

初等数学平均值和向量范数的关系

作者已匿名

本文是我（华东师范大学某社科院系本科生）跨专业修读视觉与机器智能研究所 方发明教授面向计算机科学与技术专业本科生开设的专业选修课《线性代数进阶》完成的课程作业，由于获得最高评级成绩，所以才敢公开

摘要： 本文对于列向量求多种定义的平均值，检验它们与范数、半范数、拟范数定义的关系；从定义出发，讨论了不同定义的平均值视为范数的可能性，也发现了一部分形式的平均值与范数的相似构造会怎样影响其用作距离度量的性质。未来的研究可以探索平均值在其他数学领域中的应用，以及它们在解决实际问题中的潜力。

关键词： 矩阵分析基础、范数三大公理、平均值、初等数学

一 背景

以前我们在中小学的数学课中接触过多种定义的平均值，包括但不限于——算术平均、几何平均、平方平均、调和平均、对数平均，等等。同样地，也接触过特殊定义的 L_p -范数，以及它和平均值的异同。学习到矩阵分析的基础知识[1]和范数的更一般定义[2]之后，我们就难免想知道：

模长的平均值可不可以也看作范数？考虑到中学阶段常常应用均值不等式与 L_p -范数不等式的“反向”[3]，它们的结构是否存在一些共通？以下将对于列向量 $\boldsymbol{x} = (x_1, x_2, x_3, \dots, x_n)^T$ 求多种定义的平均值，检验它们与范数、半范数、拟范数定义的关系。

二 正文

2.1 概念回顾

根据本课程使用主教材[4]，以上三种定义即实函数 $f(\boldsymbol{x}) : V \rightarrow \mathbb{R}, \forall \boldsymbol{x}, \boldsymbol{y} \in V, c \in \mathbb{K}$ （其中 \mathbb{K} 表示 \mathbb{R} 或 \mathbb{C} ）须满足以下条件：

类型	正定性	齐次性	三角不等式
范数	$f(\boldsymbol{x}) \geq 0$ $f(\boldsymbol{x}) = 0 \Leftrightarrow \boldsymbol{x} = \mathbf{0}$	$f(c\boldsymbol{x}) = c \cdot f(\boldsymbol{x})$ ($\forall c \in \mathbb{C}$)	$f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$
半范数	$f(\boldsymbol{x}) \geq 0$ $\boldsymbol{x} = \mathbf{0} \Rightarrow f(\boldsymbol{x}) = 0$ $f(\boldsymbol{x}) = 0 \not\Rightarrow \boldsymbol{x} = \mathbf{0}$	$f(c\boldsymbol{x}) = c \cdot f(\boldsymbol{x})$ ($\forall c \in \mathbb{C}$)	$f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$

类型	正定性	齐次性	三角不等式
拟范数	$f(\mathbf{x}) \geq 0$ $f(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$	$f(c\mathbf{x}) = c \cdot f(\mathbf{x})$ ($\forall c \in \mathbb{C}$)	$f(\mathbf{x} + \mathbf{y}) \leq C(f(\mathbf{x}) + f(\mathbf{y}))$ ($C \in \mathbf{R}^+, C \neq 1$)

2.2 教材正文中“请读者自证”的两条概念结论

2.2.1 命题内容

$$\|x\|_p = \left(\sum_{i=1}^m |x_i|^p \right)^{\frac{1}{p}} \text{ 在 } 0 < p < 1 \text{ 时是拟范数, 在 } p \geq 1 \text{ 时是范数。}$$

2.2.2 尝试解答

2.2.2.1 正定性

$$\left\{ \begin{array}{l} p \in \mathbb{R}^+ \\ |x_i| \geq 0 \end{array} \right. \Rightarrow |x_i|^p \geq 0 \Rightarrow \sum_{i=1}^n |x_i|^p \geq 0 \Rightarrow \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \geq 0 \Rightarrow f(\mathbf{x}) \geq 0$$

以上不等号取等当且仅当: $|x_i| = 0, \forall i \in \{1, 2, 3, \dots, n\}$, 也即 $\mathbf{x} = \mathbf{0}$, 故正定性得证。

2.2.2.2 齐次性

$$\begin{aligned} f(c\mathbf{x}) &= \left(\sum_{i=1}^n |cx_i|^p \right)^{\frac{1}{p}} \\ &= \left(|c|^p \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} = |c|^{p \cdot \frac{1}{p}} \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} = |c| \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \\ &= |c| f(\mathbf{x}) \end{aligned}$$

2.2.2.3 三角不等式

2.2.2.3.1 $p = 1$ 时

$$f(\mathbf{x} + \mathbf{y}) = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = f(\mathbf{x}) + f(\mathbf{y})$$

2.2.2.3.2 Young 不等式

若 $p, q \in \mathbb{R}$, $p > 1$, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, 则有

$$|xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q} \quad \forall x, y \in \mathbb{R}$$

证明 (作商法):

(1) 当 $xy = 0$ 时, 不妨设 $y = 0$, $0 \leq \frac{|x|^p}{p}$ 显然成立;

(2) 当 $xy \neq 0$ 时, 构造函数

$$\begin{aligned} g(t) &= \frac{t^p}{p} + \frac{t^{-q}}{q} \quad (t \in \mathbb{R}^+) \\ \Rightarrow \frac{\partial g}{\partial t} &= p \cdot \frac{t^{p-1}}{p} + (-q) \cdot \frac{t^{-q-1}}{q} \\ &= t^{p-1} - t^{-(q+1)} = t^{p-1} - \frac{1}{t^{q+1}} \\ \Rightarrow \frac{\partial^2 g}{\partial t^2} &= (p-1)t^{p-2} - (-q-1)t^{-q-2} \\ &= (p-1)t^{p-2} + (q+1) \cdot \frac{1}{t^{q+2}} > 0 \\ \Rightarrow \frac{\partial g}{\partial t} \Big|_{t=1} &= 1^{p-1} - \frac{1}{1^{q+1}} \\ &= 1 - 1 \\ &= 0 \\ \Rightarrow &\begin{cases} \frac{\partial g}{\partial t} \Big|_{t < 1} < 0 \\ \frac{\partial g}{\partial t} \Big|_{t > 1} > 0 \end{cases} \\ \Rightarrow \min_t \{g(t)\} &= g(1) \\ &= \frac{1^p}{p} + \frac{1^{-q}}{q} \\ &= \frac{1}{p} + \frac{1}{q} = 1 \\ \Rightarrow g(t) &\geq \min_t \{g(t)\} = 1 \\ \Rightarrow \frac{t^p}{p} + \frac{t^{-q}}{q} &\geq 1 \end{aligned}$$

$$\begin{aligned}
 \frac{\frac{|x|^p}{p} + \frac{|y|^q}{q}}{|xy|} &= \frac{\frac{|x|^{\frac{p}{q}+1}}{p} + \frac{|y|^{\frac{q}{p}+1}}{q}}{|xy|} \quad \Longleftarrow \quad \frac{1}{p} + \frac{1}{q} = 1 \Rightarrow p + q = pq \Rightarrow \begin{cases} p = \frac{p}{q} + 1 \\ q = \frac{q}{p} + 1 \end{cases} \\
 &= \frac{1}{p} \frac{|x|^{\frac{p}{q}+1}}{|xy|} + \frac{1}{q} \frac{|y|^{\frac{q}{p}+1}}{|xy|} \\
 &= \frac{1}{p} \frac{|x|^{\frac{p}{q}} |x|}{|x| \cdot |y|} + \frac{1}{q} \frac{|y|^{\frac{q}{p}} |y|}{|x| \cdot |y|} \\
 &= \frac{1}{p} \frac{|x|^{\frac{p}{q}}}{|y|} + \frac{1}{q} \frac{|y|^{\frac{q}{p}}}{|x|} \\
 &= \frac{1}{p} \left(\frac{|x|^{\frac{1}{q}}}{|y|^{\frac{1}{p}}} \right)^p + \frac{1}{q} \left(\frac{|y|^{\frac{1}{p}}}{|x|^{\frac{1}{q}}} \right)^q \\
 &= \frac{1}{p} \left(\frac{|x|^{\frac{1}{q}}}{|y|^{\frac{1}{p}}} \right)^p + \frac{1}{q} \left(\frac{|x|^{\frac{1}{q}}}{|y|^{\frac{1}{p}}} \right)^{-q} \\
 &= g \left(\frac{|x|^{\frac{1}{q}}}{|y|^{\frac{1}{p}}} \right) \geq 1 \\
 &\Rightarrow \frac{|x|^p}{p} + \frac{|y|^q}{q} \geq |xy| \\
 &\Rightarrow |xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q}
 \end{aligned}$$

2.2.2.3 Hölder 不等式

若 $p, q, r \in \mathbb{R}^+$, $p > 1$, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, 则有

$$\sum_{i=1}^n |x_i| \cdot |y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}} \quad \forall x_i, y_i \in \mathbb{C}, \quad i \in \{1, 2, 3, \dots, n\}$$

$$\begin{aligned} \frac{\sum_{i=1}^n |x_i| \cdot |y_i|}{\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}} &= \sum_{i=1}^n \left(\frac{|x_i|}{\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}} \cdot \frac{|y_i|}{\left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}} \right) \\ &= \sum_{i=1}^n \left| \frac{|x_i|}{\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}} \right| \left| \frac{|y_i|}{\left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}} \right| \\ &\leq \sum_{i=1}^n \left\{ \frac{\left| \frac{|x_i|}{\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}} \right|^p}{p} + \frac{\left| \frac{|y_i|}{\left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}} \right|^q}{q} \right\} \Leftarrow |xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q} \\ &= \sum_{i=1}^n \left\{ \frac{|x_i|^p}{p \left| \sum_{i=1}^n |x_i|^p \right|} + \frac{|y_i|^q}{q \left| \sum_{i=1}^n |y_i|^q \right|} \right\} \\ &= \frac{\sum_{i=1}^n |x_i|^p}{p \left| \sum_{i=1}^n |x_i|^p \right|} + \frac{\sum_{i=1}^n |y_i|^q}{q \left| \sum_{i=1}^n |y_i|^q \right|} \\ &= \frac{1}{p} + \frac{1}{q} = 1 \quad \Rightarrow \quad \sum_{i=1}^n |x_i| \cdot |y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}} \end{aligned}$$

2.2.2.3.4 $p > 1$ 时

设 $q \in \mathbb{R}^+$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$ 则有

$$\frac{1}{q} = 1 - \frac{1}{p} \Rightarrow \begin{cases} 1 - \frac{1}{q} = \frac{1}{p} \\ q = \frac{1}{1 - \frac{1}{p}} = \frac{p}{p-1} \Rightarrow q(p-1) = p \end{cases}$$

$$\begin{aligned} \sum_{i=1}^n |x_i + y_i|^p &= \sum_{i=1}^n |x_i + y_i|^{p-1} \cdot |x_i + y_i| \\ &\leq \sum_{i=1}^n |x_i + y_i|^{p-1} (|x_i| + |y_i|) \\ &= \sum_{i=1}^n |x_i + y_i|^{p-1} |x_i| + \sum_{i=1}^n |x_i + y_i|^{p-1} |y_i| \\ &= \sum_{i=1}^n |x_i| \cdot |(x_i + y_i)^{p-1}| + \sum_{i=1}^n |y_i| \cdot |(x_i + y_i)^{p-1}| \\ &\leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |(x_i + y_i)^{p-1}|^q \right)^{\frac{1}{q}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |(x_i + y_i)^{p-1}|^q \right)^{\frac{1}{q}} \\ &= \left(\sum_{i=1}^n |x_i + y_i|^{q(p-1)} \right)^{\frac{1}{q}} \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \\ &= \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{q}} \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \\ &\Rightarrow \sum_{i=1}^n |x_i + y_i|^p \leq \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{q}} \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \\ &\Rightarrow \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{1 - \frac{1}{q}} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \\ &\Rightarrow \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \\ &\Rightarrow f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \end{aligned}$$

2.3 “次数型”均值常见特例

2.3.1 算术平均 (Arithmetic Mean)

对于列向量 \mathbf{x} 求模长的算术平均值有

$$AM(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n |x_i|$$

2.3.2 平方平均 (Root Mean Square)

对于列向量 \mathbf{x} 求模长的平方平均值有

$$RM(\mathbf{x}) = \sqrt{\frac{1}{n} \sum_{i=1}^n |x_i|^2}$$

2.3.3 调和平均 (Harmonious Mean)

对于列向量 \mathbf{x} 求模长的调和平均值有

$$HM(\mathbf{x}) = \frac{n}{\sum_{i=1}^n \left| \frac{1}{x_i} \right|}$$

2.4 “次数型”一般情形

$$f(\mathbf{x}) = \left(\frac{\sum_{i=1}^n |x_i|^p}{n} \right)^{\frac{1}{p}}$$

2.4.1 正定性

$f(\mathbf{x}) \geq 0$ 总是成立，但需验证取零的等价关系。

$$f(\mathbf{x}) = 0 \Rightarrow \sum_{i=1}^n |x_i|^p = 0 \text{ 又 } |x_i| \geq 0 \Rightarrow \mathbf{x} = \mathbf{0}$$

反之是平凡的。

2.4.2 齐次性

$$f(c\mathbf{x}) = \left(\frac{\sum_{i=1}^n |cx_i|^p}{n} \right)^{\frac{1}{p}} = \left(\frac{|c|^p \sum_{i=1}^n |x_i|^p}{n} \right)^{\frac{1}{p}} = |c| \left(\frac{\sum_{i=1}^n |x_i|^p}{n} \right)^{\frac{1}{p}} = |c| f(\mathbf{x})$$

2.4.3 三角不等式

需验证

$$\begin{aligned}
 & \left(\frac{\sum_{i=1}^n |x_i + y_i|^p}{n} \right)^{\frac{1}{p}} \leq \left(\frac{\sum_{i=1}^n |x_i|^p}{n} \right)^{\frac{1}{p}} + \left(\frac{\sum_{i=1}^n |y_i|^p}{n} \right)^{\frac{1}{p}} \\
 \Leftrightarrow & \left(\frac{1}{n} \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left(\frac{1}{n} \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\frac{1}{n} \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \\
 \Leftrightarrow & \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}}
 \end{aligned}$$

这个不等式在前文已证明 $p > 1$ 时的情况。

2.5 其他常见特例

2.5.1 几何平均 (Geometric Mean)

对于列向量 \mathbf{x} 求模长的平方平均值有

$$GM(\mathbf{x}) = \left(\prod_{i=1}^n |x_i| \right)^{\frac{1}{n}}$$

2.5.1.1 正定性

若有一 $x_i = 0$ ，则无论其余元素情况，总有 $GM = 0$ —— 该函数具有半正定性。

2.5.1.2 齐次性

$$f(c\mathbf{x}) = \left(\prod_{i=1}^n |cx_i| \right)^{\frac{1}{n}} = \left(|c|^n \prod_{i=1}^n |x_i| \right)^{\frac{1}{n}} = |c| \left(\prod_{i=1}^n |x_i| \right)^{\frac{1}{n}} = |c| f(\mathbf{x})$$

2.5.1.3 三角不等式

考虑这样两个向量

$$\mathbf{x} = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix} \quad \begin{cases} f(\mathbf{x}) & \cong 2.62 \\ f(\mathbf{y}) & \cong 2.29 \\ f(\mathbf{x} + \mathbf{y}) & \cong 5.24 \end{cases}$$

$$f(\mathbf{x}) + f(\mathbf{y}) - f(\mathbf{x} + \mathbf{y}) \cong -0.33 \leq 0 \Rightarrow f(\mathbf{x}) + f(\mathbf{y}) < f(\mathbf{x} + \mathbf{y})$$

模长的几何平均值与三角不等式条件冲突。

2.5.2 对数平均 (Logarithmic Mean)

对于列向量 $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ 求模长的对数平均值有

$$LM(\mathbf{x}) = \begin{cases} |x_1| = |x_2| & (|x_1| = |x_2|) \\ \frac{|x_1| - |x_2|}{\ln|x_1| - \ln|x_2|} & (|x_1| \neq |x_2|) \end{cases}$$

2.5.2.1 正定性

由于自然对数在 \mathbb{R}^+ 单调递增, $f(\mathbf{x}) \geq 0$ 总是成立。 $\mathbf{x} = \mathbf{0}$ 在该函数上无定义。

2.5.2.2 齐次性

$$\begin{aligned} f(c\mathbf{x}) &= \frac{|cx_1| - |cx_2|}{\ln|cx_1| - \ln|cx_2|} \\ &= \frac{|c|(|x_1| - |x_2|)}{\ln|c| + \ln|x_1| - \ln|c| - \ln|x_2|} \\ &= |c| \frac{|x_1| - |x_2|}{\ln|x_1| - \ln|x_2|} = |c|f(\mathbf{x}) \end{aligned}$$

2.5.2.3 三角不等式

考虑这样两个向量

$$\mathbf{x} = \begin{pmatrix} 3 \\ 7 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 15 \end{pmatrix} \quad \begin{cases} f(\mathbf{x}) & \cong 4.72 \\ f(\mathbf{y}) & \cong 5.17 \\ f(\mathbf{x} + \mathbf{y}) & \cong 10.56 \end{cases}$$

$$f(\mathbf{x}) + f(\mathbf{y}) - f(\mathbf{x} + \mathbf{y}) \cong -0.67 \leq 0 \Rightarrow f(\mathbf{x}) + f(\mathbf{y}) < f(\mathbf{x} + \mathbf{y})$$

模长的对数平均值与三角不等式条件冲突。

2.6 图象观察

对于次数型均值表达式 $\left(\frac{1}{n} \left(\sum_{i=1}^n |x_i|^p \right) \right)^{\frac{1}{p}}$

分别取 $p = 100, 3, 2, 1, \frac{1}{2}, \frac{1}{3}, -\frac{1}{100}, -1$, 在 $n = 2$ 的特殊情况下在二维平面作图显示:

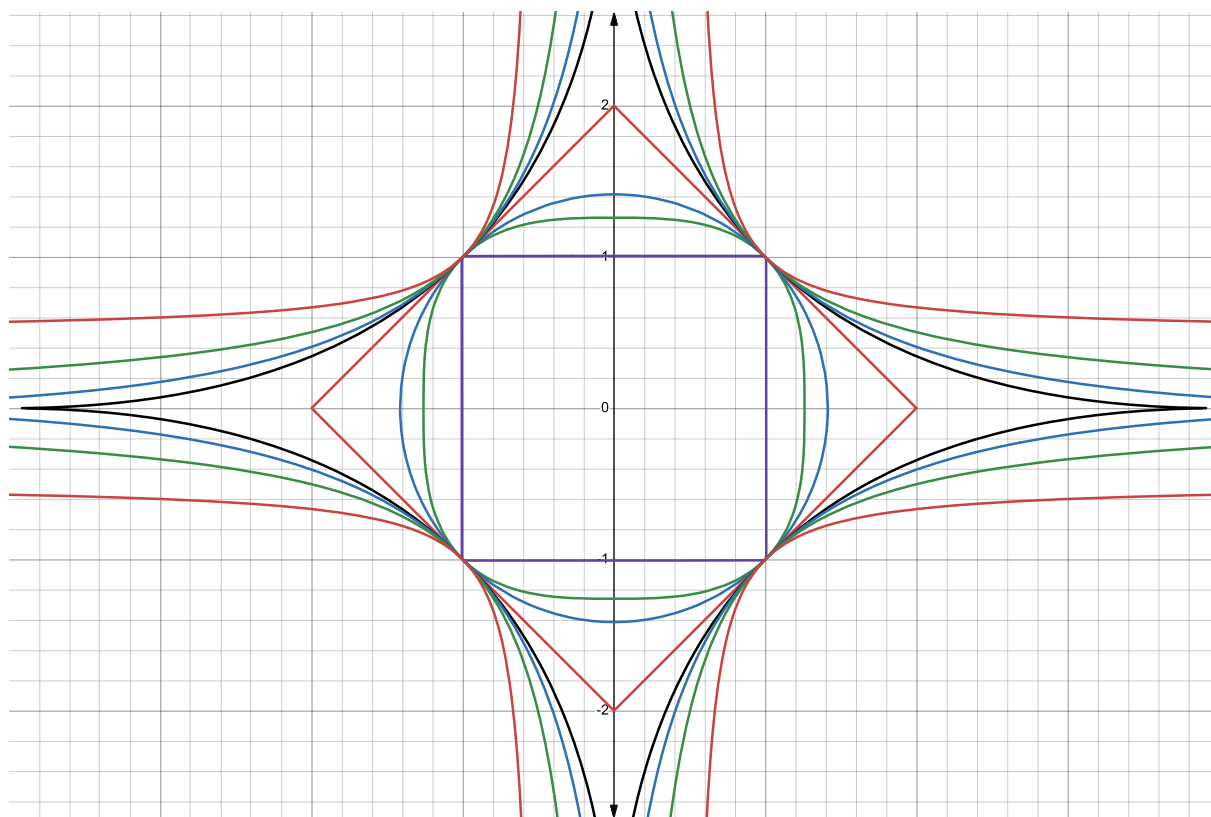


图1 用 x 和 y 作向量元素, 各均值等于 1 的图象

图象随着 p 的减小, 逐渐远离原点。全部图象切于 $|x| = 1, |y| = 1$ 处。当 $p > 0$ 时, 图象封闭; 当 $p < 0$ 时, 图象以坐标轴为渐近线, 伸向无穷远处。这辅助说明了 $p < 0$ 时, 次数型均值既不是范数, 也不是拟范数或半范数的结论。

$p > 1$ 时, 图象是凸集; $p < 1$ 时, 图象不是凸集。

几何平均值等于一的图象在 $p = -1/100$ 图象的内侧, 但非常贴近——这能够显示中学老师“可以把几何均值当成零次均值”论断的某种合理性; 对数平均值等于一的图象在 $p = 1/3$ 图象的内侧, 但非常贴近——这说明对数平均值是比几何平均“次数更大”的“零次平均”。

三 总结与展望

平均值在特定条件下可以被视为范数或拟范数, 但并非所有情况下都满足范数的全部性质。一个特别的例子是, 几何平均和对数平均在某些情况下与三角不等式的要求相悖。

在前文基础上, 我们还可以将平均值和范数的关系讨论得更加深入, 不仅对基础教育阶段的课堂内容有好处, 也对机器学习和运筹优化算法改进有潜在的启发。

参考文献

1. 马文斌, 杨彩琴, 贺艳飞, 刘媛媛, 周兰锁: 数学中基本空间的浅释. 内蒙古农业大学学报 (自然科学版). 32, 360–363 (2011)
2. 程克玲: 向量的 p -范数及向量序列的收敛性研究. 西安文理学院学报 (自然科学版). 21, 6–10 (2018). <https://doi.org/10.3969/j.issn.1008-5564.2018.06.002>
3. 冷岗松: 关于正齐次不等式. 中等数学. 12–13, 18 (2015)
4. 张贤达: 矩阵分析与应用. 清华大学出版社, 北京 (2004)