不等式选讲

余周炜*

2019 年 10 月 26 日

*Email: yuzhouwei6326@outlook.com

目录

1 均值不等式 3
1.1 证明 . . . 3
1.2 推广 . . . 4
1.2.1 加权算术-几何平均不等式 . . . 4
1.2.2 矩阵形式 . . . 4
1.2.3 极限形式 . . . 4

2 幂平均不等式 5
2.1 特例 . . . 5

3 伯努利不等式 6
3.1 证明 . . . 6
3.2 推广 . . . 6
3.2.1 实数域形式 . . . 6
3.2.2 一般形式 . . . 7

4 权方和不等式 7

5 琴生不等式 8
5.1 凸函数 . . . 8
5.2 琴生不等式 . . . 8
5.3 证明 . . . 9

6 柯西不等式 10
6.1 证明 . . . 10
6.2 推广 . . . 11
6.3 拉格朗日恒等式 . . . 11

7 赫尔德不等式 12

8 排序不等式 12
8.1 证明 . . . 12

9 切比雪夫总和不等式 13

10 三角不等式 13

11 闵可夫斯基不等式 13

1 均值不等式

设 $a_1, a_2, \cdots, a_n \in \mathbb{R}^+$, 将

$$A_n = \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{\sum_{i=1}^n a_i}{n}$$

$$G_n = \sqrt[n]{a_1 a_2 \cdots a_n} = \left(\prod_{i=1}^n a_i\right)^{\frac{1}{n}} = \exp\left[\frac{1}{n}\sum_{i=1}^n \ln a_i\right]$$

$$H_n = \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{a_i}}$$

$$Q_n = \sqrt{\frac{a_1^2 + a_2^2 + \cdots + a_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^n a_i^2}{n}}$$

分别叫做这 n 个正数的算术平均数、几何平均数、调和平均数和平方平均数，有下面的**均值不等式**成立：

$$H_n \le G_n \le A_n \le Q_n \tag{1.1}$$

当且仅当 $a_1 = a_2 = \cdots = a_n$ 时取等号。

1.1 证明

证明. 1. 当 $n = 1$ 时，显然成立。

2. 设 $n = k(k \ge 1)$ 时,$A_k \ge G_k$,

$$\underbrace{x_{k+1} + (k-1)G_{k+1}}_{k \quad numbers} \ge k\sqrt[k]{x_{k+1}G_{k+1}^{k-1}}$$

3. $n = k+1$ 时,

$$\begin{aligned}
&\frac{x_1 + x_2 + \cdots + x_k + x_{k+1} + (k-1)G_{k+1}}{k} \\
&\ge G_k + \sqrt[k]{x_{k+1}G_{k+1}^{k-1}} \\
&\ge 2\sqrt[2k]{G_k^k x_{k+1} G_{k+1}^{k-1}} = 2G_{k+1}
\end{aligned}$$

$$\begin{aligned}
\therefore (k+1)A_{k+1} &= x_1 + x_2 + \cdots + x_k + x_{k+1} \\
&\ge 2kG_{k+1} - (k-1)G_{k+1} \\
&= (k+1)G_{k+1} \implies A_{k+1} \ge G_{k+1}
\end{aligned}$$

也成立。

$\therefore \forall n \in \mathbb{N}^+, A_n \ge G_n$ □

1.2 推广

1.2.1 加权算术-几何平均不等式

不仅“均匀”的算术平均数和几何平均数之间有不等式，加权的算术平均数和几何平均数之间也有不等式。设 $x_1, \cdots, x_n$ 和 $p_1, \cdots, p_n$ 为正实数，并且 $p_1 + p_2 + \cdots + p_n = 1$, 那么:

$$p_1x_1 + p_2x_2 + \cdots + p_nx_n \geq x_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$$

1.2.2 矩阵形式

算术-几何平均不等式可以看成是一位向量系数的平均数不等式。对于二维矩阵，一样有类似的不等式：对于系数都是正实数的矩阵

$$\begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} \end{bmatrix}$$

设

$$A_j = \frac{1}{n}\sum_{i=1}^{n} a_{ij} \qquad G_i = \sqrt[k]{\prod_{j=1}^{k} a_{ij}}$$

，那么有：

$$\sqrt[k]{A_1A_2\cdots A_k} \leqslant \frac{G_1 + G_2 + \cdots + G_n}{n}$$

也就是说：对 k 个纵列取算术平均数，它们的几何平均小于等于对 n 个横行取几何平均数的算术平均。

1.2.3 极限形式

也称为**积分形式**：对任意在区间 $[0,1]$ 上可积的正值函数 f, 都有：

$$\int_0^1 f(x)\,\mathrm{d}x \geq \exp\left(\int_0^1 \ln f(x)\,\mathrm{d}x\right) \tag{1.2}$$

这实际上是在算术-几何平均值不等式取成

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \exp\left(\frac{\ln x_1 + \ln x_2 + \cdots + \ln x_n}{n}\right)$$

后，将两边的黎曼和中的 n 趋于无穷大后得到的形式。

2 幂平均不等式

幂平均 (power mean), 也叫广义平均 (generalized mean)，是毕达哥拉斯平均（包含了算术，几何，调和平均）的一种抽象化。

若 p 是一非零实数，可定义实数 $x_1, \cdots, x_n$ 的 p 次幂平均为

$$M_p(x_1, \cdots, x_n) = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i^p \right)^{\frac{1}{p}} \tag{2.1}$$

一般地，如果 $p < q$, 则 $M_p(x_1, \cdots, x_n) \leq M_q(x_1, \cdots, x_n)$ 且这两个平均相等当且仅当 $x_1 = x_2 = \cdots = x_n$。这由事实

$$\forall p \in \mathbb{R}, \frac{\partial M_p(x_1, \cdots, x_n)}{\partial p} \geq 0 \tag{2.2}$$

得出，上述不等式可由琴生不等式证明。

特别地，对 $p \in \{-1, 0, 1, 2\}$, 幂平均不等式蕴含了均值不等式。

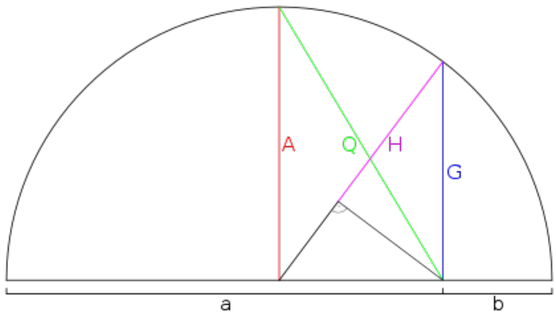

图 1: 特例:n=2 时的图形描述

2.1 特例

最小值:

$$M_{-\infty}(x_1, \ldots, x_n) = \lim_{p \to -\infty} M_p(x_1, \ldots, x_n) = \min\{x_1, \ldots, x_n\}$$

调和平均:

$$M_{-1}(x_1, \ldots, x_n) = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}}$$

几何平均:

$$M_0(x_1, \ldots, x_n) = \lim_{p \to 0} M_p(x_1, \ldots, x_n) = \sqrt[n]{x_1 \cdot \cdots \cdot x_n}$$

算术平均

$$M_1(x_1,\ldots,x_n)=\frac{x_1+\cdots+x_n}{n}$$

平方平均

$$M_2(x_1,\ldots,x_n)=\sqrt{\frac{x_1^2+\cdots+x_n^2}{n}}$$

立方平均

$$M_3(x_1,\ldots,x_n)=\sqrt[3]{\frac{x_1^3+\cdots+x_n^3}{n}}$$

最大值

$$M_{+\infty}(x_1,\ldots,x_n)=\lim_{p\to\infty}M_p(x_1,\ldots,x_n)=\max\{x_1,\ldots,x_n\}$$

3 伯努利不等式

$\forall n\in\mathbb{N}, x\geq -1,$

$$(1+x)^n\geq 1+nx \tag{3.1}$$

如果 $n\geq 0$ 且是偶数，则不等式对任意实数 x 成立。

可以看到在 $n=0,1$，或 $x=0$ 时等号成立，而对于任意正整数 $n\geq 2$ 和任意实数 $x\geq -1, x\neq 0$, 有严格不等式：

$$(1+x)^n>1+nx \tag{3.2}$$

3.1 证明

可以用数学归纳法证明：当 $n=0,1$，不等式明显成立。假设不等式对正整数 n, 实数 $x\geq -1$ 时成立，那么

$$\begin{aligned}(1+x)^{n+1}&=(1+x)(1+x)^n &(3.3)\\ &\geq (1+x)(1+nx) &(3.4)\\ &=1+(n+1)x+nx^2 &(3.5)\\ &\geq 1+(n+1)x &(3.6)\end{aligned}$$

3.2 推广

3.2.1 实数域形式

如果 $x>-1$, 那么：

- 若 $r\leq 0$ 或 $r\geq 1$, 有 $(1+x)^r\geq 1+rx$
- $0\leq r\leq 1$, 有 $(1+x)^r\leq 1+rx$

r=0,1 时，等式显然成立；除此之外，等号成立当且仅当 $x=0$。

我们用导数来证明该结论：

证明. 在 $(-1,+\infty)$ 上定义 $f(x)=(1+x)^r-(1+rx)$, 其中 $r\neq 0,1$，对 x 求导得 $f'(x)=r(1+x)^{r-1}-r$, 则 $f'(x)=0$ 当且仅当 $x=0$。分情况讨论：

1° $0<r<1$，则对 $x>0,f'(x)<0$; 对 $-1<x<0,f'(x)>0$。因此 $f(x)$ 在 $x=0$ 时取最大值 0，故得 $(1+x)^r\leq 1+rx$。

2° $r<0$ 或 $r>1$, 则对 $x>0,f'(x)>0$; 对 $-1<x<0,f'(x)<0$。因此 $f(x)$ 在 $x=0$ 时取最小值 0，故得 $(1+x)^r\geq 1+rx$。

在这两种情况下，等号成立当且仅当 $x=0$。 □

3.2.2 一般形式

$$(1+x_1)(1+x_2)\dots(1+x_n)\geq 1+x_1+x_2+\dots+x_n \tag{3.7}$$

其中 x_i 均为同号且大于等于 -1 的实数。（该证明摘自 [2]）：

证明. 构造数列

$$x_n=(1+a_1)(1+a_2)\cdots(1+a_n)-(1+a_1+a_2+\cdots+a_n)(n\geq 2)$$

则 $x_{n+1}-x_n=a_{n+1}[(1+a_1)\cdots(1+a_n)-1]$。

若 $a_i>0(i=1,2,\cdots,n+1)$, 由上式易见 $x_{n+1}>x_n$; 若 $-1<a_i<0(i=1,2,\cdots,n)$, 则 $0<1+a_i<1$，也有 $x_{n+1}>x_n$

因此 $\{x_n\}$ 是一个单调递增的序列 $(n\geq 2)$。由于 $x_2=(1+a_1)(1+a_2)-(1+a_1+a_2)=a_1a_2>0$, 则对一切 $n\geq 2,x_n>0$, 从而原不等式成立。 □

4 权方和不等式

设 $a_i,b_i>0,i=1,2,\cdots,n,k\in N^+$, 则

$$\frac{a_1^{k+1}}{b_1^k}+\frac{a_2^{k+1}}{b_2^k}+\cdots+\frac{a_n^{k+1}}{b_n^k}\geq\frac{(a_1+a_2+\cdots+a_n)^{k+1}}{(b_1+b_2+\cdots+b_n)^k} \tag{4.1}$$

证明. 令

$$s=(a_1+a_2+\cdots+a_n)^{-1} \tag{4.2}$$

$$t=(b_1+b_2+\cdots+b_n)^{-1} \tag{4.3}$$

□

则原不等式等价于 $\sum_{i=1}^{n}\frac{(sa_i)^{k+1}}{(tb_i)^k}\geq 1$

由伯努利不等式有:

$$\frac{(sa_i)^{k+1}}{(tb_i)^k} = tb_i\cdot\left(\frac{sa_i}{tb_i}\right)^{k+1} \tag{4.4}$$

$$\geq tb_i\left[1+(k+1)(\frac{sa_i}{tb_i}-1)\right] \tag{4.5}$$

$$tb_i\left[(k+1)\frac{sa_i}{tb_i}-k\right]=(k+1)sa_i-ktb_i \tag{4.6}$$

则

$$\sum_{i=1}^{n}\frac{(sa_i)^{k+1}}{(tb_i)^k}\geq\sum_{i=1}^{n}[(k+1)sa_i-ktb_i]=(k+1)-k=1 \tag{4.7}$$

此即

$$\frac{a_1^{k+1}}{b_1^k}+\frac{a_2^{k+1}}{b_2^k}+\cdots+\frac{a_n^{k+1}}{b_n^k}\geq\frac{(a_1+a_2+\cdots+a_n)^{k+1}}{(b_1+b_2+\cdots+b_n)^k} \tag{4.8}$$

5 琴生不等式

5.1 凸函数

设连续函数 $f(x)$ 的定义域为 $[a,b]$(或开区间 (a,b))，对于区间 $[a,b]$ 内任意两点 x_1,x_2 都有：

$$f(\frac{x_1+x_2}{2})\leq\frac{f(x_1)+f(x_2)}{2} \tag{5.1}$$

则称 $f(x)$ 为 $[a,b]$ 上的凸函数 (下凸函数)，若把不等号反向，则称这样的 $f(x)$ 为区间 $[a,b]$ 上的凹函数（上凸函数）。

凸函数的几何意义是，过 $y=f(x)$ 曲线上任意两点作弦，则弦的中点必在该曲线上方或在曲线上。

5.2 琴生不等式

琴生不等式 (Jensen's inequality) 以丹麦数学家约翰 · 琴生 (Johan Jensen) 命名。它给出积分的凸函数值和凸函数的积分值之间的关系。它于 1906 年被琴生证明。该不等式说明平均数的凸变换小于或等于凸变换之后的平均数。

两点形式：

$$f(tx_1+(1-t)x_2)\leq tf(x_1)+(1-t)f(x_2)$$

其中 $t\in[0,1]$。

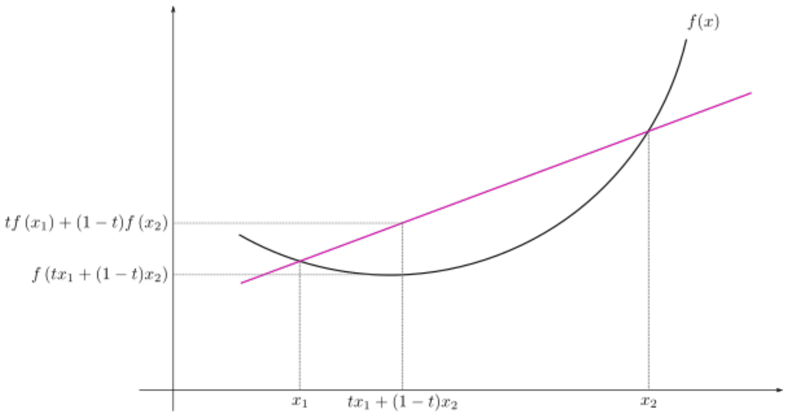

图 2: 凸函数的割线在该函数图像的上方

一般形式：对于任意点集 $\{x_i\}$, 若 $\lambda_i \geq 0$ 且 $\sum_i \lambda_i = 1$, 使用数学归纳法，可以证明凸函数 f(x) 满足：

$$f(\sum_i \lambda_i x_i) \leq \sum_i \lambda_i f(x_i) \tag{5.2}$$

该式被称为 Jensen 不等式。

在概率论中，如果 X 是一个随机变量，φ 是一个凸函数：

$$\varphi(\mathrm{E}[X]) \leq \mathrm{E}[\varphi(X)] \tag{5.3}$$

5.3 证明

证明. 1. n=1,2 时，由凸函数的定义成立。

2. 假设 $n = k$ 时，公式(5.2)成立。

3. n=k+1 时

$$f(\sum_{i=1}^{k+1} \lambda_i x_i) = f(\lambda_{k+1} x_{k+1} + \sum_{i=1}^{k} \lambda_i x_i) \tag{5.4}$$

$$= f(\lambda_{k+1} x_{k+1} + (1-\lambda_{k+1}) \sum_{i=1}^{k} \eta_i x_i) \tag{5.5}$$

其中

$$\eta_i = \frac{\lambda_i}{1-\lambda_{k+1}} \tag{5.6}$$

由公式(5.2)的结论和公式(5.5)有：

$$f(\sum_{i=1}^{k+1} \lambda_i x_i) \leq \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1}) f(\sum_{i=1}^{k} \eta_i x_i)) \tag{5.7}$$

注意到 λ_i 满足：

$$\sum_{i=1}^{k+1} \lambda_i = 1 \tag{5.8}$$

因此：

$$\sum_{i=1}^{k} \lambda_i = 1 - \lambda_{k+1} \tag{5.9}$$

因此 η_i 也满足：

$$\sum_{i}^{k} \eta_i = \frac{\sum_{i=1}^{k} \lambda_i}{1 - \lambda_{k+1}} = 1 \tag{5.10}$$

由公式(5.2)和(5.10)得到：

$$\sum_{i=1}^{k} f(\eta_i x_i) \leq \sum_{i=1}^{k} \eta_i f(x_i) \tag{5.11}$$

由公式(5.7)和式(5.11):

$$f(\sum_{i=1}^{k+1} \lambda_i x_i) \leq \lambda_{k+1} f(x_{k+1}) + (1 - \lambda_{k+1}) \sum_{i=1}^{k} \eta_i f(x_i) = \sum_{i=1}^{k+1} \lambda_i f(x_i) \tag{5.12}$$

因此 $i = k + 1$ 时，Jensen 不等式成立。

综上，Jensen 不等式成立。 □

6 柯西不等式

柯西不等式 (Cauchy-Schwarz inequality) 是数学中最重要的不等式之一，对于内积空间中的任意两个向量 $\mathbf{u}, \mathbf{v}$ 有:

$$|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \cdot \langle \mathbf{v}, \mathbf{v} \rangle \tag{6.1}$$

其中 $\langle \cdot, \cdot \rangle$ 是内积。在欧几里得空间中，即：

$$\left(\sum_{i=1}^{n} u_i v_i \right)^2 \leq \left(\sum_{i=1}^{n} u_i^2 \right) \left(\sum_{i=1}^{n} v_i^2 \right) \tag{6.2}$$

引入向量的范数，则该不等式可以写作：

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\| \tag{6.3}$$

当前仅当 $\mathbf{u}$ 和 $\mathbf{v}$ 线性相关时取等号。

6.1 证明

证明. 下面分两种情况讨论。

1° 平凡情形：$\mathbf{v}=0$，该定理显然成立。

2° 假定 $\mathbf{v}\neq 0$ 令 $\lambda=\langle u,v\rangle/\|v\|^2$

$$
\begin{aligned}
0 &\leq \|u-\lambda\cdot v\|^2 && (6.4)\\
&= \langle u,u\rangle-\langle\lambda\cdot v,u\rangle-\langle u,\lambda\cdot v\rangle+\langle\lambda\cdot v,\lambda\cdot v\rangle && (6.5)\\
&= \langle u,u\rangle-\lambda\langle v,u\rangle-\lambda\langle u,v\rangle+\lambda^2\langle v,v\rangle && (6.6)\\
&= \|u\|^2-2\lambda\langle u,v\rangle+\lambda^2\|v\|^2 && (6.7)\\
&= \|u\|^2-2\frac{|<u,v>|^2}{\|v\|^2}+\frac{|\langle u,v\rangle|^2}{\|v\|^2} && (6.8)\\
&= \|u\|^2-\frac{|\langle u,v\rangle|^2}{\|v\|^2} && (6.9)
\end{aligned}
$$

因此：$|\langle u,v\rangle|\leq\|u\|\cdot\|v\|$ 成立。

□

6.2 推广

积分形式 对于平方可积的复值函数，有：

$$\left|\int f(x)g(\bar{x})\,\mathrm{d}x\right|^2\leq\int|f(x)|^2\,\mathrm{d}x\cdot\int|g(x)|^2\,\mathrm{d}x$$

期望形式 设 X,Y 为任意两个随机变量，有：

$$[\mathbf{E}(XY)]^2\leq\mathbf{E}(X^2)\mathbf{E}(Y^2) \tag{6.10}$$

以下证明：

$$g(t)=\mathbf{E}((tX-Y)^2)=t^2(\mathbf{E}(X^2))-2t\mathbf{E}(XY)+\mathbf{E}(Y^2) \tag{6.11}$$

$$\because (tX-Y)^2\geq 0, g(t)\geq 0 \tag{6.12}$$

$$\Delta=(2\mathbf{E}(XY))^2-4(\mathbf{E}(X^2)\mathbf{E}(Y^2))\leq 0 \tag{6.13}$$

当且仅当 $tX=Y$ 时有 $g(t)=0$ 所以：

$$[\mathbf{E}(XY)]^2\leq\mathbf{E}(X^2)\mathbf{E}(Y^2) \tag{6.14}$$

6.3 拉格朗日恒等式

在三维欧氏空间中有下列关系：

$$\|\mathbf{a}\|^2 \cdot \|\mathbf{b}\|^2 = (\mathbf{a} \cdot \mathbf{b})^2 + |\mathbf{a} \times \mathbf{b}|^2 \tag{6.15}$$

7 赫尔德不等式

8 排序不等式

排序不等式又称排序原理，设 $a_1 \leq a_2 \leq \cdots a_n, b_1 \leq b_2 \leq \cdots b_n, c_1, c_2, \cdots, c_n$ 是 $b_1, b_2, \cdots, b_n$ 的任一排列，则

$$\underbrace{a_1b_n + a_2b_{n-1} + \cdots + a_nb_1}_{\text{反序和}} \leq \underbrace{a_1c_1 + a_2c_2 + \cdots a_nc_n}_{\text{乱序和}} \leq \underbrace{a_1b_1 + a_2b_2 + \cdots + a_nb_n}_{\text{顺序和}} \tag{8.1}$$

当且仅当 $a_1 = a_2 = \cdots = a_n$ 或 $b_1 = b_2 = \cdots b_n$ 时，反序和等于顺序和。

8.1 证明

设 $a_1 \leq a_2 \leq \cdots a_n, b_1 \leq b_2 \leq \cdots b_n$ 为两组实数，$c_1, c_2, \cdots c_n$ 是 $b_1, b_2, \cdots, b_n$ 的任一排列，因为 $b_1, b_2, \cdots, b_n$ 的全排列只有 $n!$ 个，所以：

$$S = a_1c_1 + a_2c_2 + \cdots + a_nc_n \tag{8.2}$$

的不同的值也只有有限个 (个数 $\leq n!$), 其中必有最大值和最小值。

考虑式(8.2), 若 $c_1 \neq b_1$, 则有某 $c_k = b_1(k > 1), c_1 > c_k$, 将(8.2)对换，得

$$S' = a_1c_k + \cdots + a_kc_1 + \cdots a_nc_n \tag{8.3}$$

(8.3)-(8.2)得:

$$S' - S = a_1c_k + a_kc_1 - a_1c_1 - a_kc_k = (a_k - a_1)(c_1 - c_k) \geq 0 \tag{8.4}$$

这说明将(8.3)调换为 a_1b_1 后，和式不减小。

若 $c_1 = b_1$, 则转而考察 c_2，并进行类似讨论。

类似地，可以证明，将(8.2)中的第一项换为 a_1b_1，第二项换位 a_2b_2 后，和式不减小。

如此下去，经过有限步调整，可以一切和数中，最大和数所对应的情况只能是数组 $\{c_i\}$ 由小到大排序的情况，最大和数是顺序和。

同理可证，最小和数是反序和。

9 切比雪夫总和不等式

10 三角不等式

11 闵可夫斯基不等式

参考文献

[1] Wikipedia

[2] 不等式的解题方法与技巧。

[3] 贝努利不等式的几个推论及应用。