏估计.

(4)由大数定律知,对任何正数ε,有

$$\lim_{n\to\infty}P(|\overline{X}-EX|<\varepsilon)=1,$$

即样本均值$\overline{X}$是总体均值EX的相合估计. 一般样本k阶原点矩A_k是总体k阶原点矩v_k的相合估计.

〖估计量的函数〗估计量取函数有如下结论:

(1)设$\hat{\theta}$是θ的矩估计,$g(\theta)$是θ的连续函数,则$g(\hat{\theta})$是$g(\theta)$的矩估计.

(2)设$\hat{\theta}$是θ的最大似然估计,$g(\theta)$是θ的单调函数,则$g(\hat{\theta})$是$g(\theta)$的最大似然估计.

(3)若$\hat{\theta}$是θ的无偏估计,则$a\hat{\theta}+b$是$a\theta+b$的无偏估计;若$\hat{\theta}_i$是θ_i的无偏估计$(i=1,2,\cdots,m)$,则$\sum\limits_{i=1}^{m}a_i\hat{\theta}_i$是$\sum\limits_{i=1}^{m}a_i\theta_i$的无偏估计,其中$a,b,a_i$均为常数.

(4)设$\hat{\theta}$是θ的相合估计,$g(\theta)$是θ的连续函数,则$g(\hat{\theta})$是$g(\theta)$的相合估计.

21.3 区间估计

【置信区间与置信度(水平)】设$X_1,X_2,\cdots,X_n$为取自总体$F(x;\theta)$的样本,θ为未知参数. 构造两个统计量$\hat{\theta}_1=\theta_1(X_1,X_2,\cdots,X_n)$和$\hat{\theta}_2=\theta_2(X_1,X_2,\cdots,X_n)$,使

$$P(\hat{\theta}_1\leqslant\theta\leqslant\hat{\theta}_2)=1-\alpha,$$

则称区间$[\hat{\theta}_1,\hat{\theta}_2]$为$\theta$的置信区间,$1-\alpha$为该区间的置信度(或置信水平),也称$\hat{\theta}_1$为$\theta$的置信下限,$\hat{\theta}_2$为$\theta$的置信上限. 当$\hat{\theta}_1\equiv-\infty$或$\hat{\theta}_2\equiv+\infty$时,称$(-\infty,$

$\hat{\theta}_2]$**和**$[\hat{\theta}_1,+\infty)$**为单侧置信区间,否则称**$[\hat{\theta}_1,\hat{\theta}_2]$**为双侧置信区间.**

【正态总体均值和方差的置信区间】

单正态总体 $N(\mu\sigma^2)$:

待估参数	条件	核心统计量	分布	置信区间
μ	σ^2 已知	$\overline{X}$	$N(0,1)$	$\left[\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right]$
	σ^2 未知		$t(n-1)$	$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right]$
σ^2	μ 已知	$\sum\limits_{i=1}^{n}(X_i-\mu)^2$	$\chi^2(n)$	$\left[\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right]$
	μ 未知	S^2	$\chi^2(n-1)$	$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1\ \alpha/2}(n-1)}\right]$

双正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$：

待估参数	条件	核心统计量	分布	置信区间
$u_1-\mu_2$	σ_1^2,σ_2^2 已知	$\overline{X}-\overline{Y}$	$N(0,1)$	$\left[(\overline{X}-\overline{Y})\pm u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$
	$\sigma_1^2=\sigma_2^2=\sigma^2$ 未知		$t(n_1+n_2-2)$	$\left[(\overline{X}-\overline{Y})\pm t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]$ $S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$
$\frac{\sigma_1^2}{\sigma_2^2}$	μ_1,μ_2 已知	$F_0=\frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}$	$F(n_1,n_2)$	$\left[\frac{n_2}{n_1}\frac{F_0}{F_{\alpha/2}(n_1,n_2)},\frac{n_2}{n_1}\frac{F_0}{F_{1-\alpha/2}(n_1,n_2)}\right]$
	μ_1,μ_2 未知	$\frac{S_1^2}{S_2^2}$	$F(n_1-1,n_2-1)$	$\left[\frac{S_1^2/S_2^2}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2/S_2^2}{F_{1\ \alpha/2}(n_1-1,n_2-1)}\right]$

〖双、单侧置信区间的转换〗正态总体参数的置信区间表中仅列出了双侧置信区间.将其中的一个置信限换为$+\infty$(或$-\infty$),另一个置信限中分位点的$\alpha/2$换为α,就可得到该参数的单侧置信区间.

常见题型·应对

【提示】矩估计和最大似然估计是求估计量的方法,无偏估计、一致估计是已有估计量的性质,这是两组不同性质的概念.

【求矩估计的要点】求矩估计时要根据实际情况进行灵活处理.

(1)在求总体原点矩时,由低阶到高阶,以够用为原则.一般情况下,仅有一个未知参数时,求一阶原点矩即可;若待估参数有k个,则应求出k个($1\sim k$阶)原点矩.以保证由k个表达式解出k个未知参数.

(2)当未知参数更容易表示为总体中心矩(比如方差为总体的二阶中心矩)的函数$\theta=h(\mu_2,\cdots,\mu_r)$时,可由样本中心矩$B_k$代替总体中心矩$\mu_k$得到矩估计$\hat{\theta}=h(B_2,\cdots,B_r)$.当所得结果与仅用原点矩的结果不一致时,以取简舍繁为原则.

例 设总体概率密度函数为$f(x)=\dfrac{\theta}{2}e^{-\theta|x|}$($\theta>0$),试求未知参数$\theta$的矩估计量.

解 注意到$f(x)$为偶函数,显然有$EX=0$,故需由二阶矩来表示θ.由

$$\nu_2=EX^2=\int_{-\infty}^{+\infty}x^2\frac{\theta}{2}e^{-\theta|x|}\mathrm{d}x$$

$$=\int_0^{+\infty}x^2\theta\mathrm{e}^{-\theta x}\mathrm{d}x=\frac{2}{\theta^2}$$

(其中后一等式由指数分布的二阶原点矩得到)解得

$$\theta=\sqrt{2/\nu_2}=(\nu_2/2)^{-\frac{1}{2}}.$$

于是θ的矩估计为

$$\hat{\theta}=\left(\frac{A_2}{2}\right)^{-\frac{1}{2}}=\left(\frac{1}{2n}\sum_{i=1}^{n}X_i^2\right)^{-\frac{1}{2}}.$$

注意到对本例总体而言,有

$$\mu_2=DX=E(X^2)=\nu_2,$$

$$\theta=(\mu_2/2)^{-\frac{1}{2}},$$

故　$$\hat{\theta}=(B_2/2)^{-\frac{1}{2}}=\left[\frac{1}{2n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]^{-\frac{1}{2}}$$

也是θ的矩估计.不过它不如前者简洁.

【求最大似然估计的要点】求最大似然估计的关键在于求似然函数的最大值点.一般采用如下方法:

(1)对于最大似然估计,由于目标函数$L(\theta)$为分布密度的连乘积,故通常由解如下对数似然方程(组)

$$\frac{\partial[\ln L(\theta)]}{\partial\theta}=0$$

得到$L(\theta)$的最大值点,即最大似然估计$\hat{\theta}$.

(2)当总体的分布密度为分段函数,且分段点与未知参数θ有关时,会造成对数似然方程无解.这时可依据$L(\theta)$对θ的单调性以及θ的取值范围来确定$L(\theta)$的最大值点,即最大似然估计$\hat{\theta}$.

例　设总体X的概率密度为

$$f(x;\alpha,\beta)=\begin{cases}\beta\alpha^{\beta}x^{-(\beta+1)}, & x\geqslant\alpha,\\ 0, & x<\alpha,\end{cases}\quad(\alpha>0,\beta\geqslant1).$$

(1)当$\alpha=1$时,求未知参数β的最大似然估计量.

(2)当$\beta=1$时,求未知参数α的最大似然估计量.

解　总体X关于样本$x_1,x_2,\cdots,x_n$的似然函数为

$$L(\alpha,\beta)=\prod_{i=1}^{n}f(x;\alpha,\beta)$$

$$=\begin{cases}\beta^{n}\alpha^{n\beta}\left(\prod\limits_{i=1}^{n}x_i\right)^{-\beta+1}, & x_1,x_2,\cdots,x_n\geqslant\alpha,\\ 0, & \text{其他}.\end{cases}$$

(1)当$\alpha=1$时,β的似然函数为

$$L(\beta)=\begin{cases}\beta^{n}(\prod\limits_{i=1}^{n}x_i)^{-(\beta+1)}, & x_1,x_2,\cdots,x_n\geqslant1,\\ 0, & \text{其他}.\end{cases}$$

解对数似然方程

$$\frac{\partial\ln L(\beta)}{\partial\beta}=\frac{n}{\beta}-\ln(\prod_{i=1}^{n}x_i)=0,$$

得β的最大似然估计

$$\hat{\beta}=\frac{n}{\ln\left(\prod\limits_{i=1}^{n}x_i\right)}=\left(\frac{1}{n}\sum_{i=1}^{n}\ln x_i\right)^{-1}.$$

(2)当$\beta=1$时,总体X的概率密度

$$f(x;\alpha)=\begin{cases}\dfrac{\alpha}{x^2}, & x\geqslant\alpha,\\ 0, & x<\alpha\end{cases}$$

的分段点与未知参数α有关.但注意到其似然函数

$$L(\alpha)=\begin{cases}\alpha^n\left(\prod_{i=1}^{n}x_i\right)^{-2}, & x_1,x_2,\cdots,x_n\geqslant\alpha,\\ 0, & \text{其他}\end{cases}$$

为 α 的单调增加函数,且要求 $\alpha\leqslant x_i(i=1,2,\cdots,n)$,故 α 的取值上界 $\alpha=\min\{x_1,x_2,\cdots,x_n\}$ 是 $L(\alpha)$ 的最大值点,即 α 的最大似然估计量为

$$\hat{\alpha}=\min\{X_1,X_2,\cdots,X_n\}=X_{(1)}.$$

【估计量优良性的判断】 判断估计量优良性的一般原则是:

(1)由数学期望 $E(\hat{\theta})$ 判断其无偏性;

(2)由方差 $D(\hat{\theta})$ 讨论其有效性;

(3)由大数定律分析其相容性.

例　设 $X_1,X_2,\cdots,X_m$ 是取自正态总体 $N(\mu,1)$ 的样本,$Y_1,Y_2,\cdots,Y_n$ 是取自正态总体 $N(\mu,9)$ 的样本,试求 a 和 b 使 μ 的估计量 $\hat{\mu}=a\sum\limits_{i=1}^{m}X_i+b\sum\limits_{j=1}^{n}Y_j$ 最有效.

解　由于有效性以无偏性为前提,故本题是要确定 a 和 b 在满足 $E\hat{\mu}=\mu$ 的条件下,使 $D\hat{\mu}$ 达到最小.

$$E\hat{\mu}=aE\left(\sum_{i=1}^{m}X_i\right)+bE\left(\sum_{j=1}^{n}X_j\right)$$
$$=am\mu+bn\mu=\mu,$$

即 a,b 应满足 $am+bn=1$ 或 $b=\dfrac{1-am}{n}$.

$$D\hat{\mu}=a^2\sum_{i=1}^{m}DX_i+b^2\sum_{j=1}^{n}DY_j=a^2m+b^2(9n)$$

$$=a^2m+\frac{9}{n}(1-am)^2$$

$$=\left(\frac{9m^2}{n}+m\right)a^2-\frac{18m}{n}a+\frac{9}{n},$$

令
$$\frac{\mathrm{d}D\hat{\mu}}{\mathrm{d}a}=\left(\frac{18m^2}{n}+2m\right)a-\frac{18m}{n}=0,$$

解得 $a=\dfrac{9}{n+9m}$. 注意到 $\dfrac{9m^2}{n}+m>0$,故当

$$a=\frac{9}{n+9m},\quad b=\frac{1-am}{n}=\frac{1}{n+9m}$$

时,$\hat{\mu}$为 μ 的无偏估计,且方差最小.

【无偏估计量的构造】如果问题是要寻求未知参数 θ 的无偏估计量,可按下面步骤解题:

(1)将 θ 表示为若干常见参数 θ_i 的线性组合 $\theta=\sum\limits_{i=1}^{m}a_i\theta_i$,而 θ_i 的无偏估计量 $\hat{\theta}_i$ 容易获得(见〖常见估计量的优良性〗).

(2)应用数学期望的线性性质,得到 θ 的无偏估计

$$\hat{\theta}=\sum_{i=1}^{m}a_i\hat{\theta}_i.$$

例　设$(X_1,X_2,\cdots,X_n)$为取自两点分布总体 $B(1,p)$的样本,试求 p^2 的无偏估计量.

解　由题设知总体的均值和方差分别为

$$E(X)=p,\quad D(X)=p(1-p),$$

故
$$p^2=E(X)-D(X).$$

而样本均值 $\overline{X}$ 和样本方差 S^2 分别为 $E(X)$和 $D(X)$

的无偏估计，于是有

$$E(\overline{X}-S^2)=E(\overline{X})-E(S^2)=p-p(1-p)=p^2,$$

即 $\widehat{p}^2=\overline{X}-S^2$ 是 p^2 的无偏估计.

【求参数的置信区间】构造未知参数的置信区间应遵循的如下步骤：

(1)明确待估参数；

(2)根据待估参数选取核心统计量，如对总体期望 μ 选样本均值 $\overline{X}$，对总体方差 σ^2 选样本方差 S^2；

(3)将核心统计量“加工”成具有常用统计分布的量(该量含有核心统计量和待估参数)；

(4)由(3)中分布的 $1-\alpha$ 概率区间解出待估参数的置信区间.

例　设一批零件的长度服从正态分布 $N(\mu,\sigma^2)$，其中 μ,σ^2 均未知，现从中随机抽取16个零件，测得样本均值 $\overline{x}=20$(cm)，样本标准差 $s=1$(cm)，则 μ 的置信度为0.90的置信区间是

(A) $\left(20-\dfrac{1}{4}t_{0.05}(16),20+\dfrac{1}{4}t_{0.05}(16)\right)$

(B) $\left(20-\dfrac{1}{4}t_{0.1}(16),20+\dfrac{1}{4}t_{0.1}(16)\right)$

(C) $\left(20-\dfrac{1}{4}t_{0.05}(15),20+\dfrac{1}{4}t_{0.05}(15)\right)$

(D) $\left(20-\dfrac{1}{4}t_{0.01}(15),20+\dfrac{1}{4}t_{0.01}(15)\right)$

解　因为是在 σ^2 未知的条件下估计 μ，所用 t 统计量的自由度为 $n-1=15$，所以正确答案在(C)和(D)中. 又因是双侧置信区间，所用分位点应为 $\alpha/2=0.05$，故应选(C).

第22章 假设检验

大纲要求

理解显著性检验的基本思想，掌握假设检验的基本步骤，了解假设检验可能产生的两类错误．掌握单个及两个正态总体的均值和方差的假设检验．

知识点·点拨

【参数假设】将总体的未知参数值作为假设提出来，由统计量检验其真伪．参数假设表现为两个：被检验的假设称为原假设，记作 H_0；与 H_0 的参数值不同的假设称为对立假设．也称为备择假设，记作 H_1．当假设的参数值只有一个时，称为简单假设，如 $H_0:\theta=\theta_0$，否则称为复合假设，如 $H_0:\theta\leqslant\theta_0$．

【检验的基本思想】依据“小概率事件在一次试验中不会发生”的原则，如果原假设 H_0 导致小概率事件发生．则应拒绝 H_0．

【拒绝域】设 W 是样本值 $(x_1,x_2,\cdots,x_n)$ 的一个区域，当样本值落入 W 时，否定原假设 H_0，称为 W 为 H_0 的拒绝域，或称为否定域．W 通常表现为检验统计量的值大于或小于某个边界值．

【两类错误】由于样本的随机性，假设检验的结果可能会犯如下两类错误：

(1) 当 H_0 为真时，由于样本落入 W 而否定了 H_0，称这种否定真实假设的错误为第一类错误，也称为去真错误．其概率为

$$P((x_1,x_2,\cdots,x_n)\in W\mid H_0\text{ 为真}).$$

(2)当 H_0 不真时，由于样本未落入 W 而接受了 H_0，称这种接受不真假设的错误为第二类错误，也称为存伪错误. 其概率为

$$P((x_1,x_2,\cdots,x_n)\overline{\in} W\mid H_1\text{ 为真}).$$

【显著水平】允许犯第一类错误的概率上界 α 称为显著水平或检验水平，即

$$P((x_1,x_2,\cdots,x_n)\in W\mid H_0\text{ 为真})\leqslant\alpha.$$

【正态总体均值和方差的假设检验(显著水平为 α)】

单总体 $N(\mu,\sigma^2)$：

原假设 H_0	检验统计量及其分布	对立假设 H_1	H_0 的拒绝域 W
$\mu=\mu_0$ (σ^2 已知)	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$	$\mu\neq\mu_0$	$\lvert U\rvert>u_{\alpha/2}$
		$\mu>\mu_0$	$U>u_\alpha$
		$\mu<\mu_0$	$U<-u_\alpha$
$\mu=\mu_0$ (σ^2 未知)	$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$	$\mu\neq\mu_0$	$\lvert T\rvert>t_{\alpha/2}(n-1)$
		$\mu>\mu_0$	$T>t_\alpha(n-1)$
		$\mu<\mu_0$	$T<-t_\alpha(n-1)$
$\sigma^2=\sigma_0^2$ (μ 已知)	$\chi_0^2=\dfrac{1}{\sigma_0^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$	$\sigma^2\neq\sigma_0^2$	$\chi_0^2<\chi_{1-\alpha/2}^2(n)$ 或 $\chi_0^2>\chi_{\alpha/2}^2(n)$
		$\sigma^2>\sigma_0^2$	$\chi_0^2>\chi_\alpha^2(n)$
		$\sigma^2<\sigma_0^2$	$\chi_0^2<\chi_{1-\alpha}^2(n)$
$\sigma^2=\sigma_0^2$ (μ 未知)	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\sigma^2\neq\sigma_0^2$	$\chi^2<\chi_{1-\alpha/2}^2(n-1)$ 或 $\chi^2>\chi_{\alpha/2}^2(n-1)$
		$\sigma^2>\sigma_0^2$	$\chi^2>\chi_\alpha^2(n-1)$
		$\sigma^2<\sigma_0^2$	$\chi^2<\chi_{1-\alpha}^2(n-1)$

双正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$：

原假设 H_0	检验统计量及其分布	对立假设 H_1	H_0 的拒绝域 W
$\mu_1=\mu_2$ (σ_1^2,σ_2^2 已知)	$U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ $\sim N(0,1)$	$\mu_1\neq\mu_2$	$\lvert U\rvert>u_{\alpha/2}$
		$\mu_1>\mu_2$	$U>u_\alpha$
		$\mu_1<\mu_2$	$U<-u_\alpha$
$\mu_1=\mu_2$ ($\sigma_1^2=\sigma_2^2=\sigma^2$ 未知)	$T=\dfrac{\overline{X}-\overline{Y}}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $\sim t(n_1+n_2-2)$ $S_w^2=[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2)$	$\mu_1\neq\mu_2$	$\lvert T\rvert>t_{\alpha/2}(n_1+n_2-2)$
		$\mu_1>\mu_2$	$T>t_\alpha(n_1+n_2-2)$
		$\mu_1<\mu_2$	$T<-t_\alpha(n_1+n_2-2)$
$\sigma_1^2=\sigma_2^2$ (μ_1,μ_2 已知)	$F_0=\dfrac{S_{x0}^2}{S_{y0}^2}\sim F(n_1,n_2)$ 其中 $S_{x0}^2=\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2$ $S_{y0}^2=\sum\limits_{j=1}^{n_2}(Y_j-\mu_2)^2$	$\sigma_1^2\neq\sigma_2^2$	$F_0<F_{1-\alpha/2}(n_1,n_2)$ 或 $F_0>F_{\alpha/2}(n_1,n_2)$
		$\sigma_1>\sigma_2^2$	$F_0>F_\alpha(n_1,n_2)$
		$\sigma_1^2<\sigma_2^2$	$F_0<F_{1-\alpha}(n_1,n_2)$
$\sigma_1^2=\sigma_2^2$ (μ_1,μ_2 未知)	$F=\dfrac{S_1^2}{S_2^2}\sim$ $F(n_1-1,n_2-1)$	$\sigma_1^2\neq\sigma_2^2$	$F<F_{1-\alpha/2}(n_1-1,n_2-1)$ 或 $F>F_{\alpha/2}(n_1-1,n_2-1)$
		$\sigma_1^2>\sigma_2^2$	$F>F_\alpha(n_1-1,n_2-1)$
		$\sigma_1^2<\sigma_2^2$	$F<F_{1-\alpha}(n_1-1,n_2-1)$

〖**拒绝域与置信区间**〗**假设检验的拒绝域** W **和接受域** $\overline{W}=R-W$ **是由检验统计量** T **刻画的关于样**

本值的取值区域. 如果将接受域$\overline{W}$看成未知参数θ的取值区域,它就是θ的置信区间. 例如,假设$H_0:\mu=\mu_0$,$H_1:\mu\neq\mu_0$(σ^2未知)的拒绝域为

$$|T|=\left|\frac{\overline{X}-\mu_0}{S}\sqrt{n}\right|>t_{\alpha/2}(n-1).$$

将接受域$\overline{W}=\{|T|\leqslant t_{\alpha/2}(n-1)\}$表示成$\mu_0$的区域就是$\mu$的置信度为$1-\alpha$的置信区间

$$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right].$$

因此,可以将正态总体均值和方差的置信区间表与假设检验表放在一起记忆.

常见题型·对应

【假设检验的基本步骤】

(1)根据实际问题,提出待检验的原假设H_0及其对立假设H_1.

(2)根据被检验参数,构造一个检验统计量$T(X_1,X_2,\cdots,X_n)$,使之在H_0为真时有确定的分布.

(3)由给定的显著水平α. 确定H_0的拒绝域W,使$P(T\in W)=\alpha$.

(4)由样本值计算T的值,当$T\in W$拒绝H_0,否则接受H_0.

〖原假设与对立假设〗假设检验的第一个步骤是将实际问题"翻译"成原假设H_0及其对立假设H_1,这时无论H_0,H_1是简单假设还是复合假设,都应将带等号的参数值作为H_0. 这是因为当H_0为真时,检验统计量要有确定的分布,而假设$H_0:\theta=\theta_0$,

$H_1:\theta<\theta_0$ 与假设 $H_0:\theta\geqslant\theta_0,H_1:\theta<\theta_0$ 有相同的拒绝域.例如,要检验参数 θ 是否小于 θ_0 时,应提出假设 $H_0:\theta\geqslant\theta_0,H_1\theta<\theta_0$ 而不是 $H_0:\theta<\theta_0,H_1:\theta\geqslant\theta_0$.

〖检验统计量〗假设检验的第二个步骤:建立检验统计量是假设检验的关键.其方法一般为:首先根据被检验参数选核心统计量(如检验 $H_0:\sigma^2=\sigma_0^2$ 时选 S^2;检验 $H_0:\mu_1=\mu_2$ 时选 $\overline{X}-\overline{Y}$),然后根据常用统计分布(如 $N(0,1),\chi^2,t,F$)的构造性定义,将核心统计量"加工"成具有确定分布的检验统计量(如对 S^2"加工"后,有 $\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$).

例　设某饮料销售机操作正常时,所放出的饮料量服从 $N(200,15^2)$(单位:mL).采用抽取 9 杯饮料算出平均容量 $\overline{x}$ 作周期性的检查,问:按显著水平 $\alpha=0.05$,$\overline{x}$ 的值为多少时可以认为销售机的操作正常?当 $\mu=180$ 时,这种检验犯第二类错误的概率 β 是多少?

解　销售机的操作是否正常可以表示为检验如下假设:

$$H_0:\mu=200,H_1:\mu\neq 200\quad(\sigma^2=15^2).$$

由于是检验正态总体的期望 μ,故选样本均值 $\overline{X}$ 为核心统计量,当 H_0 为真时,将 X 标准化后得到有确定分布的检验统计量

$$U=\frac{\overline{X}-200}{15/\sqrt{9}}\sim N(0,1).$$

于是对显著水平 $\alpha=0.05$,H_0 的拒绝域为

$$W=\{|U|>u_{0.025}\}=\{|\overline{X}-200|>5\times 1.96\},$$

即当 $|\overline{x}-200|>9.8$ 时拒绝 H_0，反之，当 $\overline{x}$ 在 $[190.2,209.8]$内时可以认为销售机的操作正常.

当 $\mu=180$ 时，$\overline{X}\sim N(180,15^2)$，上述检验第二类错误的概率为

$$\begin{aligned}\beta &= P\{\text{接受 } H_0 | H_1:\mu=180 \text{ 为真}\}\\ &= P\left(\frac{190.2-180}{5}\leqslant\frac{\overline{X}-180}{5}\leqslant\frac{209.8-180}{5}\right)\\ &= \Phi(5.96)-\Phi(2.04)=1-0.979=0.021.\end{aligned}$$