初等数学个人笔记

zhcosin<zhcosin@163.com>

created: 2016-04-06
last updated: 2017-03-10

序

鄙人中学时酷喜数学，在数学学科上花费了大部分时间，也小有成绩，然则高考后十年间，未曾认真学习过数学，不过囫囵吞枣的学了部分高数内容，数学水平始终未有什么进展。现今随波逐流于软件行业，然几年下来逐渐认识到，吾人最终兴趣仍然非数学莫属。

本书即为本人之数学试验场，其所收录之内容，或来自若干年前自行推得之结论，或来自于网络上与众高手切磋所得之精品，或来自某日某时于某书上发现之有价值之内容，林林总总，无论出自自己或他人，但凡本人有兴趣者，皆收入本书中。本书无终成之日，但随心而已，无论何时何地，兴之所致，皆可添加新内容。因本书之自由目的，不受限于中学数学教学大纲之内容，但求自由之数学，以达数学之自由。

以上算作新序，原序显得过于浮躁，人近中年，凡事冷静客观为好，特此作废。

目录

第一章　绪论

本章讲述一些与其它各章节没有太大直接关联的主题。

1.1　容斥原理

原理 1.1.1 (容斥原理). 用 $|A|$ 表示集合 A 的元素个数，则多个集合的并集的元素个数是:

$$\begin{aligned}|\cup_{i=1}^n A_i| = &\sum_{i=1}^{n}|A_i| - \sum_{1\leqslant i<j\leqslant n}|A_i\cap A_j| + \sum_{1\leqslant i<j<k\leqslant n}|A_i\cap A_j\cap A_k| \\ &-\cdots+(-1)^{n+1}|\cap_{i=1}^n A_i|\end{aligned} \tag{1.1}$$

容斥原理的含义借助韦恩图是显而易见的，其证明用数学归纳法即可，这里从略。

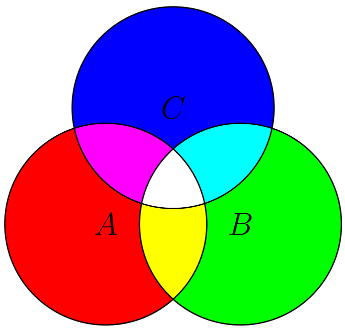

图 1.1容斥原理示意图

1.1.1) 伯努利 (Bernoulli) 信封问题　作为容斥原理的一个直接应用，这里讨论一下伯努利信封问题: 有相同数目的信封和信件若干，将这些信件装进这些信封，使得没有任何一封信件与信封搭配正确，问题是有多少种装法。

记信件数目是 n，并把信件和信封依次编号为 $1,2,\ldots,n$，信件与所属的信封编号相同。那么要计算所有信件都搭配错误的组装数，可以考虑其

反面即至少有一封信件搭配正确的组装数目，再从总数 $n!$ 中减去它即可，根据容斥原理，构造出集合 A_i 表示第 i 封信件搭配正确的组装方案，则 $|A_i| = (n-1)!$，那么至少有一封信件搭配正确的方案集合就是这些 A_i 的并集，其数目是

$$\begin{aligned}
&\sum_{i=1}^{n}|A_i| - \sum_{1\leqslant i<j\leqslant n}|A_i\cap A_j| + \sum_{1\leqslant i<j<k\leqslant n}|A_i\cap A_j\cap A_k| \\
&-\cdots+(-1)^{n+1}|\cap_{i=1}^{n}A_i| \\
=&n(n-1)! - C_n^2(n-2)! + C_n^3(n-3)! - \cdots + (-1)^{n+1}C_n^n 0! \\
=&n!\sum_{i=1}^{n}(-1)^{i+1}\frac{1}{i!}
\end{aligned}$$

所以从总数 $n!$ 中减去它就得到最后的结果，在 $0! = 1$ 的约定下，这结果可以写为:

$$\sum_{i=0}^{n}(-1)^i\frac{1}{i!} \tag{1.2}$$

欧拉对此问题有一个借助数列递推的解法，记信封数目为 n 时的结果为 a_n，那么 $a_1 = 0$，来看下这个数列的递推情况，在有 $n+1$ 个信封时，考虑编号为 1 的那封信件，除 1 号信封外它有 n 个信封可以装入，假定它装入的信封编号是 r，那再考虑编号为 r 的信件，它此时有两个选择，一是它可以装入 1 号信封，这时其它的 $n-1$ 个信件的装法是 a_{n-1}，它的另一个选择是不装入 1 号信封，这时由于 1 号信封等同于 r 号信封，所以除 1 号信件以外的 n 封信件有 a_n 种装法，于是得到该数列的递推公式为:

$$a_{n+1} = n(a_n + a_{n-1})$$

两边同除以 $(n+1)!$ 并记 $b_n = \frac{a_n}{n!}$ 可得

$$b_{n+1} = \frac{n}{n+1}b_n + \frac{1}{n+1}b_{n-1}$$

即

$$b_{n+1} - b_n = -\frac{1}{n+1}(b_n - b_{n-1})$$

所以

$$b_n - b_{n-1} = \sum_{i=2}^{n}(b_i - b_{i-1}) = \sum_{i=2}^{n}(-1)^i\frac{1}{i!}$$

于是

$$a_n = n!b_n = n!(b_1 + \sum_{i=2}^{n}(b_i - b_{i-1})) = \sum_{i=2}^{n}(-1)^i\frac{1}{i!}$$

同样在 $0! = 1$ 的约定下即有

$$a_n = \sum_{i=0}^{n} (-1)^i \frac{1}{i!}$$

∎

1.2 数学归纳法

除了常用的第一数学归纳法以外，我们还有以下的:

原理 1.2.1 (第二数学归纳法). *如果与正整数有关的命题 $P(n)$ 满足:*

1. *$P(1)$ 成立;*
2. *由 $P(1), P(2), \ldots, P(k)$ 成立能够推证出 $P(k+1)$ 成立;*

那么该命题对于一切正整数成立.

有些命题在递推过程中，依赖的是前面的所有结论而非仅仅依赖前一个结论，此时第二数学归纳法就非常适用。

原理 1.2.2 (倒推归纳法). *如果与正整数有关的命题 $P(n)$ 满足:*

1. *有无穷多个正整数 n 使命题成立;*
2. *由 $P(k)$ 成立能够推证 $P(k-1)$ 成立;*

那么该命题对于一切正整数成立。

倒推归纳法的原理也是显而易见的，对于任何一个给定的正整数 n，由于不超过 n 的正整数只有有限个，所以必然存在一个大于 n 的正整数 N，使得命题 $P(N)$ 成立，再倒推回来，知 $P(n)$ 成立。

作为一个例子，我们用倒推归纳法来证明多元均值不等式，这比用通常的第一数学归纳法来得更加容易:

命题 1.2.1 (均值不等式). *对任意 $n(n \geqslant 2)$ 个正实数 a_i，有下面不等式成立:*

$$\frac{1}{n}\sum_{i=1}^{n} a_i \geqslant \left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}}$$

证明. 对 $n = 2$ 的情形，有 $\frac{1}{2}(a_1 + a_2) - \sqrt{a_1 a_2} = \frac{1}{2}(\sqrt{a_1} - \sqrt{a_2})^2 \geqslant 0$ 知不等式成立。

反复使用 $n = 2$ 的结论，我们就可以得到当 n 是 2 的幂的时候不等式是成立的，然而 2 的幂是无穷多的，所以只要证明，不等式如果对 $n+1$ 个正实数成立就必然对 n 个正实数也成立就可以了。

对于任意 n 个正实数，我们再添加一个正实数 $a_{n+1} = \frac{1}{n}\sum_{i=1}^{n} a_i$ 构成 $n+1$ 个正实数，由假设，不等式对 $n+1$ 个正实数是成立的，所以有

$$\frac{1}{n+1}\sum_{i=1}^{n+1} a_i \geqslant \left(\prod_{i=1}^{n+1} a_i\right)^{\frac{1}{n+1}}$$

而由于 a_{n+1} 正好等于其它 n 个实数的平均数，所以

$$\frac{1}{n+1}\sum_{i=1}^{n+1} a_i = \frac{1}{n}\sum_{i=1}^{n} a_i$$

因此前一式即为:

$$\frac{1}{n}\sum_{i=1}^{n} a_i \geqslant \left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n+1}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^{\frac{1}{n+1}}$$

化简即得

$$\frac{1}{n}\sum_{i=1}^{n} a_i \geqslant \left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}}$$

即得证。 □

原理 1.2.3 (跳跃数学归纳法). *如果与正整数有关的命题 $P(n)$ 满足:*

1. *$P(1)$,$P(2)$,...,$P(m)$ 成立;*

2. *由 $P(k)$ 成立能够推证 $P(k+m)$ 成立;*

那么该命题对于一切正整数成立。

数学归纳法还有其他形式，比如解决对偶性问题（如正余弦）的螺旋归纳法，解决同时与两个正整数相关的命题的二重归纳法等，但它们的道理都是类似的。

数学归纳法的奠基也并非一定要从 1 开始。

1.3 坐标变换

在高级中学数学教材中，三角函数的伸缩变换已经为人所熟知，而本文要讨论的是坐标变换的一般性理论。

这些变换包括平移、旋转、伸缩、对称，通常所说的位似变换可以通过伸缩变换得到，其中，平移、旋转、对称都是恒等变换，即不改变平面图形的形状和大小，或者说变换前后的图形是全等的，而位似变换前后的图象则是相似的，形状不变，大小可以改变。

1.3.1 平移

平移变换由平移向量 $\vec{s}=(a,b)$ 唯一确定，平面上任意一点 P 与它在变换下的像 P' 满足 $\overrightarrow{PP'}=\vec{s}$ 为常向量，因此如果点 P 的坐标是 (x,y)，点 $P'(x',y')$ 的坐标将是 $(x+a,y+b)$，即

$$\begin{cases} x' &= x+a \\ y' &= y+b \end{cases} \tag{1.3}$$

写成矩阵形式则是:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{1.4}$$

我们把平移向量为 $\vec{s}=(a,b)$ 的平移变换记作 $S(a,b)$。

1.3.1) 本目中来建立曲线 $f(x,y)=0$ 在平移向量为 $\vec{s}=(a,b)$ 的平移变换下的新方程，假使点 $P(x,y)$ 为新曲线上任一点，则它在变换前的坐标则是 $(x-a,y-b)$，而变换前的点是满足原曲线方程的，所以得到新坐标所满足的方程 $f(x-a,y-b)=0$，此即原曲线在平移变换下的新的曲线的方程。比如说，以原点为圆心的圆 $x^2+y^2=r^2$ 在此平移变换下的新方程即为 $(x-a)^2+(y-b)^2=r^2$。对于一般函数 $y=g(x)$ 而言，把它写成 $f(x,y)=g(x)-y$ 即可应用此结论。 ■

1.3.2 旋转

假如在平面直角坐标系中，有一个点的坐标是 $P(x,y)$，现在我们把它绕着原点逆时针方向转动一个角度 θ，我们来寻求这个点的新坐标 $P'=(x',y')$ 与原坐标之间的关系。

设向量 $\overrightarrow{OP}$ 与坐标系 x 轴正向成角 α，则向量 $\overrightarrow{OP'}$ 与坐标系 x 轴正向成角 $\alpha+\theta$，记向量 $\overrightarrow{OP}$ 长度为 r，则

$$x=r\cos\alpha, y=r\sin\alpha$$

同样有

$$x'=r\cos(\alpha+\theta), y'=r\sin(\alpha+\theta)$$

于是就有

$$\begin{aligned} x' &= x\cos\theta-y\sin\theta \\ y' &= x\sin\theta+y\cos\theta \end{aligned} \tag{1.5}$$

写成矩阵形式就是:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

或者写成这样:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

我们把这个旋转变换记作 $R(\theta)$。

1.3.2) 反比例函数 $xy = 1$ 的图象位于一三象限并且关于这两个象限的角平分线对称，我们尝试把它的图象绕原点顺时针旋转 45 度，看看新的曲线方程的模样。为此目的，记反比例曲线为 C，顺时针旋转 45 度后得到的新曲线记为 C'，我们反过来把 C 看成是由 C' 绕原点逆时针旋转 45 度得到的，就有关系式 $x = \frac{\sqrt{2}}{2}(x'-y'), y = \frac{\sqrt{2}}{2}(x'+y')$，因此新曲线 C' 上的点 $P(x', y')$ 都满足方程 $\frac{\sqrt{2}}{2}(x'-y') \cdot \frac{\sqrt{2}}{2}(x'+y') = 1$，即 $x'^2 - y'^2 = 2$，这是一个离心率为 $\sqrt{2}$ 的双曲线，所以反比例函数的图象是双曲线。 ■

1.3.3 伸缩

伸缩变换是以原点为基准，把平面沿着坐标轴的两个方向进行拉伸或压缩，两个方向上各有一个 *伸缩因子*，分别用 $\lambda(>0)$ 和 $\mu(>0)$ 代表 x 和 y 方向的伸缩因子，则显然新旧坐标之间的关系是:

$$\left\{ \begin{array}{lcl} x' & = & \lambda x \\ y' & = & \mu y \end{array} \right. \tag{1.6}$$

矩阵形式为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

或者写成这样:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

如果两个方向上的伸缩因子相等，则称它为以原点为中心的 *位似变换*. 我们把伸缩因子分别为 λ 和 μ 的伸缩变换记作 $L(\lambda, \mu)$。

如果以空间的视角来看伸缩变换，它也可以看作把平面倾斜一定角度后在原平面上的投影，所以关于投影有许多跟伸缩变换完全类似的结论，比如说投影下的面积。

1.3.3) 二次曲线的离心率决定其形状 抛物线的标准方程 $y^2 = 2px(p > 0)$，对它进行位似变换，记原抛物线为 C，新的曲线为 C'，那么 C 也可以由新曲线进行位似变换得到，假定这个变换的伸缩因子为 λ，就有 $x = \lambda x', y = \lambda y'$，于是新的曲线方程是 $\lambda^2 y'^2 = 2p\lambda x'$，取 $\lambda = p$，则得 $y'^2 = 2x'$，这表明任何一个抛物线 $y^2 = 2px(p > 0)$ 的图象的形状都是一样的，不同的只是大小。事实上，同样的方法应用到圆、椭圆、双曲线上，得出的结论是: 率心率相同的二次曲线都是相似的，换句话说，离心率是确定二次曲线形状的参数，离心率相同，则形状相同，不同的只是大小。 ■

1.3.4) 伸缩变换下的面积 椭圆的面积 讨论一下伸缩变换对平面图形面积的影响，根据祖暅原理，如果两个几何体在任一水平面上的截面积都相同，那么它俩体积相等。同理也有，如果两个平面图形被某一方向上的任一直线所截得的线段长度都相等，那么它俩面积也相等，进一步，如果它俩被某一方向的任意直线所截得的线段长度都有相同的比例，那么它们面积之比也为此固定比例。而伸缩变换 $L(\lambda, \mu)$ 可以通过两次单方向的伸缩变换 $L(\lambda, 1)$ 和 $L(1, \mu)$ 来实现，易见前者把面积放大 λ 倍，后者放大 μ 倍，所以如果一个平面图形在变换前后的面积分别为 S 和 S'，那么就有

$$S' = \lambda\mu S \tag{1.7}$$

这是一个有趣的结论，比如说我们可以通过它得到椭圆的面积，因为标准椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1(a > b > 0)$ 在伸缩变换 $L(\frac{1}{a}, \frac{1}{b})$ 下将变为单位圆，单位圆面积是 π，所以椭圆面积就是 $S = \pi ab$。 ■

1.3.5) 伸缩变换下直线的斜率 椭圆的切线 在伸缩变换 $L(\lambda, \mu)$ 下，假定一条直线 l 被变换成了直线 l'(直线在伸缩变换下仍旧是直线这一事实，可以通过直线方程在变换之后仍旧是一个二元一次方程这一点上看出)，记原直线斜率为 k，新直线斜率为 k'，那么对直线上两点有

$$k' = \frac{y_1' - y_2'}{x_1' - x_2'} = \frac{\mu}{\lambda} \cdot \frac{y_1 - y_2}{x_1 - x_2} = \frac{\mu}{\lambda} k \tag{1.8}$$

因此新旧直线斜率之比是与伸缩因子有关的固定比例。有了这个结论，我们可以利用它求出椭圆曲线在任一点处的切线方程来，因为椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1(a > b > 0)$ 在伸缩变换 $L(\frac{b}{a}, 1)$ 下成为圆 $x^2 + y^2 = b^2$，因为圆的切线与过切点处的半径互相垂直，也就是斜率之积为 -1，因此椭圆上任一点处的切线的斜率，与该点与椭圆中心连线的斜率之积为 $-\frac{b^2}{a^2}$，这是一个椭圆中非常有用的结论，在椭圆上任取一点 $P(x_0, y_0)$，那么可以写出过该点的切线方程

$y-y_0=-\frac{b^2x_0}{a^2y_0}(x-x_0)$，利用 $\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}=1$ 即可将它化简为 $\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1$。■

1.3.6) 如果允许伸缩因子是虚数，那么在伸缩变换 $L(1,i)$ 下 (i 是虚数单位)，双曲线将被变换为椭圆! 然而这对现阶段的我们似乎也没有什么用处。■

还有一种变换称为 *反射变换*，也称为对称变换，即把平面上所有的点关于某一直线作对称，如果该对称轴与坐标轴成倾斜角度，可以先做旋转变换，使得对称轴与坐标轴平行或者垂直，再通过平移使对称轴与坐标轴重合，这时只需要关于坐标轴作对称之后再反平移和反旋转，即可实现一般的反射变换，所以只要讨论关于坐标轴的反射变换就可以了，然而这一点如果允许伸缩变换中的伸缩因子为负就可以得到，所以反射变换本质上是平移、旋转和伸缩的叠加，就不单独讨论了。

1.4 多项式乘幂定理

从中学数学教材中熟知有如下的二项式定理

定理 1.4.1 (二项式定理). *对于任意两个实数 x 和 y，以及任意的正整数 n，有如下等式:*

$$(x+y)^n=\sum_{i=0}^{n}C_n^ix^iy^{n-i} \tag{1.9}$$

其中每一项的系数 $C_n^i=\frac{n!}{i!(n-i)!}$ 称为 二项式系数。

利用数学归纳法，证明是很容易的，此处略去。

把这定理推广到多个数相加的情况，就有如下的多项式乘幂定理

定理 1.4.2 (多项式乘幂定理). *对于任意 m 个实数 $x_i(i=1,2,\ldots,m)$，以及任意的正整数 n，有如下等式:*

$$\left(\sum_{i=1}^{m}x_i\right)^n=\sum_{r_i\geqslant 0,\sum_{i=1}^{m}r_i=n}\frac{n!}{r_1!r_2!\cdots r_m!}x_1^{r_1}x_2^{r_2}\cdots x_m^{r_m} \tag{1.10}$$

其中的指数组合 $(r_1,r_2,\ldots,r_m)$ 要遍及方程 $\sum_{i=1}^{m}r_i=n$ 的所有非负整数解，每一项的系数 $\frac{n!}{r_1!r_2!\cdots r_m!}$ 称为 多项式系数。

对加数的个数 m 使用数学归纳法，并利用二项式定理，便可以证明此定理，此处同样略去。

这公式随着次数和加数个数的增加会迅速变长，它的项数就是方程 $\sum_{i=1}^{m}r_i=n$ 的非负整数解的个数，这个值是 C_{n+m-}^{m-1}(参见6.1小节)，比如说，对于 $n=3,m=3$ 的情况，把这公式写出来就是:

$$(a+b+c)^3=a^3+b^3+c^3+3a^2b+3a^2c+3b^2a+3b^2c+3c^2a+3c^2b+6abc$$

为了以后书写方便，我们引入所谓轮换求和和对称求和的符号，对于三个实数 a, b, c 而言，轮换求和是

$$\sum_{cyc} f(a,b,c) = f(a,b,c) + f(b,c,a) + f(c,a,b)$$

而对称求和是

$$\sum_{sym} f(a,b,c) = f(a,b,c) + f(a,c,b) + f(b,c,a) + f(b,a,c) + f(c,a,b) + f(c,b,a)$$

即轮换求和是将序列 a, b, c 首尾相接后轮换进行求和，而对称求和是要对所有可能的排列进行求和。

在这种符号下，$(a+b+c)^3$ 的展开可以写成如下这样:

$$(a+b+c)^3 = \sum_{sym} a^3 + 3\sum_{sym} a^2b + 6abc$$

而 $(a+b+c+d)^4$ 的展开则是

$$(a+b+c+d)^4 = \sum_{sym} a^4 + 4\sum_{sym} a^3b + 6\sum_{sym} a^2b^2 + 12\sum_{sym} a^2bc + 24abcd$$

现在利用这些简化的符号来表达多项式乘幂定理，可以看出，定理公式中因子次数组合相同的项都具有相同的系数，把系数提出来就是对称求和，所以这定理可以改写为

$$\left(\sum_{i=1}^{m} x_i\right)^n = \sum_{r_i \geqslant 0, \sum_{k=1}^{m} r_k = n} \frac{n!}{r_1! r_2! \cdots r_n!} \sum_{sym} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$$

1.4.1) 均值不等式 作为多项式乘幂定理的一个应用，我们来证明均值不等式: 对任意 n 个正实数 $x_i (i = 1, 2, \ldots, n)$，有下面不等式成立:

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leqslant \frac{x_1 + x_2 + \cdots x_n}{n}$$

由于证明过程写起来比较晦涩，先写出 $n = 3$ 的过程示例，以帮助理解，以 a, b, c 标记这三个数，由 $(a^k + b^k) - (a^{k-1}b + ab^{k-1}) = (a-b)(a^{k-1} - b^{k-1}) \geqslant 0$ 得 $a^k + b^k \geqslant a^{k-1}b + ab^{k-1}$，于是

$$\sum_{sym} a^2b = \frac{1}{2}\sum_{sym}(a^2b + bc^2) = \frac{1}{2}\sum_{sym} b(a^2 + c^2) \geqslant \frac{1}{2}\sum_{sym} 2bac = 6abc$$

同样

$$\sum_{sym} a^3 = \sum_{sym} \frac{a^3 + b^3}{2} \geqslant \sum_{sym} \frac{a^2b + ab^2}{2} = \frac{1}{2}\sum_{sym} a^2b \geqslant 3abc$$

所以最终

$$(a+b+c)^3=\sum_{sym}a^3+3\sum_{sym}a^2b+6abc\geqslant 3abc+18abc+6abc=27abc$$

于是三元均值不等式得证。

可以看出，这就是一个不断平衡各个因子次数的过程，它基于 $a^k+b^k\geqslant a^{k-1}b+ab^{k-1}$ 这个基本的不等式，下面我们把这个过程一般化。

只需要证明如下的不等式

$$(x_1+x_2+\cdots+x_n)^n\geqslant n^nx_1x_2\cdots x_n$$

我们考虑左边按照多项式乘幂定理展开后的对称求和的通项

$$\frac{n!}{r_1!r_2!\cdots r_n!}\sum_{sym}x_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$$

这个对称求和是针对 $r_1,r_2,\ldots,r_n$ 这个组合的各种可能的排列的，把右边的对称求和（不要系数）记作 $\sigma(r_1,r_2,\ldots,r_n)$，我们来平衡这些次数。

理想的次数平衡是每个因子的次数 r_i 都是 1，如果不是如此的话，因为这 $\sum_{i=1}^n r_i=n$，所以必然存在某两个 r_i 和 r_j 使得 $r_i-1\geqslant r_j+1$ 即 $r_i-r_j\geqslant 2$，我们把这两个次数平衡一次，变成 r_i-1 和 r_j+1，基本不等式是

$$(x_1^{r_i}y_2^{r_j}+x_2^{r_i}x_1^{r_j})-(x_1^{r_i-1}x_2^{r_j+1}+x_2^{r_i-1}x_1^{r_j+1})=x_1^{r_j}x_2^{r_j}(x_1-x_2)(x_1^{r_i-r_j-1}-x_2^{r_i-r_j-1})\geqslant 0$$

所以

$$\begin{aligned}\sigma(r_1,\ldots,r_i,\ldots,r_j,\ldots,r_n) &= \sum_{sym}x_i^{r_i}x_j^{r_j}\cdots\\ &= \frac{1}{n-1}\sum_{sym}(x_i^{r_i}x_j^{r_j}+x_i^{r_j}x_j^{r_i})\cdots\\ &\geqslant \frac{1}{n-1}\sum_{sym}(x_1^{r_i-1}x_2^{r_j+1}+x_2^{r_i-1}x_1^{r_j+1})\cdots\\ &= \frac{1}{2(n-1)}\sum_{sym}x_1^{r_i-1}x_2^{r_j+1}\cdots\\ &= \frac{1}{2(n-1)}\sigma(r_1,\ldots,r_i-1,\ldots,r_j+1,\ldots,r_n)\end{aligned}$$

■

1.5 无理指数幂的定义问题

在中学数学的课本中，没有定义指数为无理数的幂，然而却提出了定义在全实数域上的指数函数，这是因为无理指数幂的定义要用到极限，本文的

目的是为了在初等数学范围内给无理指数幂作一个解释，以解答中学生对此问题可能的疑惑。

在高等数学中，对无理指数幂的定义是，对于一个无理数 r 和一个实数 $a>0$，用任意一个以 r 为极限的有理数的序列 $r_i(i=1,2,\ldots)$ 去逼近它，无理指数幂 a^r 的定义为 $a^r=\lim_{n\to\infty}a^{r_n}$.

这个极限值是与序列 r_n 无关的，用任何一个以 r 为极限的有理数序列，所得的那个极限都是相同的。这就是高等数学中无理指数幂的定义。

接下来我们尝试在初等数学范围内解释一下无理指数幂定义。

先回顾一下在中学数学范围内是如何定义出一个实数的。

对于字面上能够写出来的数，比如整数 2 与小数 3.2，我们清楚它的每一个数位上的数字是多少，我们就认为我们定出了一个数。比如 3.2 的个位是 3，十分位为 2，其余为零。

有些数我们是不可能把它的每一位都写出来的，比如说无限循环小数 $\frac{1}{3}$，但是我们清楚它的小数部分每一位都是 3，这样我们也认为我们定出了一个数。

圆周率 π 是一个自然界中存在的常数，这个无限不循环小数我们也不能把它的每一位都写出来，甚至我们为了确定它的小数部分第 100 位数是几都得经过一番计算，但无论如何，它的每一位数我们也是能够确定的，只是这需要一些计算而已[1]。

还有一些数，比如 $\ln 2$，我们是用它所满足的一些性质来刻画它的，如果要问 $\ln 2$ 的某一数位上的数字是几，我们也需要通过一些计算步骤才能得出[2]。

所以我们可以说，我们认为我们定出了一个数，当且仅当我们能够回答出来这个数的每一个数位上的数字是多少，无论这个回答是立刻就能作出的，还是需要经过一系列的运算。

有了这点认识，我们要定义如像 $2^{\sqrt{3}}$ 这样的无理指数幂，我们只要能够按照某种规则确定出它的每一个数位上的数值就行了。

把所有实数都用小数的形式来表示，即

$$x=\cdots x_2x_1x_0.x_{-1}x_{-2}\cdots(0\leqslant x_i\leqslant 9) \tag{1.11}$$

在这种表示下，对于两个实数 x 和 y，如果从数从高到低的顺序（双端无穷序列），第一个不相同的数位上，x 的该数位大于 y 的该数位，则有 $x\geqslant y$，如果所有数位都相同，则必有 $x=y$。

上面为何是大于等于而不是大于呢，考察一下 1 与 $0.\dot{9}$ 就知道了，这两个数是相等的，因为它们的差可以小于任何一个正实数，从而这个差值

[1]这计算需要利用高等数学中的级数展开式

[2]这同样需要利用高等数学中的级数展开式

就只能是零，所以有 $1 = 0.\dot{9}$。为了能在上段结论中去掉这个等号，我们约定，如果一个无限循环小数的循环部分是 9，我们就把它收上来，也就是说 $1.12\dot{9} = 1.13$，在这样的约定下，刚才比较大小时如果第一个不相同的数位上谁大，谁的值就大。

提一下一个实数的不足近似值和过剩近似值的概念，对于任何一个实数，它的 n 位 不足近似值是

$$\underline{x_n} = \cdots x_2x_1x_0.x_{-1}x_{-2}\cdots x_{-n} \tag{1.12}$$

它的 n 位 过剩近似值是

$$\overline{x_n} = \cdots x_2x_1x_0.x_{-1}x_{-2}\cdots x_{-n} + 10^{-n} \tag{1.13}$$

通俗的说法就是，n 位不足近似值是舍掉小数点后第 n 位以后的数位，而 n 位过剩近似值则是在把这以后的数位收上来。易知序列 $\underline{x_n}$ 单调不减而序列 $\overline{x_n}$ 单调不增，而且有 $\overline{x_n} - \underline{x_n} = 10^{-n}$。

从定义可以看出，对于任何一个实数 x，不等式 $\underline{x_n} \leqslant x \leqslant \overline{x_n}$ 永远成立，但是两边的等号不能同时成立。更进一步，对于任意两个正整数 m 和 n，实际上有 $\underline{x_m} < \overline{x_n}$ 成立，这个很容易说明，假定 $m < n$，就有 $\underline{x_m} \leqslant \underline{x_n} \leqslant x \leqslant \overline{x_n}$，而式中的后两个等号又不能同时取到，所以有此结论。

现在来考虑无理指数幂，对于一个实数 $a > 0$ 和一个无理数 r，我们作出 r 的不足近似值序列 $\underline{r_n}$ 和过剩近似值序列 $\overline{r_n}$，因为 $\underline{r_n}$ 和 $\overline{r_n}$ 都是有理数，我们作两个序列 $L_n = a^{\underline{r_n}}$ 和 $M_n = a^{\overline{r_n}}$。

现在我们来证明，存在唯一一个实数 $P = P(a, r)$，它能够介于所有的 L_n 和所有的 M_n 之间，具体来说，就是在 $a > 1$ 的情况下不等式 $L_n < P < M_n$ 对于任意正整数都成立，在 $0 < a < 1$ 的情况下是 $M_n < P < L_n$ 对一切正整数恒成立。我们将会把这个实数 P 定义为 a^r，这样定义将会使指数函数在整个实数集上保持它在有理数集上同样的单调性。

只就 $a > 1$ 的情况进行论证，至于 $0 < a < 1$ 的情况，有 $\frac{1}{a} > 1$，于是可以验证 $\frac{1}{P(\frac{1}{a}, r)}$ 就符合要求。

在 $a > 1$ 的假定下对任意两个正整数 m 和 n 就有 $L_m < M_n$。这就是说，任何一个 L_m 都小于所有的 M_n，任何一个 M_n 都大于所有的 L_m。

因为

$$M_m - L_m = a^{\overline{x_m}} - a^{\underline{x_m}} = a^{\underline{x_m}}(a^{\overline{x_m} - \underline{x_m}} - 1) < a^{\overline{x_0}}(a^{\overline{x_m} - \underline{x_m}} - 1) < a^{\overline{x_0}}(a^{10^{-m}} - 1)$$

为了定出实数 P 在 10^{-n} 数位上的数值，我们让 $M_m - L_m < 10^{-n}$，这只要下式成立就可以了

$$a^{10^{-m}} < 1 + \frac{1}{10^n a^{\overline{x_0}}} \tag{1.14}$$

但是要注意的是我们现在还没有无理指数幂，所以对数的定义也就不完整，暂时也就不要尝试使用对数来从上式中解出 m 了。

现在我们需要一个引理，这个引理的内容是，对于一个实数 $a>1$ 和任何一个大于 1 的实数 T，存在一个正有理数 t，使得对任何一个满足 $0<x<t$ 的有理数 x 恒有 $a^x<T$。

这个引理也是显而易见的，只是我们现在需要把指数函数限制在有理数上，这时指数函数的单调性仍然同我们所熟知的，其它性质也与我们所熟知的在实数集上的指数函数一样。首先我们可以找到一对正整数 u 和 v，使得 $a^u<T^v$，也就是 $a^{\frac{u}{v}}<T$，于是根据定义在有理数集上的指数函数单调性，在区间 $(0,\frac{u}{v})$ 上的有理数 x 将恒有 $a^x<T$，所以引理成立。

现在回到前面的问题上来，式1.14的右端即是一个大于 1 的与 m 无关的正实数 T_n，根据引理，存在与 n 有关的有理数 t_n，使得对区间 $(0,t_n)$ 上的所有有理数 x 都成立 $a^x<T_n$，于是只要 $10^{-m}<t_n$ 即 $10^m>\frac{1}{t_n}$ 时式1.14就能成立，而 $\frac{1}{t_n}$ 显然也是一个与 m 无关的常数，所以也可以找到一个与 n 有关的有理数 s_n，使得 $m>s_n$ 时便有 $10^m>\frac{1}{t_n}$。

所以我们得到的结论就是，存在与正整数 n 有关的有理数 s_n，使得当 $m>s_n$ 时便恒有 $M_m-L_m<10^{-n}$，于是我们的要找的实数 P 就这样定义，它在 10^{-n} 这一位上的数值，与 $m>s_n$ 时 L_m 在这一位上的数值相同（容易知道 m 继续增大时，这个数位上的值不会变）。通俗的说就是，在 L_m 随着 m 增长（单调不减）的过程中，我们取它某一位上稳定下来的数值（稳定是指随着 m 无限增大都不会改变）。这样我们确实定出了一个实数，接下来需要做的事情就是，证明对于任意正整数 n 都有 $L_n<P<M_n$。

从实数 P 的作法即可知 $L_n\leqslant P$，但是由于 r 是一个无理数，它不可能从某一数位开始后面全是零，所以 $\underline{r_m}$ 随着 m 增大，虽然中途可能有所停留，但必将持续增大，因而实数 P 不可能与某个 L_m 相等，否则就与 $L_n\leqslant P$ 恒成立相矛盾了。所以对于所有正整数 n，恒有 $L_n<P$ 成立，不等式左边得证。

为了证明不等式 $P<M_n$ 对所有正整数成立，先证明实数 P 与序列 L_m 是无限接近的，也即是说，$P-L_n$ 随着 n 的增大可以小于任何正实数，这个是明显的，根据实数 P 的作法即知，存在有理数 s_n，使得 $m>s_n$ 时，$P-L_m<10^{-n}$，所以有此结论。有了这个，利用反证法就可以说明 $P<M_n$ 对所有正整数 n 成立，若不然，假定 P 能够大于某个 M_{n_0}，则有 $P-L_n\geqslant P-M_{n_0}$，右端为一常数，这与 $P-L_n$ 可以小于任何正实数相矛盾，所以实数 P 确实是介于所有的 L_n 和所有的 M_n 之间的一个实数，并且由于 M_n-L_n 可以任意小，位于这两个序列之间的实数也是唯一的。

现在，我们将把这个实数 P 定义为 a^r，因为 $L_n<P<M_n$，这个定义

使得指数函数在实数集上仍然是单调性的。

需要说明的是，以上有一些不甚严格的地方，这严格性需要建立在实数公理化体系之上。另外这基本上已经接触到确界的概念和戴德金的实数分划理论了。

第二章　函数与方程

2.1　函数概要

2.1.1　函数概念

函数是两个数集之间的一个映射，根据对应法则，数集 A 中每一个数在数集 B 中都有唯一一个数与之对应。函数通常写为 $y=f(x)$，但这并不是说，所有函数都能表示成自变量的式子，比如黎曼函数就没有解析式，而隐函数 $f(x,y)=0$ 甚至不能将 y 解成关于 x 的式子。

在讨论一些函数时，为了方便，将它的对应法则分解成嵌套的多个法则，于是得到复合函数的概念，但它并不是一类新的函数，只是认识函数对应法则的一个视角而已。

对于某些函数，由于它的对应法则的逆法则也正好满足函数定义（只要原法则下不存在多对一，则逆法则就不存在一对多，从而符合函数定义），因此自变量也就可以看成因变量的函数，这就是反函数，反函数与其原来函数是同一对应法则的两种表示方法，图象也是完全重合的，只有在互换 x 和 y 后，两者图象关于一三象限角平分线对称。

2.1.2　函数的性质

比较通用的性质是单调性，对称性（含奇偶性），周期性，凸凹性等。

单调性

单调性反应了两个变量的变化趋势，如果变化趋势一致，则为增函数，变化趋势相反则为减函数。但函数在某一区间上并不必然有某种单调性，有些函数无论你把区间划分得多么小，都没有单调性，比如狄利克雷函数和黎曼函数。

这里讨论下函数 $f(x)=x+\frac{a}{x}$ 的单调性，这里 a 是任何固定的正实数。

因为它是奇函数，奇函数在关于原点对称的区间上单调性情况相同，所以只要讨论 $x>0$ 的情况即可，此时由于

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

括号中部分是关于 x 的增函数，但是外面有平方，还得考虑它的符号，在 $x=\sqrt{a}$ 左侧为负右侧为正，所以 $f(x)$ 在 $(0,\sqrt{a}]$ 上单调减少，在 $[\sqrt{a},+\infty)$ 上单调增加，在 $x=\sqrt{a}$ 处有极小值 $f(\sqrt{a})=2\sqrt{a}$，在 x 趋近于 0 和正无穷大时，函数值亦趋向于正无穷大，而且在这两个情形下，它分别与反比例函数 $y=\frac{1}{x}$ 和正比例函数 $y=x$ 无限逼近，因此它的图象如图所示。

对称性

奇偶性是对称性的特殊情况，更一般的情况是，若函数 $f(x)$ 的图象关于直线 $x=a$ 对称，则 $f(a+x)=f(a-x)$，若它的图象关于点 (a,b) 中心对称，则 $f(a+x)+f(a-x)=2b$。

周期性

周期性反映了函数值重复取值的规律，三角函数的周期性人尽皆知。此处需要说明的是周期函数并不一定存在最小正周期，除了最为特殊的常量函数以外，狄利克雷函数（在任何有理数处函数值为 1，而任何无理数处函数值为零）也可以说明这一点，任何有理数都是它的周期，而最小的正有理数是不存在的。

凸性

凸凹性反映了函数图像的拱形特征，这个性质是一大批不等式的本源，比如说，由对数函数的上凸性即得均值不等式，再由琴生（Jensen）不等式可推得多元均值不等式，更为宽泛的加权均值不等式仍然从对数函数的上凸性获得。本章有专门讨论这一性质的小节。

2.2 关于指对函数与多项式函数的增长阶

最近人教数学官方 QQ 群接连抛出如下两道与指对函数与多项式函数的增长阶有关的题目，

题目 2.2.1. *(辽 G 教师 LiBy)* 设 $a>0$ 为实常数，讨论函数 $f(x)=ax^2-e^x$ 的极值点个数.

题目 2.2.2. *(辽 B 爱好者 1bk3)* 已知 $a>0$ 为实常数，函数 $f(x)=\ln x$，$g(x)=ax-1$
(1). 讨论函数 $h(x)=f(x)-g(x)$ 的单调性.
(2). 若函数 $f(x)$ 与 $g(x)$ 有两个不同的交点 $A(x_1,y_1)$ 和 $B(x_2,y_2)$，其中 $x_1<x_2$，求实数 a 的取值范围.

两个问题到最后引申出了如下两个子问题.

问题 2.2.1. 已知实数 $a>0$，证明存在实数 $s>\ln 2a$，使得 $e^s>2as$ 成立.

问题 2.2.2. 已知实数 $a>0$, 证明存在实数 $s>\frac{1}{a}$，使得 $\ln s<as-1$ 成立.

这两个问题的背景都是高等数学中指对函数与多项式函数的增长阶，标准答案都是通过导数解决的，现在，我们利用初等方法证明下面这两个更强的结论，它们是本文的主要结论.

命题 2.2.1. 设实数 $a>1$，则存在实数 T，使得对于任意满足 $x>T$ 的实数 x 均成立 $a^x>x^m$，这里 m 是任何正整数.

这结论的意义就是：指数函数能够在充分远的区间 $(T,+\infty)$ 上，大于任何多项式函数，只要 T 选择的足够大。

命题 2.2.2. 设实数 $a>0$，则存在实数 T，使得对于任何满足 $x>T$ 的实数 x，均成立 $\ln x<ax$，这里 m 是任何正整数.

这结论就是说：对数函数能够在充分远的区间 $(T,+\infty)$ 上比多项式中增长最慢的一次函数都还要小。

先证明命题2.2.1，其实只要证明 $m=1$ 的情况就行了，因为

$$\frac{a^x}{x^m}=\left(\frac{(a^{\frac{1}{m}})^x}{x}\right)^m$$

只要 $a>1$，就能保证 $a^{\frac{1}{m}}>1$，所以由 $m=1$ 的情况便可证明 $m>1$ 的情况。在 $m=1$ 时，不等式简化成 $a^x>x$。

因为 $a>1$，命 $a=1+\lambda(\lambda>0)$，对于自变量 x 为正整数的时候，利用二项式定理，我们有

$$a^n=(1+\lambda)^n=\sum_{k=0}^{n}C_n^k\lambda^k>C_n^2\lambda^2=\frac{1}{2}n(n-1)\lambda^2$$

所以只要 $n>1+\frac{2}{\lambda^2}$，便能保证 $a^n>n$ 成立.

根据前面所说，显然我们也已经证明了 $m>1$ 时的限于正整数情形的结论: 存在实数 T，使得对于任何满足 $n>T$ 的正整数 n，均有 $a^n>n^m$ 成立，m 是任何正整数.

现在来考虑自变量 x 的一般情形，设其整数部分为 n，$x = n + \alpha(0 \leqslant \alpha < 1)$，在刚才正整数的结论中取 $m = 2$，就有，存在实数 T，使得对于任何满足 $n > T$ 的正整数 n 都成立

$$a^x = a^{n+\alpha} = a^\alpha a^n \geqslant a^n > n^2 > n + 1 > n + \alpha = x$$

于是命题得证.

接下来再看对数函数的结论: 因为 $\ln ax < ax$ 等价于 $ax < e^{ax}$，利用指数函数的结论，存在实数 T 使得任何满足 $x > T$ 的实数 x 都成立 $e^{ax} > (ax)^2$，为了使其大于 ax，只要再限制 $x > \frac{1}{a}$ 就行，于是我们取一个新的 T'，使其同时大于 T 和 $\frac{1}{a}$ 就行了，于是命题得证.

现在来证明问题2.2.2，只要在命题2.2.2中将 a 换成 $\frac{a}{2}$ 便知存在实数 T，使得不等式 $\ln x < \frac{1}{2}ax$ 能在区间 $(T, +\infty)$ 上恒成立，显然当 $x > \frac{2}{a}$ 时有 $\frac{1}{2}ax < ax - 1$ 成立，所以只要取 T' 使得同时大于 T 和 $\frac{2}{a}$ 就行了，证毕.

今天就写到这里，以后从这里引申开，给出极限的 $\epsilon - N$ 定义后，利用本文的结论就可以证明如下两个极限:

$$\lim_{x\to+\infty} \frac{x^m}{e^x} = 0$$

和

$$\lim_{x\to+\infty} \frac{\ln x}{x} = 0$$

2.3 函数的凸性

本文讨论上凸函数与下凸函数以及与之有关的不等式。

绝大多数函数的图像都不是一条笔直的直线，它们在某个区间上经常向上凸起或者向下凹陷。现在，我们要用数学语言来给函数的凸性下个定义：

定义 2.3.1. 如果定义在区间 D 上的 连续 函数 $f(x)$ 满足：对区间 D 上任意两个不相等的实数 a、b 都成立

$$f(\frac{a+b}{2}) \geqslant \frac{f(a) + f(b)}{2} \tag{2.1}$$

则称函数 $f(x)$ 在区间 D 上是上凸的。把这不等式中的不等号反向，则得到下凸函数的定义。如果这不等式中的等号不成立，则称为严格上凸或者严格下凸。

这定义的几何意义如图2.1，本文主要讨论上凸函数，因为下凸函数只要乘以 -1，即可变为上凸函数，但本文关于上凸函数的结论都可以类似的得到下凸函数的结论。

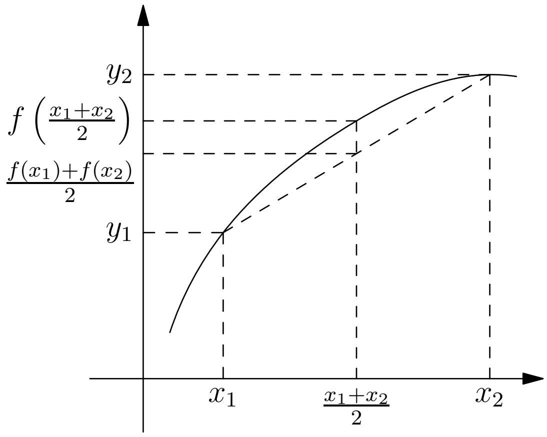

图 2.1

需要特别强调的是上述定义中对函数连续性的要求，这并不是说不连续的函数就没有凸性，而是因为这是一个不完善的定义，只有连续函数的上凸才能用上述不等式刻画，对于非连续函数而言，满足上述不等式还不足以得出它的凸性，这从下面这个函数就可以看出:

$$f(x) = \begin{cases} 0 & x \in Q \\ \sqrt{x} & x \notin Q \end{cases}$$

可以验证，这个函数满足上述定义中的不等式，然而它显然不是上凸的。关于凸函数完善的定义将在本节后文给出。

下面是一些例子。

2.3.1) 讨论下面这些函数的凸性。

1. $f(x) = x^n$(n 为正整数, $x > 0$)

2. $f(x) = \ln x$

3. $f(x) = \sin x, 0 \leqslant x \leqslant \pi$

■

解. 因为对于整数 $k(0 < k < n)$ 有 $(a^n + b^n) - (a^{n-k}b^k + a^k b^{n-k}) = (a^{n-k} -$

$b^{n-k})(a^k-b^k)>0$，因此

$$\begin{aligned}\left(\frac{a+b}{2}\right)^n+\left(\frac{b+a}{2}\right)^n &= \frac{1}{2^n}\sum_{k=0}^{n}C_n^k a^{n-k}b^k+\frac{1}{2^n}\sum_{k=0}^{n}C_n^k b^{n-k}a^k \\ &= \frac{1}{2^n}\sum_{k=0}^{n}C_n^k\left(a^{n-k}b^k+a^kb^{n-k}\right) \\ &< \frac{1}{2^n}\sum_{k=0}^{n}C_n^k\left(a^n+b^n\right) \\ &= a^n+b^n\end{aligned} \tag{2.2}$$

所以 $\frac{a^n+b^n}{2}>\left(\frac{a+b}{2}\right)^n$，函数 $f(x)=x^n$ 在正实数区间上是下凸函数。

因为 $\frac{a+b}{2}-\sqrt{ab}=\frac{1}{2}(\sqrt{a}-\sqrt{b})^2>0$，所以 $\ln\frac{a+b}{2}-\frac{\ln a+\ln b}{2}=\ln\frac{a+b}{2}-\ln\sqrt{ab}>0$，所以对数函数在正实数区间上是上凸函数。

因为 $\frac{1}{2}(\sin a+\sin b)=\sin\frac{a+b}{2}\cos\frac{a-b}{2}<\sin\frac{a+b}{2}$，所以正弦函数在区间 $(0,\pi)$ 上是上凸函数。 □

利用数学归纳法，可以很容易的将定义中的两个数推广到任意个数的情形，这就是著名的琴生 *(Jensen)* 不等式:

定理 2.3.1 (琴生 (Jensen) 不等式). *如果函数 $f(x)$ 是区间 D 上的上凸函数，则对于此区间上的任意 n 个实数 $x_i(i=1,2,\dots,n)$，成立着不等式:*

$$f(\frac{1}{n}\sum_{i=1}^{n}x_i)\geqslant\frac{1}{n}\sum_{i=1}^{n}f(x_i) \tag{2.3}$$

证明. 事实上，与第一数学归纳法相比，倒推归纳法更容易证明本定理，读者不妨一试，本文仍使用第一数学归纳法。

由定义知 $n=2$ 的情形成立。现在假定对 n 的情形成立，则对于 $n+1$ 个数的情形，记 $X_{n+1}=\frac{1}{n+1}\sum_{i=1}^{n+1}x_i$，则

$$\begin{aligned}f(X_{n+1}) &= f\left(\frac{x_1+x_2+\cdots+x_{n+1}}{n+1}\right) \\ &= f\left(\frac{(x_1+\cdots+x_n)+(x_{n+1}+(n-1)X_{n+1})}{2n}\right) \\ &= f\left(\frac{\frac{x_1+\cdots+x_n}{n}+\frac{x_{n+1}+(n-1)X_{n+1}}{n}}{2}\right) \\ &\geqslant \frac{f\left(\frac{x_1+\cdots+x_n}{n}\right)+f\left(\frac{x_{n+1}+(n-1)X_{n+1}}{n}\right)}{2} \\ &\geqslant \frac{\frac{f(x_1)+\cdots+f(x_n)}{n}+\frac{f(x_{n+1})+(n-1)f(X_{n+1})}{n}}{2}\end{aligned} \tag{2.4}$$

化简即得定理结论。 □

琴生不等式是一个非常重要的不等式，它是一系列不等式的重要来源，例如把它应用到对数函数身上，就得出多元均值不等式：

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geqslant \sqrt[n]{\prod_{i=1}^{n} x_i}$$

另一个例子，因为正弦函数在区间 $(0,\pi)$ 上是上凸函数，所以对于一个三角形的三个内角，由琴生不等式就有：

$$\sin A+\sin B+\sin C \leqslant 3\sin\frac{A+B+C}{3}=\frac{3\sqrt{3}}{2}$$

最后我们来证明：圆的内接凸多边形中，以正多边形的面积为最大。只要记每一条边所对应的圆心角为 $\theta_i(i=1,2,\ldots,n)$，凸多边形的面积

$$\begin{aligned}
S &= \frac{1}{2}R^2\sum_{i=1}^{n}\sin\theta_i \\
&\leqslant \frac{n}{2}R^2\sin\left(\frac{1}{n}\sum_{i=1}^{n}\theta_i\right) \\
&= \frac{n}{2}R^2\sin\frac{2\pi}{n}
\end{aligned}$$

而等号在 $\theta_i=\frac{2\pi}{n}(i=1,2,\ldots,n)$ 时成立，即为正多边形。

从函数图象的几何意义上看，前面的定义仅仅刻画了区间中点处的函数值与区间两个端点函数值的关系，但从直觉上看，作为上凸函数，它在区间上任意一点的函数值，都大于连接函数曲线段 (由区间所截得) 两端点连线所代表的一次函数的函数值，换句话说，定义中的 $\frac{1}{2}$ 应该可以推广到 0 与 1 之间的任意实数。但是下面的定理表明，仅仅由前面的定义，只能将这个 $\frac{1}{2}$ 推广到 0 与 1 之间的有理数，而无理数的情形要依赖于函数的连续性才能得出。也正是由于这个原因，通常把下面定理中这个更强的不等式来作为上凸函数的定义。

定理 2.3.2. 如果函数 $f(x)$ 是定义在区间 D 上的上凸函数，并且在此区间上连续，则对于区间上任意两个实数 a、b，以及任意两个满足 $\alpha+\beta=1(\alpha\geqslant 0,\beta\geqslant 0)$ 的实数 α、β，都成立不等式[1]

$$f(\alpha a+\beta b)\geqslant \alpha f(a)+\beta f(b) \tag{2.5}$$

证明. 先证明 α 是有理数的情形 (β 也就同时为有理数)，对于满足 $\alpha+\beta=1$ 的非负有理数 α 和 β，必定存在不同时为零的非负整数 m 和 n，使得

$$\alpha=\frac{m}{m+n},\beta=\frac{n}{m+n}$$

[1] 也有的书上是用这个定理来作为凸函数定义的。

因此，在下面中将 $ma+nb$ 视为 m 个 a 与 n 个 b 分别相加之后求和，再利用琴生不等式，得到

$$\begin{aligned} f(\alpha a+\beta b) &= f\left(\frac{ma+nb}{m+n}\right) \\ &\geqslant \frac{mf(a)+nf(b)}{m+n} \\ &= \alpha f(a)+\beta f(b) \end{aligned}$$

因此有理数的情形得证。

对于 α 为无理数的情形,存在有理数的序列 α_n 和 β_n,使得 $\lim_{n\to\infty}\alpha_n=\alpha$ 和 $\lim_{n\to\infty}\beta_n=\beta$，由

$$f(\alpha_n a+\beta_n b)\geqslant \alpha_n f(a)+\beta_n f(b)$$

不等式两端令 $n\to\infty$，并利用函数 f 的连续性即得无理数时不等式仍成立。 □

现在来看一些例子，前面已经证明过对数函数在其定义域上是上凸函数，因此套用刚证明过的定理，就得到，对于任意正实数 a,b 以及满足 $\alpha+\beta=1(\alpha>0,\beta>0)$ 的正实数 α,β，都成立不等式

$$\alpha a+\beta b\geqslant a^{\alpha}b^{\beta} \tag{2.6}$$

这称为加权均值不等式（二元），而通常情况下的均值不等式 $\frac{a+b}{2}\geqslant\sqrt{ab}$，则不过是其中 $\alpha=\beta=\frac{1}{2}$ 的特殊情况而已。

这里再为这个重要定理提供另外一个证明，它不使用琴生不等式，而是直接从定义出发，不过，篇幅稍长些。

先提出如下两个引理:

引理 2.3.1. *区间上的凸函数，加上或者减去同一区间上的一次函数，不改变凸性。*

这个结论是明显的，因为对于一次函数，式2.1中的等号总是成立的，因此在一个凸函数的式2.1两端同时减去一次函数的该式两端，不等式仍然成立。

引理 2.3.2. *如果函数 $f(x)$ 是区间 I 上的上凸函数，a 和 b 是区间 I 上两个不相等的固定实数，则函数 $h(t)=f((1-t)a+tb)$ 是关于 t 的定义在区间 $[0,1]$ 上的上凸函数。*

证明. 记 $g(t)=(1-t)a+tb$，则 $h(t)=f(g(t))$, 而 $g(t)$ 作为一次函数，显然对于任意 t_1 和 t_2 成立着不等式:

$$g\left(\frac{t_1+t_2}{2}\right)=\frac{g(t_1)+g(t_2)}{2}$$

因此有

$$\begin{aligned} h\left(\frac{t_1+t_2}{2}\right) &= f\left(g\left(\frac{t_1+t_2}{2}\right)\right) \\ &= f\left(\frac{g(t_1)+g(t_2)}{2}\right) \\ &\geqslant \frac{f(g(t_1))+f(g(t_2))}{2} \\ &= \frac{h(t_1)+h(t_2)}{2} \end{aligned}$$

即得证。 □

从这证明过程可以看出，实际上我们可以得到更为通用的关于复合函数凸性的结论：

定理 2.3.3. *若函数 $g(x)$ 是区间 I 上的上凸函数，并且在该区间上的值域为 D，而函数 $f(t)$ 是区间 D 上的单调增加的上凸函数，则复合函数 $h(x)=f(g(x))$ 是区间 I 上的上凸函数，其它条件组合也有类似的结论。*

它的证明仿照引理2.3.2即可，这里从略。

现在来证明定理2.3.2

证明. 将不等式左边减去右边的差记为 $g(\alpha)$，则只要证明 $g(\alpha)\geqslant 0$ 对于一切 $0\leqslant\alpha\leqslant 1$ 恒成立即可。

由于完整的证明需要用到高等数学中连续性的定义，所以这里仅就 α 是有理数的情形进行证明 (当然 β 也就同时为有理数，而且函数连续的条件也就用不上)。

现在我们证明：对于任意正整数 n，不等式 $g(\frac{i}{n})\geqslant 0(i=0,1,\dots,n)$ 都成立。

首先 $g(0)=g(1)=0$，并且利用上面两个引理可以得知 $g(\alpha)$ 仍是一个关于 α 的上凸函数，因此

$$g(\frac{i}{n})\geqslant\frac{1}{2}\left(g(\frac{i-1}{n})+g(\frac{i+1}{n})\right)$$

也即是

$$g(\frac{i+1}{n})-g(\frac{i}{n})\leqslant g(\frac{i}{n})-g(\frac{i-1}{n})$$

令 $\lambda_i=g(\frac{i+1}{n})-g(\frac{i}{n})$，则有 $\lambda_i\leqslant\lambda_{i-1}(i=1,2,\dots,n-1)$，而

$$0=g(1)-g(0)=\sum_{i=0}^{n-1}\left(g(\frac{i+1}{n})-g(\frac{i}{n})\right)$$

即是 $\sum_{i=0}^{n-1}\lambda_i=0$，因此存在正整数 $k(0<k<n)$，使得

$$\lambda_0 \geqslant \cdots \geqslant \lambda_{k-1} \geqslant 0 \geqslant \lambda_k \geqslant \ldots \geqslant \lambda_{n-1}$$

也即是

$$0=g(0) \leqslant g(\frac{1}{n}) \ldots \leqslant g(\frac{k}{n}) \geqslant \ldots \geqslant g(\frac{n-1}{n}) \geqslant g(1)=0$$

因此，$g(\frac{i}{n}) \geqslant 0(i=0,1,\ldots,n)$ 成立，有理数的情形得证。而无理数的情形仍同前面的证明。 □

同样的，我们还可以把定理2.3.2从两个数的情况推广到多个数的情形:

定理 2.3.4. *设函数 $f(x)$ 在区间 D 上是上凸函数，那么对于该区间上任意 n 个实数 $x_i(i=1,2,\ldots,n)$ 以及相应的权重 $\alpha_i(\sum_{i=1}^n\alpha_i=1,\alpha_i\geqslant 0)$，有下面的不等式成立:*

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \geqslant \sum_{i=1}^n \alpha_i f(x_i) \tag{2.7}$$

其证明与定理2.3.1类似，此处从略。

我们还是把这定理套用到对数函数身上，得到如下的加权均值不等式：对于一组实数 $x_i(i=1,2,\ldots,n)$ 及一组权值 $\alpha_i>0(i=1,2,\ldots,n),\sum_{i=1}^n\alpha_i=1$，成立着不等式:

$$\sum_{i=1}^n \alpha_i x_i \geqslant \prod_{i=1}^n x_i^{\alpha_i} \tag{2.8}$$

等号成立的条件是 x_i 全部都相等。它也可被视为均值不等式的推广。

第三章　数列

3.1　数列概要

3.1.1　数列概念

数列就是一个数的序列，可以是有限序列也可以是无限序列，按下标记法可以写成 $\cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots$，如果有一个公式可以把数列的每一项表成它的下标的函数，就称此公式为该数列的通项公式。数列本质上是定义在整数集或者正整数集上的函数。

3.1.2　等差数列与等比数列

等差数列和等比数列是最常见的两种数列。

等差数列的每一项减去它前一项所得的差值都相同，这差值称为此等差数列的公差。

对于等差数列中的任意两项（无论顺序）有

$$a_n - a_m = (n-m)d$$

证明是容易的，只要证明 $n \geqslant m$ 的情况即可（否则可以反过来相减），首先可以验证 n 与 m 相等时该等式成立，而后有

$$a_n - a_m = \sum_{k=m}^{n-1}(a_{k+1} - a_k) = (n-m)d$$

假定我们知道 a_0 和公差 d，在上式中取 $m=0$，我们就得到等差数列的通项公式

$$a_n = a_0 + nd$$

它是关于 n 的一次函数，实际上，通项是关于下标的一次函数的数列必然是等差数列。如果从 a_0 开始，向正下标或者负下标进行累加，n 是任意一

个整数（无论正负），记 $S_n=\sum_{i=0}^{n}a_n$，则

$$S_n=(n+1)a_0+d\sum_{i=0}^{n}i=(n+1)a_0+\frac{1}{2}n(n+1)d$$

对于等比数列，如果它的各项都是正的，取对数即可变身为一个等差数列。对于所有的等比数列，同样可以得出类似于等差数列的结论来，即是说，对于公比为 q 的等比数列 $\cdots,a_{-2},a_{-1},a_0,a_1,a_2,\cdots$，仍然有

$$\frac{a_n}{a_m}=q^{n-m}$$

并由此得通项 $a_n=a_0q^n$。

3.1.3 递推数列的通项

线性递推数列的通项

分式型递推数列的通项

3.2 线性递推数列的通项

本节讨论常系数线性递推数列的通项求法问题，这个都是有固定结论的内容，本文只是粗略转叙一番而已。

所谓常系数线性递推数列，是指它的递推公式形如 $\sum_{k=1}^{n}\lambda_k a_k=c$ 的数列，例如斐波那契数列 $a_{n+2}=a_{n+1}+a_n$。

先看最简单的一种，递推式为 $a_{n+1}=pa_n+q$ 的数列，当 $p=1$ 时它成为等差数列，当 $q=0$ 则成为等比数列，所以此处限定 $p\neq 1,q\neq 0$。

要求它的通项，只要在它两端同时除以 p^{n+1}，就有

$$\frac{a_{n+1}}{p^{n+1}}=\frac{a_n}{p^n}+\frac{q}{p^{n+1}}$$

因此有

$$\begin{aligned}\frac{a_n}{p^n} &= \frac{a_1}{p}+\sum_{k=2}^{n}\left(\frac{a_k}{p^k}-\frac{a_{k-1}}{p^{k-1}}\right)\\ &= \frac{a_1}{p}+\sum_{k=2}^{n}\frac{q}{p^k}\end{aligned}$$

剩下的就是对一个等比数列进行求和了。

另外一种方法比较巧妙，假定存在一个实数 λ，使得 $a_{n+1}+\lambda=p(a_n+\lambda)$，展开与原递推式比较即得 $\lambda=\frac{q}{p-1}$，于是数列 $a_n-\lambda$ 就成为一个等比数列了。

现在看高一阶的例子,每一项需要它前面两项才能确定:$a_{n+2}=pa_{n+1}+qa_n$，假想有两个实数 r 和 s 能够使得

$$a_{n+2}-ra_{n+1}=s(a_{n+1}-ra_n) \tag{3.1}$$

展开后与原递推式比较可得

$$r+s=p$$
$$rs=-q$$

因此 r 和 s 是方程 $x^2=px+q$ 的两个根，所以在此方程确实有两个根 (可以相等) 的情况下，数列 $a_{n+1}-ra_n$ 成为等比数列，求出它的通项 $a_{n+1}-ra_n=f(n)$ 后，只要两端同时除以 r^{n+1} 即可求出 a_n 的通项。

如果这两个根不相等，则还有另一种求法，因为 r 和 s 都是方程 $x^2=px+q$ 的根，因此既然有3.1成立，也就必然有

$$a_{n+2}-sa_{n+1}=r(a_{n+1}-sa_n) \tag{3.2}$$

成立，令 $b_n=a_{n+1}-ra_n, c_n=a_{n+1}-sa_n$，则 $b_{n+1}=sb_n, c_{n+1}=rc_n$，所以 $b_n=s^{n-1}b_1, c_n=r^{n-1}c_1$，于是

$$a_n-ra_{n-1}=s^{n-2}b_1$$
$$a_n-sa_{n-1}=r^{n-2}c_1$$

从中解出 a_n 来:

$$a_n=\frac{r^nc_1-s^nb_1}{r-s}$$

若取 $\alpha_1=\frac{c_1}{r-s}, \alpha_2=-\frac{b_1}{r-s}$，则得到

$$a_n=\alpha_1r^n+\alpha_2s^n$$

这两个系数 α_1 和 α_2 也不太容易记住，但现在我们知道了通项的最终形式是这两个根的幂的线性组合，所以可以把这两个系数作为待定系数，利用 $a_1=\alpha_1r+\alpha_2s$ 和 $a_2=\alpha_1r^2+\alpha_2s^2$ 把它们解出来。

更一般的情形是:对于线性递推数列 $\sum_{k=1}^n\lambda_ka_k=c$,称方程 $\sum_{k=1}^n\lambda_kx^k=c$ 为它的特征方程，在复数范围内这个特征方程必有 n 个根（重根按重数计算)，假定这些根是 $x_i(i=1,2,\cdots,m,m\le n)$，相应根的重数是 $r_i(i=1,2,\cdots,m,\sum_{i=1}^mr_i=n)$，则它的通项是：

$$a_n=\sum_{i=1}^mP_{r_i-1}(n)x_i^n$$

上式中 $P_{r_i-1}(n)$ 表示一个关于 n 的次数是 r_i-1 的多项式，如果哪个根是单重根，则它的系数是常数。式中的系数都可以利用待定系数法来求。

作为一个例子，现在来求斐波那契数列的通项，递推公式为 $a_{n+2}=a_{n+1}+a_n$，特征方程是 $x^2=x+1$，其两个根是 $x_{1,2}=\frac{1}{2}(1\pm\sqrt{5})$，于是通项应为：

$$a_n=\alpha_1x_1^n+\alpha_2x_2^n$$

将 $a_1=1$ 和 $a_2=1$ 带入求出两个系数，最后得：

$$a_n=\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n-\left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

令人惊讶的是一个所有项都是正整数的数列，其通项居然出现了无理数，事实上，利用二项定理可以证明，这个表达式将永远是正整数。

3.3 正整数的幂和公式

本文来计算正整数的幂和，也就是求 $S_m(n)=\sum_{k=1}^n k^m$ 的关于 n 的公式。

显然有 $s_0(n)=n$，然后利用二项式展开有

$$(k+1)^m-k^m=\sum_{i=0}^{m-1}C_m^ik^i$$

于是

$$\begin{aligned}(n+1)^m-1&=\sum_{k=1}^n((k+1)^m-k^m)\\&=\sum_{i=0}^{m-1}C_m^i\sum_{k=1}^nk^i\\&=\sum_{i=0}^{m-1}C_m^iS_i(n)\end{aligned}\tag{3.3}$$

因此，在上式中取 $m=2$ 得

$$(n+1)^2-1=S_0(n)+2S_1(n)$$

于是（结果已经进行因式分解）

$$S_1(n)=\frac{1}{2}n(n+1)$$

继续取 $m=3$ 得

$$(n+1)^3-1=S_0(n)+3S_1(n)+3S_2(n)$$

并将 $S_0(n)$ 和 $S_1(n)$ 代入即得

$$S_2(n) = \frac{1}{6}n(n+1)(2n+1)$$

同样可以继续得到

$$S_3(n) = \frac{1}{4}n^2(n+1)^2$$

继续下去，理论上可以得到正整数的任意正整数次方的和式。

从3.3还可以得出：$S_m(n)$ 是关于 n 的 $m+1$ 次多项式。

3.4 关于正整数倒数和的一些讨论

本节讨论与正整数倒数和相关的一些问题。

首先是 $S_n = \sum_{k=1}^{n} \frac{1}{k}$，这是一个递增的数列，但是增量逐次减小趋于零，所以一个很自然的问题就是：这数列是有一个上限呢，还是可以无限增加到正无穷。

结论是它在增加过程中可以超过任何一个正实数，无论多大的正实数，也就是说，它是一个正无穷大。

先叙述一个在许多书上都可以找到的证明，这个证明是这样的，把正整数序列进行分段，1 为第零段，2 为第一段，3 和 4 为第二段，5 到 8 为第三段，一般的第 n 段是从 $2^{n-1}+1$ 到 2^n，即每一段的最后一个数都刚好是 2 的幂，于是对于每一段上的正整数们的倒数，显然最后一个正整数的倒数是最小的，所以该段上的正整数的倒数和大于 $2^{n-1}\cdot\frac{1}{2^n}=\frac{1}{2}$，而按照这个方法，这个段是可以有无穷多的，只是越往后，段的长度就越长，所以这就说明它了 S_n 是无限增加的。

再提一个思路完全一样只是分段不同的证明，可以证明，对于正整数 $n>3$，下面不等式成立:

$$\frac{1}{n+1}+\frac{1}{n+2}\cdots+\frac{1}{3n}>1$$

实际上，左边是一个递增的（作差即知），而在 $n=4$ 时左边已经大于 1，所以得出。

如此 1、2、3、4 各自独立为一段，以后 5 到 12 为一段，13 到 36 为一段，如此下去，也能达到同样的目标。

3.5 题集

题目 3.5.1. *已知各项都是正实数的数列 x_n 对一切正整数 n 都成立 $x_n+\frac{1}{x_{n+1}}<2$，求证该数列所有项都满足 $x_n<1$.*

解答. 如果用上极限理论，则可以很容易的得出它单调增加并以 1 为极限，结论不证自明，所以这里主要讨论的是初等证明。

因为

$$x_n + \frac{1}{x_{n+1}} < 2 \leqslant x_{n+1} + \frac{1}{x_{n+1}}$$

所以 $x_n < x_{n+1}$，即该数列单调增加。

又显然 $x_n < 2$，所以

$$2 > x_n + \frac{1}{x_{n+1}} > x_n + \frac{1}{2}$$

于是 $x_n < 2 - \frac{1}{2}$，我们得到一个更加好的上限，重复这个过程，我们由 $x_n < y_m$ 就可以得到

$$x_n < 2 - \frac{1}{y_m}$$

所以我们作数列 y_m，它由 $y_1 = 2$ 和

$$y_{m+1} = 2 - \frac{1}{y_m}$$

来确定。

数列 y_m 的每一项都大过数列 x_n 的全部项，所以它的下标特意用 m 而不是 n 来表示，以示不相关。

现在来求 y_m 的通项公式，由于

$$\frac{1}{y_{m+1} - 1} = 1 + \frac{1}{y_m - 1}$$

因此数列 $\frac{1}{y_m - 1}$ 是等差数列，它的通项为 $y_m = 1 + \frac{1}{m}$[1]，于是 $x_n < y_m$ 对一切正整数 n 和 m 都成立，所以必定有 $x_n \leqslant 1$（用反证法），而 x_n 的单调性则保证了等号是不能取的。

下面关于 y_m 再给个不求通项的玩法[2]，$y_m > 1$ 这一点根据数学归纳法是明显成立的。下面证明它可以任意接近 1，也就是要证明，对于无论多么小的正实数 δ，总存在 y_m 中的某一项 y_M，使得 $y_M < 1 + \delta$。采用反证法，假定存在某个正实数 δ，使得 y_m 中的所有项都满足 $y_m \geqslant 1 + \delta$，则

$$y_{m+1} - 1 = \frac{1}{y_m}(y_m - 1) \leqslant \frac{1}{1+\delta}(y_m - 1)$$

于是

$$y_m - 1 \leqslant \frac{1}{(1+\delta)^m}$$

显然与假设矛盾，故得证。 □

[1]这是分式型递推数列求通项的不动点解法。

[2]其实这是高数的玩法，就差提到确界二字了。

题目 3.5.2. 数列 a_n 满足：$a_1=\frac{1}{2}$，$a_{n+1}=\frac{na_n+a_n^2}{n+1}$，

1. 求证该数列是递减的。

2. 求证 $a_n<\frac{7}{4n}$

解答. 根据数列归纳法易知 $0<a_n<1$，所以

$$a_{n+1}=\frac{na_n+a_n^2}{n+1}<\frac{na_n+a_n}{n+1}=a_n$$

所以数列递减，第二问，只要证明 $n>3$ 时有如下更强的不等式即可（使用数学归纳法，过程略去）

$$a_n\leqslant\frac{7}{4n}-\frac{4}{n^2}$$

□

题目 3.5.3. 记 $I_n=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+\frac{1}{2n-1}-\frac{1}{2n}$，求证，在正整数 $n\geqslant 100$ 时，有 $0.68<I_n<0.7$.

证明. 记 $J_n=\sum_{k=1}^{n}(\frac{1}{2k}-\frac{1}{2k+1})$，由于 $z_n=\frac{1}{n-1}-\frac{1}{n}=\frac{1}{n(n-1)}$ 是递减的，并且相邻两项也相差越来越小，所以有不等式 $z_{2k}<\frac{1}{2}(z_{2k-1}+z_{2k+1})$, 也就是如下的:

$$\frac{1}{2k-1}-\frac{1}{2k}<\frac{1}{2}\left[\left(\frac{1}{2k-2}-\frac{1}{2k-1}\right)+\left(\frac{1}{2k}-\frac{1}{2k+1}\right)\right]$$

对上式左边进行累加，但从 $k=3$ 到 $k=n$ 使用右边放缩，得

$$I_n<\frac{1}{2}+\frac{1}{12}+\frac{1}{2}\left[\left(J_n-\frac{1}{6}-\frac{1}{2n(2n+1)}\right)+\left(J_n-\frac{1}{6}-\frac{1}{20}\right)\right]$$

化简

$$I_n<J_n+\frac{47}{120}-\frac{1}{4n(2n+1)}$$

利用 $I_n+J_n=1-\frac{1}{2n+1}$ 从上式中换掉 J_n 得

$$I_n<\frac{167}{240}-\frac{1}{2(2n+1)}-\frac{1}{8n(2n+1)}<\frac{167}{240}<\frac{168}{240}=0.7 \tag{3.4}$$

于是不等式的右边得证，接下来考虑左边不等式，同样因为 z_n 是递减的，有不等式 $z_{2k}>\frac{1}{2}(z_{2k}+z_{2k+1})$，也就是

$$\frac{1}{2k-1}-\frac{1}{2k}>\frac{1}{2}\left[\left(\frac{1}{2k-1}-\frac{1}{2k}\right)+\left(\frac{1}{2k}-\frac{1}{2k+1}\right)\right]$$

对左边进行累加，在 $k\geqslant 3$ 时使用右边放缩，得到

$$I_n>\frac{1}{2}+\frac{1}{12}+\frac{1}{2}\left(\frac{1}{5}-\frac{1}{2n+1}\right)$$

也就是

$$I_n > \frac{41}{60} - \frac{1}{2(2n+1)} \tag{3.5}$$

在 $n \geqslant 100$ 时，有

$$I_n \geqslant I_{100} > \frac{41}{60} - \frac{1}{402} = 0.680845771... > 0.68$$

所以不等式左边得证.

其实证明左边所用的放缩是比较松的，实际上因为 $\lim_{n\to\infty} I_n = \ln 2 = 0.693147...$，所以左边不等式的放缩余地较大，所以这样的放缩也能达到要求，现在来尝试使用更强的放缩，看看能得到一个什么样的结果。

对于 z_n，不等式 $z_n > 2z_{n+1} - z_{n+2}$ 将是一个更强的放缩，所以我们有 $z_{2k} > 2z_{2k+1} - z_{2k+2}$，也就是下面的不等式

$$\frac{1}{2k-1} - \frac{1}{2k} > 2\left(\frac{1}{2k} - \frac{1}{2k+1}\right) - \left(\frac{1}{2k+1} - \frac{1}{2k+2}\right)$$

对上式左边进行累加，但只在 $k \geqslant 3$ 时使用右边放缩，即得

$$I_n > \frac{1}{2} + \frac{1}{12} + 2\left(J_n - \frac{1}{6} - \frac{1}{20}\right) - \left(I_n - \frac{1}{2} - \frac{1}{12} - \frac{1}{30} + \frac{1}{(2n+1)(2n+2)}\right)$$

将其中的 J_n 用 $1 - I_n - \frac{1}{2n+1}$ 替换掉，即得

$$I_n > \frac{83}{120} - \frac{1}{2(2n+1)} - \frac{1}{4(2n+1)(2n+2)}$$

因此在 $n \geqslant 100$ 时，便有

$$I_n \geqslant I_{100} > \frac{83}{120} - \frac{1}{2(2\times 100+1)} - \frac{1}{4(2\times 100+1)(2\times 100+2)} = 0.6891729471.....$$

这个值已经非常接近 0.69 了。 □

题目 3.5.4. *已知数列* $\{a_n\}$ *满足:* $a_1 = 1$，$a_{n+1} = a_n + \frac{1}{a_n^2}$，*求证* $a_{2015} > 18$.

证明. 易见这是一个递增的正项数列，在递推式两边同时三次方:

$$a_{n+1}^3 = \left(a_n + \frac{1}{a_n^2}\right)^3 = a_n^3 + 3 + \frac{3}{a_n^3} + \frac{1}{a_n^6} > a_n^3 + 3$$

所以 $a_{2015} > a_1^3 + 3 \times 2014 = 6043 > 5832 = 18^3$. □

第四章　不等式

4.1　一些重要的不等式

本节叙述一些在初等数学领域非常重要的不等式，这些不等式应用广泛，刻画的不等式关系非常深刻。

4.1.1　均值不等式

定理 4.1.1 (均值不等式). 对于任意 n 个正实数 $x_i(i = 1, 2, \ldots, n)$，记

$$H_n = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \tag{4.1}$$

$$G_n = (\prod_{i=1}^{n} x_i)^{\frac{1}{n}} \tag{4.2}$$

$$A_n = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.3}$$

$$Q_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \tag{4.4}$$

称 H_n 为这 n 个正实数的 调和平均数，而 A_n 为 算术平均数，G_n 为 几何平均数，Q_n 为 平方平均数，则有下面不等式成立

$$H_n \leqslant G_n \leqslant A_n \leqslant Q_n \tag{4.5}$$

其中的等号只有在 x_i 全部都相等时才成立。

我们在1.4节中曾经利用多项式乘幂定理证明过中间的 $A_n \geqslant G_n$，在这一节，我们使用其它的方法来证明。

证明. 先来证明 $A_n \geqslant G_n$.

我们在小节1.2中已经使用倒推形式的数学归纳法证明过这个不等式，所以在这一节我们使用第一数学归纳法。

因为 $\frac{x_1+x_2}{2}-\sqrt{x_1x_2}=\frac{1}{2}(\sqrt{x_1}-\sqrt{x_2})^2\geqslant 0$，所以不等式 $A_2\geqslant G_2$ 成立。

假定 $A_n\geqslant G_n$，来尝试证明 $A_{n+1}\geqslant G_{n+1}$，因为 A_{n+1} 是一个算术平均数，对一组数而言，如果新添加的数等于原来这些数的算术平均数，则算术平均数保持不变，于是有如下的策略

$$\begin{aligned}
A_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1}x_i\\
&= \frac{\sum_{i=1}^{n}x_i+[x_{n+1}+(n-1)A_{n+1}]}{2n}\\
&= \frac{1}{2}\left(\frac{x_1+\cdots+x_n}{n}+\frac{x_{n+1}+(n-1)A_{n+1}}{n}\right)\\
&\geqslant \frac{1}{2}\left(\sqrt[n]{x_1\cdots x_n}+\sqrt[n]{x_{n+1}A_{n+1}^{n-1}}\right)\\
&\geqslant \sqrt{\sqrt[n]{x_1\cdots x_n}\cdot\sqrt[n]{x_{n+1}A_{n+1}^{n-1}}}\\
&= \sqrt[2n]{x_1\cdots x_{n+1}A_{n+1}^{n-1}}
\end{aligned}$$

也就是 $A_{n+1}^{2n}\geqslant G_{n+1}^{n+1}A_{n+1}^{n-1}$，整理即得 $A_{n+1}\geqslant G_{n+1}$，得证，等号成立的条件也是显而易见的。

在 $A_n\geqslant G_n$ 中，把每个数 x_i 都换成 $\frac{1}{x_i}$，立即便得到 $\frac{1}{H_n}\geqslant\frac{1}{G_n}$，所以 $H_n\leqslant G_n$，至于 $A_n\leqslant Q_n$，利用数学归纳法和基本不等式 $2ab\leqslant a^2+b^2$ 就可以轻松获证，这里就不啰嗦了。

下面再提供 $A_n\geqslant G_n$ 的另一证法.

这个不等式可以改写为如下形式

$$\frac{x_1}{G_n}+\frac{x_2}{G_n}+\cdots+\frac{x_n}{G_n}\geqslant n$$

因为左边各项之积为 1，所以我们只要能够证明：如果 n 个正实数 $x_i(i=1,2,\ldots,n)$ 满足 $x_1x_2\cdots x_n=1$，则 $x_1+x_2+\cdots+x_n\geqslant n$ 就可以了。

对于 $n=2$ 的情况是很容易验证的，假定对于 n 个数也成立，现在来看 $n+1$ 个数的情况，先把 x_nx_{n+1} 绑在一起视为一个数，这样 $n+1$ 个数就成了 n 个数，根据归纳假设可得 $x_1+\cdots+x_{n-1}+x_nx_{n+1}\geqslant n$，于是我们只要证明 $x_n+x_{n+1}\geqslant x_nx_{n+1}+1$ 就可以了，也就是要证明 $(x_n-1)(x_{n+1}-1)\leqslant 0$，这个似乎不一定成立，但是乘积为 1 的 $n+1$ 个正实数，必定至少有一个大于等于 1，同时至少有一个小于等于 1，我们可以重新排序这些正实数，使得这两个数放在最后（或者干脆一开始绑定两个数的时候就绑这两个数），于是不等式 $(x_n-1)(x_{n+1}-1)\leqslant 0$ 就成立了，所以得证。 □

4.1.1) 均值不等式中的几个项都有相似的构成形式: $f^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}f(x_i)\right)$,

对于 A_n，$f(x)=x$，对于 H_n，$f(x)=\frac{1}{x}$，对于 G_n，唔...$f(x)=\ln x$，对于 Q_n，$f(x)=x^2$，所以这些平均数的大小关系完全取决于对应函数的性质，从几何意义上看，比算术平均数小的平均数所对应函数都是向左凸起的，而比算术平均数大的平均数所对应函数则是向右凸起的。 ■

4.1.2) 匀变速直线运动中间时刻瞬时速度和中间位移瞬时速度的大小 我们来考察一下匀变速直线运动中，中间时刻的瞬时速度和经过中间位移处瞬时速度的大小关系。因为加速度是恒定的，速度是时间的一次函数，所以对于起始时刻 t_1 和终止时刻 t_2 来说，中间时刻的瞬时速度是起始速度 v_1 和终止速度 v_2 的算术平均 $\frac{1}{2}(v_1+v_2)$，根据速度公式 $v(t)=v_1+at$ 和位移公式 $s(t)=v_1t+\frac{1}{2}at^2$(其中 a 是加速度) 可得出经过中间位移处的瞬时速度是 $\sqrt{\frac{v_1^2+v_2^2}{2}}$，由均值不等式即知中间位移处的瞬时速度要大一些。 ■

对于均值不等式中最基本的 $G_n \leqslant A_n$，将每个数的权重一般化，得如下的加权形式

定理 4.1.2 (加权平均值不等式). *对于任意 n 个正实数 $x_i(i=1,2,\ldots,n)$ 和一组权值 $\alpha_i \geqslant 0(i=1,2,\ldots,n), \sum_{i=1}^n \alpha_i = 1$，有如下的不等式*

$$\sum_{i=1}^n \alpha_i x_i \geqslant \prod_{i=1}^n x_i^{\alpha_i} \tag{4.6}$$

等号也只有在 x_i 全部都相等时才成立.

证明. 先证 $\alpha_i(i=1,2,\ldots,n)$ 都是有理数的情况，此时存在非负整数 $p_i(i=1,2,\ldots,n)$ 使得 $\alpha_i=\frac{p_i}{\sum_{k=1}^n p_k}$，只要把 p_ix_i 看成 p_i 个 x_i 相加，就有

$$\begin{aligned}\sum_{i=1}^n \alpha_i x_i &= \sum_{i=1}^n \frac{p_i}{\sum_{k=1}^n p_k} x_i \\ &= \frac{1}{\sum_{i=1}^n p_i} \cdot \sum_{i=1}^n p_i x_i \\ &\geqslant \left(\prod_{i=1}^n x_i^{p_i}\right)^{\frac{1}{\sum_{i=1}^n p_i}} \\ &= \prod_{i=1}^n x_i^{\alpha_i}\end{aligned}$$

所以不等式得证。

对于某个 α_i 为无理数的情况，使用有理数序列去逼近它，再两边取极限即得证。 □

4.1.2 柯西 (Cauchy) 不等式

定理 4.1.3 (柯西不等式). 对于任何两组实数 $x_i(i=1,2,\ldots,n)$ 和 $y_i(i=1,2,\ldots,n)$，有下面不等式成立

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \tag{4.7}$$

其中等号成立的唯一场景是：存在一个共同的实数 λ 使得 $x_i=\lambda y_i(i=1,2,\ldots,n)$ 都成立。

等号成立的条件，通俗的说就是两组实数成比例，如果我们约定分数中的分母可以为零，并且此时分子也必须为零，则可以把取等条件改写为

$$\frac{x_1}{y_1}=\frac{x_2}{y_2}=\cdots=\frac{x_n}{y_n}$$

如果记向量 $\vec{a}=(x_1,x_2,\ldots,x_n)$, $\vec{b}=(y_1,y_2,\ldots,y_n)$，则柯西不等式即表明 $|\vec{a}\cdot\vec{b}|\leqslant|\vec{a}|\cdot|\vec{b}|$.

证明很简单:

$$\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} x_i^2\right)-\left(\sum_{i=1}^{n} x_i y_i\right)^2=\sum_{1\leqslant i<j\leqslant n}(x_i y_j-x_j y_i)^2\geqslant 0$$

但更为流行的是下面的证明:

证明. 构造二次函数

$$f(t)=t^2\sum_{i=1}^{n} x_i^2+2t\sum_{i=1}^{n} x_i y_i+\sum_{i=1}^{n} y_i=\sum_{i=1}^{n}(x_i t+y_i)^2\geqslant 0$$

既然是恒为非负的二次函数，其判别式必小于等于零，于是即得柯西不等式，等号成立的条件就是二次函数有零点 t_0，使得 $x_i t_0+y_i=0(i=1,2,\ldots,n)$，也就是两组数成比例。 □

4.1.3 排序不等式

定理 4.1.4 (排序不等式). 对于任何两组按照相同的大小顺序排列好（均从小到大或者均从大到小）的两组实数 $x_i(i=1,2,\ldots,n)$ 和 $y_i(i=1,2,\ldots,n)$，有下面的不等式成立

$$\begin{aligned}
&\quad x_1 y_n+x_2 y_{n-1}+\cdots+x_n y_1\\
&\leqslant\ x_1 y_{r_1}+x_2 y_{r_2}+\cdots+x_n y_{r_n}\\
&\leqslant\ x_1 y_1+x_2 y_2+\cdots+x_n y_n
\end{aligned} \tag{4.8}$$

其中 $r_i(i=1,2,\ldots,n)$ 是 $1,2,\ldots,n$ 的任意一个排列。

不等式的通俗说法就是，反序和最小，同序和最大，乱序和居中。

证明. 对于中间的乱序和，如果 $r_1 \neq 1$，那么 y_1 就必定跟某个 x_m 搭配着，也就是 $x_1 y_{r_1}$ 和 $x_m y_1$ 这两项，我们尝试交换一下这两个搭配，看看和会怎么变化

$$(x_1 y_1 + x_m y_{r_1}) - (x_1 y_{r_1} + x_m y_1) = (x_1 - x_m)(y_1 - y_{r_1}) \geqslant 0$$

因此这样做的结果是使得和式变大了，照此逐步调整下去，所得和式将会继续变大，直到所有 $r_i = i(i = 1, 2, \ldots, n)$ 都成立时得到同序和为最大，这就证明了不等式的左边，这个方法称为 *逐步调整法*。

至于不等式的右边，只要把第二组数 $y_1, y_2, \ldots, y_n$ 换为 $-y_n, -y_{n-1}, \ldots, -y_1$，则它与 x_i 单调性相同，把左边不等式应用到这两组数上即可得到右边不等式。 □

排序不等式刻画的是在乘法运算中所蕴含的规律，本书倾向于认为在均值不等式、柯西不等式、排序不等式中，排序不等式是最基本的不等式，因为从1.4节利用多项式乘幂定理证明均值不等式的过程中可以看出，均值不等式赖以成立的基础是不等式 $x^n + y^n \geqslant x^{n-1}y + xy^{n-1}(n \in N_+)$，而柯西不等式则是得益于不等式 $x^2 + y^2 \geqslant 2xy$，而这两个简单不等式都是排序不等式的直接推论。

4.1.4 切比雪夫 (Chebyshev) 不等式

定理 4.1.5 (切比雪夫不等式). *对于任意两组按照相同的大小顺序排列好的实数 $x_i(i = 1, 2, \ldots, n)$ 和 $y_i(i = 1, 2, \ldots, n)$，成立着下面不等式*

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_{n+1-i} \leqslant \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) \leqslant \frac{1}{n}\sum_{i=1}^{n} x_i y_i \tag{4.9}$$

其中的等号成立的唯一场景是：两组数中至少有一组数的值都相同。

切比雪夫不等式是排序不等式的直接推论，这从如下的证明过程即可看出。

证明. 先证明右边不等式，由排序不等式，可得如下的 n 个等式或不等式:

$$\begin{aligned}
x_1 y_1 + x_2 y_2 + \cdots + x_n y_n &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\
x_1 y_2 + x_2 y_3 + \cdots + x_n y_1 &\leqslant x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\
x_1 y_3 + x_2 y_4 + \cdots + x_n y_2 &\leqslant x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\
&\cdots \\
x_1 y_n + x_2 y_1 + \cdots + x_n y_{n-1} &\leqslant x_1 y_1 + x_2 y_2 + \cdots + x_n y_n
\end{aligned}$$

将这些等式或不等式全部相加，即得切比雪夫不等式的右边。

至于不等式的左边，只要在把序列 $y_1, y_2, \ldots, y_n$ 替换为序列 $-y_n, -y_{n-1}, \ldots, -y_1$ 后应用不等式的右边即可得出 (当然也可以使用排序不等式的左边也像刚才那样累加得出)。 □

4.1.5 琴生不等式

琴生不等式是一个与函数凸性有关的不等式。

定理 4.1.6 (琴生不等式). *设函数 $f(x)$ 在区间 D 上是上凸函数，则对于该区间上任意 n 个实数 $x_i(i=1,2,\ldots,n)$，有如下不等式成立:*

$$f\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right) \geqslant \frac{1}{n}\sum_{i=1}^{n}f(x_i) \tag{4.10}$$

同样有加权形式的琴生不等式

定理 4.1.7 (加权形式的琴生不等式). *设函数 $f(x)$ 在区间 D 上是上凸函数，则对于该区间上任意 n 个实数 $x_i(i=1,2,\ldots,n)$ 和一组权值 $\alpha_i \geqslant 0(i=1,2,\ldots,n), \sum_{i=1}^{n}\alpha_i = 1$，有如下不等式成立:*

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) \geqslant \sum_{i=1}^{n}\alpha_i f(x_i) \tag{4.11}$$

我们在2.3节已经证明过琴生不等式，所以这里就省去了。

4.1.6 幂平均值不等式

定理 4.1.8 (幂平均值不等式). *对于任意 n 个正实数 $x_i(i=1,2,\ldots,n)$ 和两个正实数 $0<p<q$，成立着不等式*

$$\left(\frac{1}{n}\sum_{i=1}^{n}x_i^p\right)^{\frac{1}{p}} \leqslant \left(\frac{1}{n}\sum_{i=1}^{n}x_i^q\right)^{\frac{1}{q}} \tag{4.12}$$

等号当且仅当 $x_i(i=1,2,\ldots,n)$ 全部相等时成立。

取 $p=1, q=2$，便可得算术平均数小于等于平方平均数。

将每个数的权重 $\frac{1}{n}$ 一般化，得如下的加权形式

定理 4.1.9 (加权幂平均值不等式). *对任意 n 个正实数 $x_i(i=1,2,\ldots,n)$ 和任意两个正实数 $0<p<q$, 以及一组权值 $\alpha_i \geqslant 0(i=1,2,\ldots,n), \sum_{i=1}^{n}\alpha_i = 1$，成立着下面不等式*

$$\left(\sum_{i=1}^{n}\alpha_i x_i^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n}\alpha_i x_i^q\right)^{\frac{1}{q}} \tag{4.13}$$

4.1.7 赫尔德 (Holder) 不等式

定理 4.1.10 (赫尔德不等式). 设 $x_i > 0, y_i > 0 (i = 1, 2, \ldots, n)$ 且 $\frac{1}{p} + \frac{1}{q} = 1$，那么在 $p > 1$ 时成立下面的不等式

$$\sum_{i=1}^{n} x_i y_i \leqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} x_i^q\right)^{\frac{1}{q}} \tag{4.14}$$

在 $p < 1, p \neq 0$ 不等式反向，等号当且仅当向量 $(x_1^p, \ldots, x_n^p)$ 和 $(y_1^q, \ldots, y_n^q)$ 共线时成立。

在赫尔德不等式中，取 $p = q = 2$，即得柯西不等式。

4.1.8 闵可夫斯基 (Minkowski) 不等式

定理 4.1.11 (闵可夫斯基不等式). 设 $x_i > 0, y_i > 0 (i = 1, 2, \ldots, n)$，那么在 $p > 1$ 时成立下面的不等式

$$\left(\sum_{i=1}^{n} (x_i + y_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}} \tag{4.15}$$

在 $p < 1$ 时不等式反向，等号当且仅当存在一个共同的实数 λ 使得 $x_i = \lambda y_i (i = 1, 2, \ldots, n)$ 时成立。

在闵可夫斯基不等式中取 $n = 2, p = 2$，便是三角形两边之和大于第三边的坐标表达。

4.1.9 嵌入不等式

定理 4.1.12 (嵌入不等式). 对 $\triangle ABC$ 和任意的实数 x, y, z 均有下面不等式成立

$$x^2 + y^2 + z^2 \geqslant 2yz \cos A + 2zx \cos B + 2xy \cos C \tag{4.16}$$

其中，当且仅当 $x : y : z = \sin A : \sin B : \sin C$ 时等号成立。

第五章　向量与复数

用定比分点公式解决线线相交的平面几何问题

5.1　用向量定比分点公式解决与交点相关的平面几何问题

本文是人教数学官方论坛 QQ 群里某网友发上来的题目，我一瞅，这不又是这类问题么，之前我用向量定比分点公式证明过梅涅劳斯定理和塞瓦定理，少不得再卖弄一番。当然了，其实葫芦里也没什么东西，不过是用了向量的共线、定比分点和内积方面的几个小知识点而已，借此也保证不超出高中教学大纲的范围。

向量的定比分点公式，其实更加有用的是另一种形式 $\vec{x}=t\vec{a}+(1-t)\vec{b}$，因为只要取 $t=1/2$ 便定位出中点，而分别取 0 和 1，则分别得出两个基向量，这是用分比的形式所得不出的。

这个方法就是，对于两条线段的交点，先在其中一条线段上使用定比分点公式，而公式中的这个实数就是一个待定的参数，怎么定出这个参数呢，再通过该点在另一线段上，就有相关向量的共线关系，于是得出关于这个参数的方程，以后的事情就没什么技术含量了。

题目：三角形 $\triangle ABC$ 中，AD 是 $\angle BAC$ 的平分线，点 D 在 BC 边上，DE、DF 分别与边 AB、AC 垂直，E 和 F 是垂足，BF、CE 相交于点 P，求证：$AP\perp BC$。

证明：设 $\overrightarrow{AB}=\vec{a}$, $\overrightarrow{AC}=\vec{b}$，并记 $|\vec{a}|=a$，$|\vec{b}|=b$，则由于 $\frac{BD}{DC}=\frac{AB}{AC}$，根据向量定比分点公式得

$$\overrightarrow{AD}=\frac{b}{a+b}\vec{a}+\frac{a}{a+b}\vec{b} \tag{5.1}$$

再设 $\overrightarrow{AE}=\lambda_1\vec{a}$，则由 $\overrightarrow{DE}\cdot\vec{a}=0$ 得

$$(b-\lambda_1(a+b))a+(\vec{a}\cdot\vec{b})=0 \tag{5.2}$$

或者

$$(1-\lambda_1)ab-\lambda_1 a^2+m=0 \tag{5.3}$$

同理，设 $\overrightarrow{AF}=\lambda_2\overrightarrow{b}$，则有

$$(a-\lambda_2(a+b))b+(\overrightarrow{a}\cdot\overrightarrow{b})=0 \tag{5.4}$$

或者

$$(1-\lambda_2)ab-\lambda_2 b^2+m=0 \tag{5.5}$$

(题外话：据此便有 $a\lambda_1=b\lambda_2$，这与 $AE=AF$ 相吻合)

为确定点 P 的位置，设 $\overrightarrow{BP}=t\overrightarrow{PF}$，则

$$\begin{aligned}\overrightarrow{AP} &= \frac{1}{1+t}\overrightarrow{AB}+\frac{t}{1+t}\overrightarrow{AF} &(5.6)\\ &= \frac{1}{1+t}\overrightarrow{a}+\frac{t}{1+t}\lambda_2\overrightarrow{b} &(5.7)\end{aligned}$$

为便于计算，记 $r=\frac{t}{1+t}$，则 $\overrightarrow{AP}=(1-r)\overrightarrow{a}+r\lambda_2\overrightarrow{b}$

接着便要确定 r 的值，由于点 P 也在 CE 上，所以 $\overrightarrow{CP}$ 与 $\overrightarrow{CE}$ 共线，而 $\overrightarrow{CE}=\overrightarrow{AE}-\overrightarrow{AC}=\lambda_1\overrightarrow{a}-\overrightarrow{b}$，$\overrightarrow{CP}=\overrightarrow{AP}-\overrightarrow{AC}=(1-r)\overrightarrow{a}+(r\lambda_2-1)\overrightarrow{b}$

设 $\overrightarrow{CP}=s\overrightarrow{CE}$，消去 s 得到

$$1-r=\lambda_1(1-r\lambda_2) \tag{5.8}$$

解之得

$$r=\frac{1-\lambda_1}{1-\lambda_1\lambda_2} \tag{5.9}$$

同样为了方便，记 $m=(\overrightarrow{a}\cdot\overrightarrow{b})$，则

$$\begin{aligned}\overrightarrow{CB}\cdot\overrightarrow{AP} &= (\overrightarrow{a}-\overrightarrow{b})\cdot((1-r)\overrightarrow{a}+r\lambda_2\overrightarrow{b})\\ &= (1-r)a^2-r\lambda_2 b^2+(r\lambda_2-(1-r))m\end{aligned}$$

由5.5有 $\lambda_2 b^2=(1-\lambda_2)ab+m$，代入上式得

$$\begin{aligned}\overrightarrow{CB}\cdot\overrightarrow{AP} &= (1-r)a^2-r(1-\lambda_2)ab-rm+(r\lambda_2-1+r))m\\ &= (1-r)a^2-r(1-\lambda_2)ab+(r\lambda_2-1)m\end{aligned}$$

又由式5.8得 $r\lambda_2-1=-\frac{1-r}{\lambda_1}$，并将5.9代入上式得

$$\begin{aligned}\overrightarrow{CB}\cdot\overrightarrow{AP} &= (1-r)\left(a^2-\frac{m}{\lambda_1}\right)-\frac{(1-\lambda_1)(1-\lambda_2)}{1-\lambda_1\lambda_2}ab\\ &= \frac{\lambda_1(1-\lambda_2)}{1-\lambda_1\lambda_2}\left(a^2-\frac{m}{\lambda_1}\right)-\frac{(1-\lambda_1)(1-\lambda_2)}{1-\lambda_1\lambda_2}ab\\ &= \frac{1-\lambda_2}{1-\lambda_1\lambda_2}(\lambda_1 a^2-m-(1-\lambda_1)ab)\\ &= \frac{1-\lambda_2}{1-\lambda_1\lambda_2}a(\lambda_1 a(a+b)-m-ab)\end{aligned}$$

再由式5.2得 $\lambda_1 a(a+b) = ab + m$，代入上式立刻得到 $\overrightarrow{CB} \cdot \overrightarrow{AP} = 0$，证毕。

第六章　组合

方程 $\sum_{i=1}^n x_i = m$ 的正整数解与非负整数解的个数错位排列问题的数列解法

6.1　方程 $\sum_{i=1}^n x_i = m$ 的解的个数

排列组合中有很多问题都可以归结为以下问题：方程 $\sum_{i=1}^n x_i = m$ 的正整数解的个数和非负整数解的个数，这里 n 和 m 都是正整数。

先看正整数解，只要想象一下，把 m 个圆圈画在一横排，中间就会形成 $m-1$ 个空隙，只要在这些空隙中选取 $n-1$ 个插入分隔符号，则这些圆圈被划分成了 n 个组，每一组中的圆圈个数就能对应是 x_i 的值，因此整数解的个数是 C_{m-1}^{n-1}。

而对于非负整数解，只要令 $y_i = x_i + 1$，则问题转化为求方程 $\sum_{i=1}^n y_i = m+n$ 的正整数解的个数，借用前述结论，这个结果应是 C_{n+m-1}^{n-1}。

6.2　用数列方法解决错位排列问题

此文要解决的问题是：有编号为 $1, 2, \ldots, n$ 的 n 个人和同样编号的 n 个座位，如果每个人都不允许坐与自身编号相同的座位，有多少种坐法？此即所谓的错位排列问题。

这个用容斥原理是可以瞬间解决的，高中时期年少轻狂，愣是用定义数列求通项的方法给求出来了，此文即是重新写的，原稿已经遗失。

考虑编号为 1 的人，假如他坐在编号为 r_1 的座位上，又看编号为 r_1 的人，他可以坐在 1 号座位上，也可以坐在 r_2 号座位上，如果是坐在 1 号座位上，那么 1 号和 r_1 号这两个人由于交叉坐，与其他人无涉，如果是坐在 r_2 号座位上，则再继续看 r_2 号人，依此顺着链条 $1 \to r_1 \to r_2 \to \cdots$，这样下去的结果是，从 1 号人开始，能够找到一个环，使得该环内的人和座位与环外的人和座位无涉，可以独立开来，如果找不到这样的环，那无非是所

有人一起构成了一个整体环，即是无法分割。这样的环可能不止一个，但我们仅考虑包含 1 号人的那个环。

假定 n 个人的错位排列数为 a_n，设包含 1 号人的环共含有 $m(2 \leqslant m \leqslant n-2)$ 个人，那么只要从除 1 号之外的人中选出 $m-1$ 个人排在这个座位链上，因此座法是 $(m-1)!C_{n-1}^{m-1} = \frac{(n-1)!}{(n-m)!}$，而剩下的 $n-m$ 个人又组成了一个较小的错位排列，因此其坐法数是 a_{n-m}，此外，当 $m=n$ 时，所有人构成一个环，只要把他们按照座位链排成一排即可，所以此时的坐法是 $(n-1)!$，于是总共的坐法是：

$$a_n = \sum_{m=2}^{n-2} \frac{(n-1)!}{(n-m)!} a_{n-m} + (n-1)! = (n-1)!(\sum_{m=2}^{n-2} \frac{a_{n-m}}{(n-m)!} + 1)$$

将式中的求和顺序倒过来，就有

$$a_n = (n-1)!(\sum_{m=2}^{n-2} \frac{a_m}{m!} + 1) \tag{6.1}$$

此即其作为数列的递推公式，自然的，$a_1 = 0, a_2 = 1$。

下面来求它的通项公式，上式可以变形为：

$$\frac{a_n}{n!} = \frac{1}{n}(1 + \sum_{m=2}^{n-2} \frac{a_m}{m!}) \tag{6.2}$$

只要记 $b_n = \frac{a_n}{n!}$，则有 $b_1 = 0$ 和

$$b_n = \frac{1}{n}(1 + \sum_{m=2}^{n-2} b_m) \tag{6.3}$$

于是有

$$nb_n - (n-1)b_{n-1} = b_{n-2} \tag{6.4}$$

也就是

$$b_n - b_{n-1} = -\frac{1}{n}(b_{n-1} - b_{n-2}) \tag{6.5}$$

而 $b_2 - b_1 = \frac{1}{2}$，所以

$$b_n - b_{n-1} = (-1)^{n-2}\frac{1}{n!} = (-1)^n \frac{1}{n!} \tag{6.6}$$

于是

$$b_n = b_1 + \sum_{m=2}^{n} (b_m - b_{m-1}) = \sum_{m=2}^{n} (-1)^m \frac{1}{m!} \tag{6.7}$$

所以最终

$$a_n = n!b_n = n! \sum_{m=2}^{n} (-1)^m \frac{1}{m!} \tag{6.8}$$

在 $0! = 1$ 的规定下，也可以把求和指标从零开始

$$a_n = n!\sum_{m=0}^{n}(-1)^m\frac{1}{m!} \tag{6.9}$$

这就是最终的结果。

6.3 题选

题目 6.3.1. 四个人玩传球游戏，每个人接到球后将球传给别人，首先由甲发球，在经过 n 次传球后球又回到甲手中，求这样的传球方法数。

由于结果与自然数 n 有关，将这结果作为一个数列是合理的。

记结果数为 a_n，那么在 n 次传球过程中，传球方法总数是 3^n。考虑数列的递推情况，对于 $n+1$ 次传球而言，只要前 n 次传球后球不在甲手中，则最后一次传球只有一种固定的方式，所以有 $a_{n+1} = 3^n - a_n$，而 $a_1 = 0$。

有了递推公式，现在来求通项，在递推式两边同时乘以 $(-1)^n$ 得

$$(-1)^n a_n - (-1)^{n+1}a_{n+1} = (-3)^n$$

于是令 $b_n = (-1)^n a_n$ 则有 $b_n - b_{n+1} = (-3)^n$，所以

$$b_n = b_1 - \sum_{k=1}^{n-1}(b_k - b_{k+1} = -\sum_{k=1}^{n-1}(-3)^k = \frac{3}{4}(1-(-3)^{n-1})$$

所以 $a_n = \frac{(3^n+3(-1)^n)}{4}$.

第七章 平面几何

7.1 题集

题目 7.1.1. 如图，在三角形 ABC 中，AO 是 BC 边上的高，D 是 AO 的中点，点 E 是其内切圆 $\odot I$ 与 BC 边的切点，连接 ED 并延长与内切圆相交于另一点 F，过点 C 并与 BC 垂直的直线与 BI 的延长线相交于 G，求证:(1) $EF \perp FG$,(2)EF 平分角 BFC。

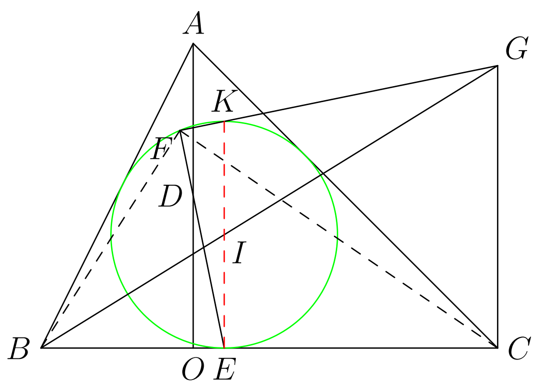

图 7.1

题目 7.1.2. 如图7.2，分别以三角形 ABC 的边 AB 和 AC 向形外作正方形，两个正方形的中心分别记为 O_1 和 O_2，X 是 BC 边中点，其余各点的意义如图所示，求证:

1. 三角形 O_1XO_2 是等腰直角三角形.
2. $ST \parallel PQ$.

证明. 易见 $\triangle ACE \cong \triangle AGB$，而 $O_1X \parallel EC$ 并且 $|O_1X| = \frac{1}{2}|EC|$，同样 $O_2X \parallel BG$ 并且 $|O_2X| = \frac{1}{2}|BG|$，但是 $EC = BG$，并且 $EC \perp BG$，所以 $O_1X = O_2X$ 并且 $O_1X \perp O_2X$，所以 $\triangle O_1XO_2$ 是等腰直角三角形.

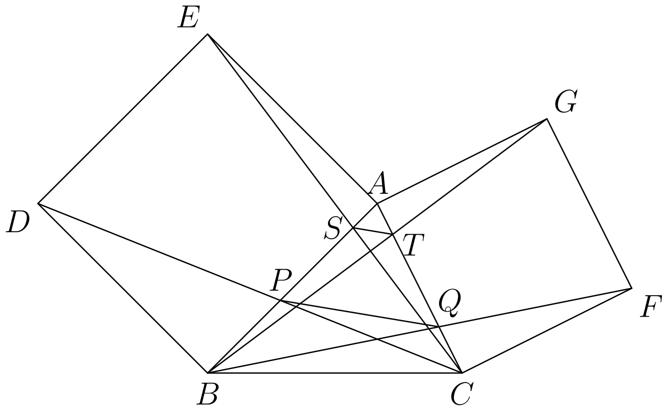

图 7.2

第二问的最简单证法是设 DE 延长线与 CA 延长线交于 U，FG 延长线与 BA 延长线交于 V，则有 $\triangle AEU \sim \triangle AGV$，于是

$$\frac{AS}{SP} = \frac{UE}{ED} = \frac{UE}{EA} = \frac{VG}{GA} = \frac{VG}{GF} = \frac{AT}{TQ}$$

所以 $ST \parallel PQ$.

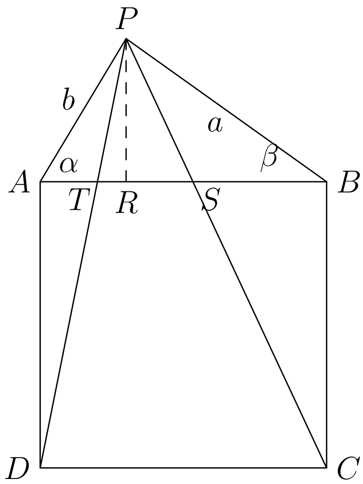

图 7.3

如果注意到图中包含着两个同样的基本图形，则可以有如下的证明，先提一个引理：如图7.3，设点 P 为正方形 $ABCD$ 的边 AB 外侧一点，PC 与 PD 分别与边 AB 相交于 S 和 T，记 $\angle PAB = \alpha, \angle PBA = \beta$，则有

$$AT : TS : SB = \cos\alpha\sin\beta : \sin\alpha\sin\beta : \sin\alpha\cos\beta$$

利用此引理知，$AT : TS$ 仅与角 α 有关，因此在原题中立刻便有 $\frac{AS}{SP} = \frac{AT}{TQ}$，

因此两线平行，而这个引理的证明也很简单，记点 P 在边 AB 上的投影为 R，并假定正方形边长为 1，并记 $PB=a, PA=b$，根据正弦定理可得

$$\frac{a}{\sin\alpha}=\frac{1}{\sin\alpha+\beta}=\frac{b}{\sin\beta}$$

于是

$$a=\frac{\sin\alpha}{\sin\alpha+\beta}, b=\frac{\sin\beta}{\sin\alpha+\beta}$$

根据几何关系有

$$AR=b\cos\alpha=\frac{\cos\alpha\sin\beta}{\sin\alpha+\beta}, PR=b\sin\alpha=\frac{\sin\beta}{\sin\alpha+\beta}$$

而由 $\frac{TR}{AT}=\frac{PR}{AD}$ 得

$$AT=\frac{AT}{AT+TR}AR=\frac{1}{1+\frac{TR}{AT}}AR=\frac{1}{1+\frac{\sin\alpha\sin\beta}{\sin\alpha+\beta}}\cdot\frac{\cos\alpha\sin\beta}{\sin\alpha+\beta}=\frac{\cos\alpha\sin\beta}{\sin\alpha\sin beta+\sin\alpha+\beta}$$

根据对称性，互换 α 和 β 可得

$$BS=\frac{\sin\alpha\cos\beta}{\sin\alpha\sin\beta+\sin\alpha+\beta}$$

于是

$$TS=\frac{\sin\alpha\sin\beta}{\sin\alpha\sin\beta+\sin\alpha+\beta}$$

于是引理得证。 $\square$

第八章　解析几何

8.1　椭圆初步

椭圆曲线形状优雅，方程优美，性质优酷[1]...... 拜托我不能再吹下去了，不然全世界的汽车车轮都该改成椭圆形了，想象一下坐在这样的车上是什么感受......

本节所谓的标准椭圆是指方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的椭圆。

8.1.1　焦半径公式

椭圆上任一点与焦点的连线称为焦半径,在标准椭圆上任取一点 $P(x_0,y_0)$，那么它到左准线的距离 $d=x_0+\frac{a^2}{c}$，根据第二定义，它到左焦点 F_1 的距离是 $PF_1=ed=a+ex_0$，再根据第一定义，它到右焦点的距离是 $PF_2=2a-PF_1=a-ex_0$，于是得坐标形式的 焦半径公式:

$$PF_1=a+ex_0,\ PF_2=a-ex_0 \tag{8.1}$$

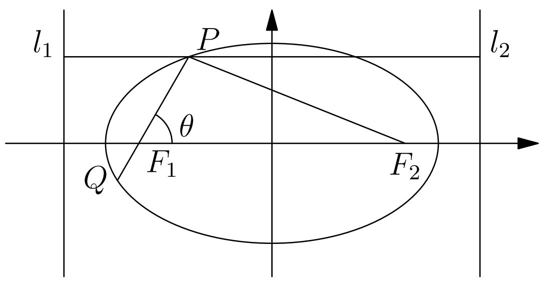

图 8.1椭圆的焦半径

假定向量 $\overrightarrow{F_1P}$ 与 x 轴正方向单位向量成角 θ，记其模 $r=|\overrightarrow{F_1P}|$ 在 $\triangle PF_1F_2$ 中使用余弦定理可得 $(2a-r)^2=r^2+4c^2-4cr\cos\theta$，解得

$$r=\frac{b^2}{a-c\cos\theta} \tag{8.2}$$

[1] 又美又酷。

这是倾斜角形式的 焦半径公式。如果 PF_1 延长线与椭圆相交于另一点 Q，易知向量 $\overrightarrow{F_1Q}$ 与 x 轴正方向单位向量成角 $\theta+\pi$，因此在上式中将 θ 换成 $\theta+\pi$，即得可知

$$F_1P=\frac{b^2}{a-c\cos\theta},\ F_1Q=\frac{b^2}{a+c\cos\theta} \tag{8.3}$$

从上式立刻得出，如果过焦点 F 的直线与椭圆相交于 P、Q 两点，那么有

$$\frac{1}{FP}+\frac{1}{FQ}=\frac{2a}{b^2} \tag{8.4}$$

既然倒数和为定值，那么根据均值不等式便有

$$PQ=PF+QF\geqslant\frac{4}{\frac{1}{FP}+\frac{1}{FQ}}=\frac{2b^2}{a}$$

椭圆上任意两点间的连线称为弦，过焦点的弦称为焦点弦，那么上式就表明，焦点的长度有最小值，而根据焦半径公式，这个最小值只有在焦点弦与长轴垂直时取得。

实际上根据倾斜角形式的焦半径式，我们可以得出焦点弦的准确范围:

$$PQ=PF+QF=\frac{b^2}{a-c\cos\theta}+\frac{b^2}{a+c\cos\theta}=\frac{2ab^2}{a^2-c^2\cos^2\theta}$$

而 $0\leqslant\cos^2\theta\leqslant1$，所以焦点弦的范围是 $[\frac{2b^2}{a},2a]$ 最大值是长轴长，这一点从直观上是明显的。

8.1.2 第一定义与第二定义的等价性

椭圆两个定义的等价性其实从根据各自的定义得出相同的椭圆方程就可以看出，但这里感兴趣的是纯几何角度的证明，这只要证明下面这个命题就可以了：

命题 8.1.1. 设 a 和 c 是两个正实数，且 $a>c$，F_1 和 F_2 是平面上相距为 $2c(c>0)$ 的两个定点，O 为它们的中点，L 为射线 OF_1 上满足 $OL=\frac{a^2}{c}$ 的点 (由于 $\frac{a^2}{c}>c$，因此 L 比 F_1 距离 O 更远一些)，直线 l 过点 L 且与直线 F_1F_2 垂直，则平面上的动点 P 到 F_1、F_2 的距离之和为 $2a$ 的充分必要条件是 $\frac{PF_1}{d}=\frac{c}{a}$，其中 d 是动点 P 到直线 l 的距离。

证明. 先证必要性。设直线 PF_1 的倾斜角为 θ，记 $PF_1=r$，则 $PF_2=2a-r$，在三角形 PF_1F_2 中由余弦定理得

$$(2a-r)^2=r^2+(2c)^2-2r\cdot2c\cdot\cos\theta$$

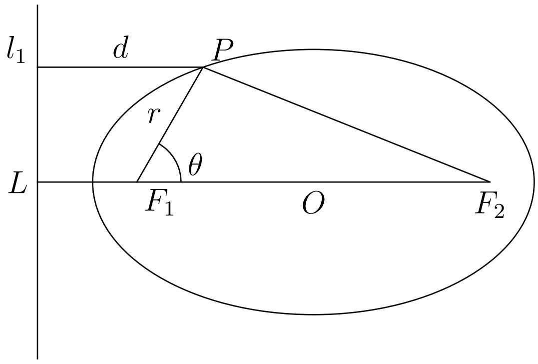

图 8.2椭圆的第一定义与第二定义

记 $b^2 = a^2 - c^2$，则由上式解得

$$r = \frac{b^2}{a - c\cos\theta} \tag{8.5}$$

另外，根据几何关系显然有

$$d = LF_1 + PF_1\cos\theta = \frac{a^2}{c} - c + r\cos\theta = \frac{b^2}{c} + r\cos\theta \tag{8.6}$$

因此

$$\frac{d}{r} = \frac{1}{r}\frac{b^2}{c} + \cos\theta$$

将式8.5代入上式右端，立即得到 $\frac{d}{r} = \frac{a}{c}$，必要性得证。

下证充分性，倾斜角及 r 的定义同前，此时式8.6仍然成立，因此

$$\frac{a}{c} = \frac{d}{r} = \frac{b^2}{c} \cdot \frac{1}{r} + \cos\theta$$

解出 r 即得到8.5，由此之前的余弦定理式也是成立的，因此必然 $PF_2 = 2a - r$，充分性得证。 □

8.1.3 切线与光学性质

椭圆曲线将其所在平面分成了三个部分：椭圆内部、椭圆曲线上、椭圆外部，对于椭圆曲线上的点，正如大家所熟知的，它们的坐标都满足椭圆的方程：

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{8.7}$$

而对于椭圆内部的点，它们的坐标都满足不等式：

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1 \tag{8.8}$$

这个很好理解，对于椭圆内部任何一点，总能在椭圆上找到一个对应点，使其横坐标相同而纵坐标的绝对值大于椭圆内的点，而椭圆曲线上的这个对应点的坐标满足方程8.7，那么椭圆内的点的坐标满足不等式8.8就是显然的事情了。至于椭圆外部的点，那就只有成立不等式：

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}>1 \tag{8.9}$$

因为之前椭圆曲线与其方程之间以及椭圆内部的平面区域与其不等式之间都是可以互推的。

接下来讨论椭圆的切线问题，所谓切线，仿照圆的定义，按照直线与椭圆的交点数目，把直线与椭圆的位置关系划分为三类：相切、相交、相离。

在椭圆上任取一个点 $P(x_0,y_0)$，我们在例1.3.5中利用伸缩变换曾经得出过如下的切线方程:

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1 \tag{8.10}$$

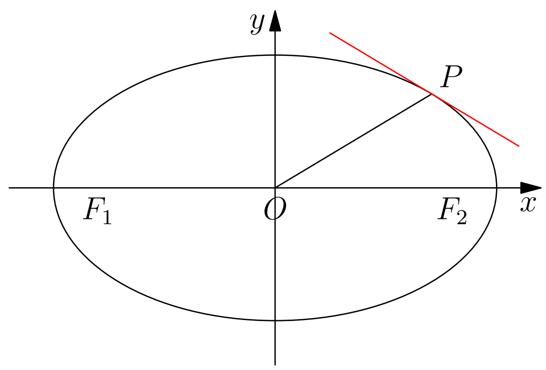

图 8.3椭圆的切线

在这里，我们换一个角度来阐明它就是切线，在这直线上任取一点 $T(x_T,y_T)$，有：

$$(\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2})+(\frac{x_T^2}{a^2}+\frac{y_T^2}{b^2})\geqslant 2(\frac{x_0x_T}{a^2}+\frac{y_0y_T}{b^2})=2 \tag{8.11}$$

所以得到:

$$\frac{x_T^2}{a^2}+\frac{y_T^2}{b^2}\geqslant 1 \tag{8.12}$$

这表明直线8.10上除点 P 外任何一点都在椭圆外，与椭圆只有 P 一个交点，所以它理所当然就是点 P 处的切线方程。

椭圆具有一个漂亮的光学性质：从椭圆的焦点之一发出的光线，经椭圆反射之后，必经过另一个焦点。需要说明的是光线经椭圆反射，等同于经椭圆上相应点处的切线反射。

为着证明这一点，提出如下定理:

定理 8.1.1 (椭圆切线定理)**.** 椭圆上某点处的切线，是该点与椭圆两个焦点连线所形成的焦点三角形在该点处的外角平分线。

证明. 分两步证明，先证明定理中所说的外角平分线是椭圆的切线，再证明切线的唯一性。

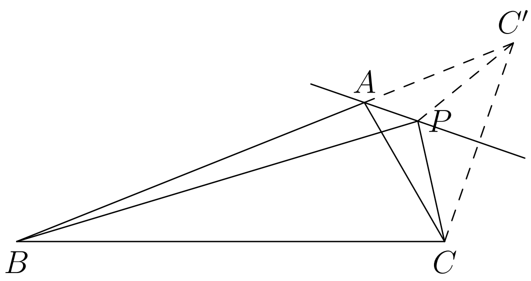

图 8.4三角形外角平分线的一个性质

要证明此外角平分线与椭圆相切，只需证明它与椭圆只有唯一公共点即可。这实际上来自于平面几何中的一个结论：如图8.4，点 P 是三角形 $\triangle ABC$ 在顶点 A 处的外角平分线上异于点 A 的任意一点，那么有不等式 $PB+PC>AB+AC$ 成立。

证明也很简单，作出点 C 关于此外角平分线的对称点 C'，那么有 $PB+PC=PB+PC'>BC'=AB+AC'=AB+AC$.

这表明点 P 不可能在椭圆上，所以此外角平分线与椭圆除点 A 外无其它公共点，即为切线。

再继续证明切线的唯一性，这只要证明在上述三角形 ABC 中，过点 A 的其它直线上，必存在另一个异于点 A 的点 Q，使得 $QB+QC=AB+AC$ 即可，因为这样点 Q 就是该直线与椭圆的另一个交点，从而直线与椭圆是相交而非相切。

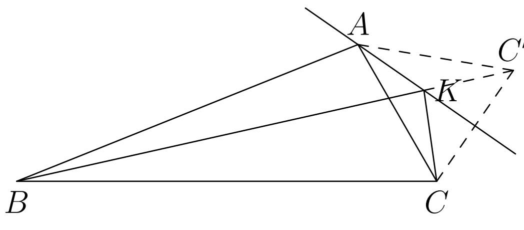

图 8.5椭圆切线唯一性

如图8.5，同样作点 C 关于这直线的对称点 C'，记 BC' 与直线交点为 K，则 $KB+KC=KB+KC'=BC'<AB+AC'=AB+AC$，这表明

点 K 到 B 和 C 两点的距离之和小于 $AB+AC$，显然直线上无穷远处的点到 B 和 C 的距离之和大于 $AB+AC$，所以直线上必存在一点 Q，使得 $QB+QC=AB+AC$，于是得证。 □

由上述定理立即可得椭圆的光学性质，这里我们再利用刚刚到得的切线方程来证明这一事实，假定从焦点发出的光线到达椭圆上一点 $P(x_0,y_0)$，而该点处的切线便是前文所求者，于是得出法线[2]的方程为

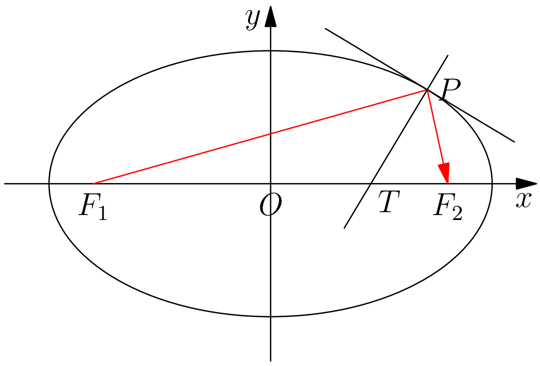

图 8.6椭圆的光学性质

$$y-y_0=\frac{a^2y_0}{b^2x_0}(x-x_0)$$

容易求得法线与 x 轴交点是 $T(e^2x_0,0)$，于是

$$\frac{TF_1}{TF_2}=\frac{c+e^2x_0}{c-e^2x_0}=\frac{a+ex_0}{a-ex_0}=\frac{PF_1}{PF_2}$$

这意味着法线平分了反射点处的两条焦半径，所以由一个焦点处发出的光线经切线反射后必然经过另一焦点。

8.1.4 极点与极线

现在在椭圆外任选一个点 $P(x_0,y_0)$，并由它向椭圆引两条切线，切点分别为 $T_1(x_1,y_1)$ 和 $T_2(x_2,y_2)$，由于点 P 同时在两条切线上，所以同时成立着:

$$\frac{x_0x_1}{a^2}+\frac{y_0y_1}{b^2}=1,\frac{x_0x_2}{a^2}+\frac{y_0y_2}{b^2}=1 \tag{8.13}$$

由此立刻得知切点弦 T_1T_2 的方程是:

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1 \tag{8.14}$$

[2]法线是经过反射点并与反射面垂直的直线。

这一点令人惊讶，要注意这时点 P 是在椭圆外，所以不要与刚才所证的切线方程混淆了。同样的方程，因为点 P 位置的不同，它代表的直线也不同。相信读者此时也会不失时机的提出如下问题：如果点 P 是在椭圆内部，那么方程8.14又代表什么样的直线呢？此刻还不好回答，但后文自然而然的会得出结论。

接着讨论弦的中点的问题，在椭圆上任取两点 $T_1(x_1, y_1)$ 和 $T_2(x_2, y_2)$，线段 T_1T_2 中点为 $Q(x_Q, y_Q)$，那么首先成立着等式：

$$\frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} = 1, \frac{x_2^2}{a^2} + \frac{y_2^2}{b^2} = 1 \tag{8.15}$$

两式相减得到:

$$\frac{x_1^2 - x_2^2}{a^2} + \frac{y_1^2 - y_2^2}{b^2} = 0 \tag{8.16}$$

进一步变形为:

$$\frac{y_1 - y_2}{x_1 - x_2} \frac{\frac{y_1+y_2}{2}}{\frac{x_1+x_2}{2}} = -\frac{b^2}{a^2} \tag{8.17}$$

所以弦 T_1T_2 及中点与椭圆中心连线 OQ 两者斜率之积为定值[3]:

$$k_{T_1T_2} k_{OQ} = -\frac{b^2}{a^2} \tag{8.18}$$

据此可以知道若另作一弦使其所在直线平行于 OQ，那么其中点与椭圆中心的连线必平行于直线 T_1T_2，即两个方向存在着相互性，这称之为椭圆的一组“共轭方向”，而这个结论则可以视为椭圆中的垂径定理。此外还可以知道，对于椭圆内部任何一个点，在经过它的所有弦中，只有唯一一条能使它成为弦的中点，而该弦的方向，就是该点与椭圆中心连线的共轭方向。而如果在这里平行移动弦 T_1T_2，直到它成为椭圆的切线 l，假如切点为 $P(x_0, y_0)$，那么我们应该得出：

$$k_l k_{OP} = -\frac{b^2}{a^2} \tag{8.19}$$

这与我们刚才所推导的点 P 处的切线方程正好吻合。

回到刚才过椭圆外一点 $P(x_0, y_0)$ 引两条切线 PT_1 及 PT_2 的场景中，以 Q 标记切点弦 T_1T_2 的中点，则点 Q 的坐标必然满足切点弦 T_1T_2 的方程：

$$\frac{x_0 x_Q}{a^2} + \frac{y_0 y_Q}{b^2} = 1 \tag{8.20}$$

这一点就很有意思了，因为这意味着点 P 位于下面的直线上：

$$\frac{x_Q x}{a^2} + \frac{y_Q y}{b^2} = 1 \tag{8.21}$$

[3]本文中都不考虑斜率不存在等特殊情况，只就一般性的情况进行叙述，以得出一般性的结论。

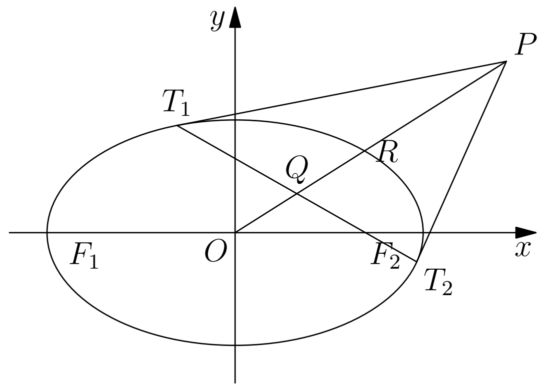

图 8.7椭圆的极点与极线

这是一条什么样的直线呢？容易发现，它的方向就是 OQ 的共轭方向！现在回想一下前面所提的悬而未决的问题：对于椭圆内的一点，方程8.14代表什么样的直线，现在就可以理直气壮的回答说，先找出以椭圆内的这点为中点的弦，然后过此弦两端点引椭圆切线，两切线有一个交点，这条直线就经过了这个点，那么方向呢？将椭圆内的这点与椭圆中心相连，其共轭方向就是该直线的方向！这段长长的描述显得有些不符合数学美感，但至少我们解决了如下问题：针对椭圆上、椭圆内、椭圆外的点 $P(x_0, y_0)$，方程8.14各自分别代表什么样的直线，而且还可以知道，椭圆内 (除去中心) 与椭圆外的点之间，存在着一种对称关系，借此我们可以在这两大平面区域之间，建立一对一的映射关系。

进一步，直线 OQ 与 T_1T_2 是共轭的，而现在直线8.21的方向也与 OQ 共轭，这表明直线8.21与切点弦 T_1T_2 是平行的！比较一下方程8.21与方程8.14就可以得出结论：直线 PQ 是经过椭圆中心的！进而记 R 为线段 PQ 与椭圆的交点，那么显然：

$$\frac{x_R^2}{a^2} + \frac{y_R^2}{b^2} = 1 \tag{8.22}$$

比较一下方程8.20与8.22，再加之我们知道 P、Q、R 三点共线，而且还过椭圆中心，所以：

$$x_0 x_Q = x_R^2, y_0 y_Q = y_R^2 \tag{8.23}$$

由此得到：

$$OP \cdot OQ = OR^2 \tag{8.24}$$

一切都是如此的完美。

前面曾经考虑过，当点 $P(x_0, y_0)$ 在椭圆内部时，方程8.10所代表的直线似乎难以描述，现在我们给它取一个好听的名字：极线。准确的说，是点 P 关于该椭圆的极线，而点 P 则称为直线8.10关于椭圆的极点。当点 P 在椭

圆曲线上时，它关于椭圆的极线就是椭圆在该点处的切线（相切），当点 P 在椭圆外部时，它关于椭圆的极线就是对应的切点弦所在直线（相交），而当点 P 在椭圆内时，对应的极线就在椭圆外部了（相离）。而且我们看到，点 P 无论位于哪个位置（椭圆中心除外），它关于椭圆的极线的方向都是 OP 的共轭方向，所以如果让点 P 在一条从椭圆中心出发的射线上由椭圆内到椭圆外缓慢移动，我们将看到它关于椭圆的极线由远及近平行移动，当点 P 移到椭圆上的那一刻，极点刚好位于极线上，而极线也成为过极点的切线，此后，当点 P 移动到椭圆外时，极线则与椭圆椭圆相交，当点 P 向无穷远处移动时，极线则向椭圆中心无限靠近。容易看出，对于平面上任何一点（椭圆中心除外），它关于椭圆的极线都是唯一存在的，反过来，对于平面上任何一条不通过椭圆中心的直线，它关于椭圆的极点也是唯一存在的。而切点与切线，不过是极点刚好位于极线上的一种特殊情况罢了。

第九章　立体几何

三面角

9.1　三面角的余弦公式

参考文献

[1] 菲赫金哥尔茨. 微积分学教程. 高等教育出版社, 2006.

[2] 钱展望, 朱华伟. 奥林匹克数学. 湖北教育出版社, 2002.

[3] 陈传理, 张同君. 竞赛数学教程. 高等教育出版社, 2002.

[4] 北京大学数学系几何与代数教研室前代数小组. 高等代数. 高等教育出版社, 2004.

[5] 华东师范大学数学系, 数学分析. 高等教育出版社, 2004.

[6] 闵嗣鹤, 严士健. 初等数论. 人民教育出版社, 1983. 1983.

[7] kuing. kuing 网络撸题集. 未出版, 2015.