一个一般的 Lagrange 定理

Giovanni Panti

1. 引言

一个实数通常的连分数展开基于欧几里得算法.后者的变种产生前者的变种,而这所有的变种都被涵盖在更普遍的动力系统框架中.对于所有这些变种,Lagrange (拉格朗日) 定理成立:一个数有一个最终循环的展开式,当且仅当它是一个二次无理数.对于特殊的展开式,这个事实肯定是知道的,但是对于一般情形的证明,我在文献中所能追踪到的,只是作为 Boshernitzan 和 Carroll 关于区间交换变换 [2] 的甚为深刻结果的一个隐含的推论.因而,对于一般的 Lagrange 定理有一个简单的和几乎不需要计算的证明是有用的.

令 \mathcal{D} 是实单位区间的任一 \mathfrak{L} 4 模划分. 这意味着 \mathcal{D} 是一族 (有限或可数,基数至少是 2) 半开区间 $\Delta_a = (p/q, r/s]$, 其中 a 在一个固定的指标集合 $I \subseteq \mathbb{Z}$ 中变动,使得

- (i) 每个区间 $\Delta_a \in \mathcal{D}$ 是 幺模的, 即, 它的端点 $0 \le p/q < r/s \le 1$ 是有理数 (总写成既约形式), 并且 $|\frac{p}{q}|_s^r| = -1$ (这等于说,列向量 $(p \ q)^t$ 和 $(r \ s)^t$ 构成 \mathbb{Z}^2 的一个 \mathbb{Z} 基);
 - (ii) D 中两个不同的区间不相交;
 - (iii) 集合 $X = \bigcup \mathcal{D} \setminus \{1\}$ 包含 [0,1] 中所有的无理数.

固定一个任意函数 $\varepsilon: I \to \{-1, +1\}$, 令 $\Delta_a = (p/q, r/s] \in \mathcal{D}$, 并令

$$G_a = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{(\varepsilon(a)+1)/2} \begin{pmatrix} p & r \\ q & s \end{pmatrix}^{-1}.$$

 $\operatorname{GL}_2\mathbb{Z}$ 中任一矩阵 $T=\begin{pmatrix} t & u \\ v & w \end{pmatrix}$, 通过 $\phi_T(x)=(tx+u)/(vx+w)$, 得到了一个从射影 实直线 $\mathbb{R}\cup\{\infty\}$ 到其自身上的分式线性同胚 ϕ_T . 利用射影坐标 $(x\ y)^t$ 来表示实数 x/y $((1\ 0)^t$ 表示 $\infty)$, ϕ_T 就等于用 T 左乘. 我们把上述应用到 $T=G_a$, 注意到 $G_a(p\ q)^t$ 和 $G_a(r\ s)^t$ 是 $(0\ 1)^t$ 和 $(1\ 1)^t$ (依照该次序或其反序,取决于 $\epsilon(a)$ 的取值). 这样,我们就能 容易地验证: 若 $\epsilon(a)=-1$, 则 $\phi_{G_a}\upharpoonright\Delta_a$ 是单调增的,值域为 (0,1]; 若 $\epsilon(a)=+1$, 则它 是单调减的,值域为 [0,1). 所得到的由对于 $x\in\Delta_a$, $Gx=\phi_{G_a}(x)$ 定义的分段分式映射 $G:X\to[0,1]$ 是由 \mathcal{D} 和 ϵ 确定的 Gauss (高斯) 映射.

用 $\psi_a=\phi_{G_a^{-1}}\upharpoonright [0,1]$ 表示 G 的第 a 个逆分支; ψ_a 的值域是 Δ_a 的拓扑闭包. 计算 G_a^{-1} ,我们显式地得到

译自: The Amer. Math. Monthly, Vol.116 (2009), No.1, p.70–74, A General Lagrange Theorem, Giovanni Panti, figure number 2. Copyright ©the Mathematical Association of America 2009. All rights reserved. Reprinted by permission of Taylor & Francis Ltd on behalf of the Mathematical Association of America.

作者的邮箱地址是 panti@dimi.uniud.it.

$$\psi_a(x) = \begin{cases} \frac{(r-p)x+p}{(s-q)x+q}, & \text{ } \sharp \varepsilon(a) = -1; \\ \frac{(p-r)x+r}{(a-s)x+s}, & \text{ } \sharp \varepsilon(a) = +1. \end{cases}$$

如果 $x, Gx, G^2x, \ldots, G^{n-1}x$ 都在 X 中,则对 $1 \le t \le n$,令 $G^{t-1}x \in \Delta_a$,我们就有恒等式

$$x = \psi_{a_1} \psi_{a_2} \cdots \psi_{a_n} (G^n x). \tag{1}$$

用动力系统的语言, (有限或无限的) 序列 a_1,a_2,\ldots 是 x 的 符号序列. 由 (1), 如果 x 有一个有限的符号序列 (即,对某个 n, $G^nx \in [0,1]\setminus X \subset \mathbb{Q}$), 那么 x 是有理的. 反之,令 $x=u/v\in \Delta_a$ 是有理的,并令 p/q < r/s 是 Δ_a 的端点. 那么对某些满足 $m\geq 1$ 的非负整数 l,m, 有 $(u\ v)^t=l(p\ q)^t+m(r\ s)^t$. 用 G_a 左乘,我们看到 G_x 的分母等于 l+m,它严格小于 x 的分母 v=lq+ms. 因而,沿着 x 的 G 轨道,分母序列是严格减的,因而 x 最终必定离开 X (注意,0 和 1 是仅有的两个分母为 1 的有理点,而它们都不在 X 中).

例 1 第一个和主要的例子理所当然地由通常的连分数给出: $I = \{1, 2, ...\}$, $\Delta_a = (1/(a+1), 1/a]$, X = (0, 1), 以及在此例中 $\varepsilon = +1$. Gauss 映射是 $Gx = 1/x - \lfloor 1/x \rfloor$, 图 1中展示了它的图形.

我们有 $\psi_a(x) = 1/(a+x)$, 因而 (1) 取其熟悉的形状

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{ \cfrac{\ddots}{a_{n-1} + \cfrac{1}{a_n + G^n x}}}}}.$$



图 1

例 2 在奇连分数中, 我们有 I, X 和诸 Δ_a 如上例, 以及 $\varepsilon(a) = (-1)^{a+1}$. Gauss 映射是 Gx = |1/x - b|, 其中 $b \in \{|1/x|, |1/x| + 1\}$ 中的奇整数, 图 2 中展示了它的图形.

对于奇的 $a_t = \lfloor 1/G^{t-1}x \rfloor$, 令 $(b_t, \varepsilon_t) = (a_t, +1)$; 对于偶的 a_t , 令 $(b_t, \varepsilon_t) = (a_t + 1, -1)$, 我们得到 $\psi_{a_t}(x) = 1/(b_t + \varepsilon_t x)$, 因而

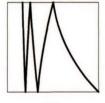


图 2

在被我们的框架所覆盖的连分数算法中,有偶连分数,最近整数分数,Farey (法里)分数,等等;更多的例子,以及相应的 Gauss 映射的遍历性质的详细分析,见 [1,7]. 其它一些算法—— 诸如印度人于 12 世纪为了解 Pell (佩尔) 方程而发现的循环法 (cyclic method) 的各种变形—— 在此并未涉及;有关这些算法请见 [8, §2-3] 及那里的文献.

2. 收敛性 ...

幺模区间有一个奇特的性质. (下文中, length (Γ_n) 表示 Γ_n 的长度 —— 译注)

观察 3 令 $\Gamma_1 \supset \Gamma_2 \supset \cdots$ 是一个闭幺模区间套,它的每一个区间都严格地包含在前一个区间中. 则

$$\lim_{n\to\infty} \operatorname{length}(\Gamma_n) = 0.$$

证明 $\Gamma_n = [p_n/q_n, r_n/s_n]$ 的长度是 $(q_n s_n)^{-1}$, 它小于或等于 $[\max(q_n, s_n)]^{-1}$. 如果 u/v 在 Γ_n 的拓扑内点集中是有理的, 那么 v 严格大于 $\max(q_n, s_n)$ (这又是因为对某些正的 $l, m \in \mathbb{Z}$, 有 $(u \ v)^t = l(p_n \ q_n)^t + m(r_n \ s_n)^t$). 由此即得,序列 $\max(q_1, s_1)$, $\max(q_2, s_2)$, ... 是严格单调增的,并且趋于无穷,因而其倒数序列 —— 它是区间套长度序列的上界 —— 趋于零.

作为推论, 我们有

- (i) 每个把无理数与其符号序列联系起来的映射 $k:[0,1]\setminus\mathbb{Q}\to I^{\mathbb{N}}$, 是单射的, 连续的和开的;
 - (ii) 如果 x 是无理的, 其符号序列为 $a_1, a_2, ...$, 并且 $y \in [0, 1]$, 则

$$\lim_{n\to\infty}\psi_{a_1}\cdots\psi_{a_n}(y)=x.$$

事实上,令 $k(x)=a_1,a_2,\ldots$; 由 $\S1$ (1),那些符号序列与 k(x) 直到 a_n 都一致的数的集合被包含在闭幺模区间 Γ_n 中, Γ_n 的端点是 $\psi_{a_1}\cdots\psi_{a_n}(0)$ 和 $\psi_{a_1}\cdots\psi_{a_n}(1)$. 因为每个 $\Delta_a\in\mathcal{D}$ 是真 (properly) 包含在 (0,1] 中 (即 $\Delta_a\subsetneq(0,1]$ —— 译注),因而包含关系 $\Gamma_1\supset\Gamma_2\supset\cdots$ 是真的,并且观察 3 蕴涵着 (ii) 和 k 的单射性.我们把 k 的连续性和开性作为练习留给读者;k 可以是满射的 (例如,在通常连分数的情形),但是当 I 为有限的时,k 不可能是满射的 (因为此时 $I^{\mathbb{N}}$ 是紧的,而 $[0,1]\setminus\mathbb{Q}$ 是非紧的).

3. ... 和周期性

1769 年 8 月 25 日,Lagrange 在柏林皇家科学院的一份研究报告中讨论了关于通过连分数来解决代数方程的方法 [5]. 自从 1766 年起,Lagrange 就已经在柏林,经其前任 Leonhard Euler (欧拉) 的推荐,出任柏林皇家科学院数学部主任. 在这份研究报告中,他证明了: 无理数 x 最终有一个周期连分数表达式,当且仅当它在有理域上阶为 2. 即使在我们一般的框架中,"仅当"的结论是明显的. 如果 x 在任何 Gauss 映射 G 下有符号序列 $a_1,\ldots,a_t,\overline{a_{t+1},\ldots,a_{t+r}}$,那么 G^tx 和 $G^{t+r}x$ 有相同的符号序列,即 $\overline{a_{t+1},\ldots,a_{t+r}}$,因而它们是相等的. 这就蕴涵着向量

$$G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}$$
 $\forall a G_{a_{t+r}} \cdots G_{a_{t+1}} G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}$

是射影相等的,即,相差一个非零乘法常数. 然而此时 $(x\ 1)^t$ 是整数矩阵 $(G_{a_t}\cdots G_{a_1})^{-1}$ $G_{a_{t+r}}\cdots G_{a_{t+1}}G_{a_t}\cdots G_{a_1}$ 的一个本征向量,因而 x 是二次的.

"当" 部分较棘手. 通常的证明 [4,6] 基本上是 Lagrange 的方法. 我们考虑 $1/G^nx$ 的极小多项式 $c_nX^2+d_nX+e_n\in\mathbb{Z}[X]$, 并来证明: 从某个 n 起,系数 c_n,d_n,e_n 必须满足某些不等式. 有时,这个证明把 $1/G^nx$ 改写成既约形式

$$\frac{1}{G^n x} = \frac{P_n + \sqrt{D}}{Q_n},$$

并且,对于足够大的 n,利用上述诸多项式的公共判别式 D 来作为 P_n , Q_n 的界. 所有这些都与简化二次形式的 Gauss 理论密切相关,见 [3, §5.7].

还不清楚,如何在任意分段分式展开的情形中应用上述证明.我们转到更几何一些的看法来叙述另一个观察.

观察 4 令 x 是一个二次无理数. 那么存在一个以 $(x\ 1)^t$ 为一个本征向量,相应的本征值 $\lambda > \overline{\lambda} > 0$ ($\overline{\lambda}$ 是另一本征值) 的整数矩阵 H.

证明 令 $cX^2 + dX + e \in \mathbb{Z}[X]$ 是 x 的极小多项式,并令 \overline{x} 是其代数共轭.显然

$$\begin{pmatrix} -d & -e \\ c & 0 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = cx \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

令 K 是上面的矩阵与 $t imes (\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$ 之和. 对于充分大的正整数 t, K 的本征值 cx + t 和 $c\overline{x} + t$ 都是正的. 如果 $c\overline{x} + t < cx + t$, 我们取 H = K, 即得证. 否则,我们取 $H = |K|K^{-1}$, 并注意到 $0 < |K|(c\overline{x} + t)^{-1} = cx + t < c\overline{x} + t = |K|(cx + t)^{-1}$.

现在我们可以对任意分段分式展开来证明 Lagrange 定理了. 令 $x \in [0,1]$ 是一个二次 无理数,它在某个 Gauss 映射 G 下的符号序列为 a_1,a_2,\ldots ,并令 $H,\lambda,\overline{\lambda},\overline{x}$ 如上所述. 令 $h:\mathbb{R}^2 \to \mathbb{R}^2$ 是关于标准基其矩阵为 H 的线性变换. 由 $\S1$ (1), 对于每个 $n \geq 0$, x 属于端点 为 $\psi_{a_1}\cdots\psi_{a_n}(0)$ 和 $\psi_{a_1}\cdots\psi_{a_n}(1)$ 的区间 Γ_n . 向量 $(\psi_{a_1}\cdots\psi_{a_n}(0)\ 1)^t$ 和 $(\psi_{a_1}\cdots\psi_{a_n}(1)\ 1)^t$ 正比于矩阵

$$B_n = G_{a_1}^{-1} G_{a_2}^{-1} \cdots G_{a_n}^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

的列. 令 C_n 是由这些列正张成的锥. 由 $\S 2$ (ii), 存在 m, 使得对每个 $n \geq m$ 有 $\overline{x} \notin \Gamma_n$. 因为 $0 < \overline{\lambda} < \lambda$, $(x \ 1)^t \in C_n$, 以及 $(\overline{x} \ 1)^t \notin C_n \cup -C_n$, 再把 C_n 中的任一向量写成本征向量 $(x \ 1)^t$ 和 $(\overline{x} \ 1)^t$ 的线性组合,那么我们容易看到,对 $n \geq m$, $h[C_n] \setminus \{0\}$ 被包含在 C_n 的拓扑内点集中. 这蕴涵着,矩阵 $H_n = B_n^{-1}HB_n$ (即,h 关于由 B_n 的诸列所给出的基的矩阵) 当 $n \geq m$ 时的元素为正的. 因而 H_m, H_{m+1}, \ldots 是正整数矩阵,在 $GL_2\mathbb{Z}$ 中都是相互共轭的. 因为矩阵的行列式和迹在共轭下是不变的,所以这些矩阵的数量是有限的. 因而,对于某个 $t \geq m$ 和 t > 0,有 $H_t = H_{t+r}$. H_t 的 λ 本征空间是一维的,其基为 $B_t^{-1}(x \ 1)^t$,对 H_{t+r} 是类似的. 用 $\binom{0}{1}$ 力 左乘,我们看到,向量

$$G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}$$
 $n G_{a_{t+r}} \cdots G_{a_{t+1}} G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}$

是射影等价的. 这意味着 $G^t x = G^{t+r} x$; 因而在 Gauss 映射下 x 是预周期的 (preperiodic), 并且其符号序列最终是周期的.

参考文献

- V. Baladi and B. Vallée, Euclidean algorithms are Gaussian, J. Number Theory 110 (2005) 331–386.
- [2] M. D. Boshernitzan and C. R. Carroll, An extension of Lagrange's theorem to interval exchange transformations over quadratic fields, J. Anal. Math. 72 (1997) 21–44.
- [3] H. Cohen, A Course in Computational Algebraic Number Theory, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, Berlin, 1993. (下转 271 页)

- [5] Y. L.i and S. Dù, Chinese mathematics (Translated from the Chinese by J. N. Crossley and A. W.-C. Lun), Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1987. MR932966 (90m:01003).
- [6] J.-C. Martzloff, Histoire des mathématiques chinoise, Masson, Paris, 1987, A History of Chinese Mathematics (Translated by S. S. Wilson). Springer, Berlin, 1997.
- [7] C. Meyer, Matrix an alysis and applied linear algebra, SIAM, Philadelphia, 2000. MR1777382.
- [8] K. Chemla and G. Shuchun, Les neuf chapitres, Dunod, Paris, 2005. MR2111394 (2005h:01004).
- [9] K. Shen, J. N. Crossley, and A. W.-C. Lun, The nine chapters of the mathematical art companion and commentary, Oxford Univ. Press, New York, 1999. MR1740507 (2001g:01012).
- [10] J. Brennan, NSF proposal preparation, Notices Amer. Math. Soc. 54 (2007), no. 9, 1153–1157.
- [11] D. A. Grier, When computers were humans, Princeton Univ. Press, 2005.
- [12] J. Buteo, Logistica, Paris, 1560.
- [13] A. Heeffer, From the second unknown to the symbolic equation, Philosophical Aspects of Symbolic Reasoning in Early Modern Mathematics (A. Heeffer and M. Van Dyck, eds.), College Publications, London, 2011, pp. 57–101.
- [14] J. F. Grear, How ordinary elimination became Gaussian elimination, Historia Math. 38 (2011), no. 2, 163–218.
- [15] ——, Mathematicians of Gaussian ellimination, Notices Amer. Math. Soc. 58 (2011), no. 6.
- [16] I. Newton, Universal arithmetick, Senex, Taylor, et al., London, 1720.
- [17] F. Chiò, Mémoire sur les fonctions connues sous le nom de résultantes ou de déterminants, Turin, 1853.
- [18] L. E. Fuller and J. D. Logan, On the evaluation of determinants by Chiò's method, College Math. J. 6 (1975), no. 1, 8–10.
- [19] E. H. Bareiss, Sylvester's identity and multistep integer-preserving Gaussian elimination, Math. Comp. 22 (1968), no. 103, 565–578. MR0226829 (37:2416).
- [20] S. C. Althoen and R. McLaughlin, Gauss-Jordan Reduction: A Brief History, Amer. Math. Monthly 94 (1987), 130–142. MR874013 (88d:01017).
- [21] V. J. Katz, Who is the Jordan of Gauss-Jordan?, Math. Mag. 61 (1988), no. 2, 99–100. MR934825 (89f:01034).

	(赵振江	译	陆柱家	校)
*************	******	*****	******	***

(上接 282 页)

- [4] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Number, 5th ed., Oxford University Press, Oxford, 1985.
- [5] J. L. Lagrange, Additions au Mémorie sur la Résolution des Équations Numerériques, Mémoires de l'Académie royale des Sciences et Belles-Letters de Berlin XXIV (1770) 581-652; Also availabe at the Göttingen Digitalisierungszentrum, http://www.gdz-cms.de.
- [6] A. M. Rockett and P. Szüsz, Continued Fractions, World Sciendific, River Edge, NJ, 1992.

(下转 278 页)

有很多有用的构造使我们能把一类组合对象转换成另一类. 例如, 给出一个有向图, 我们可以通过抹去 (有向) 边的定向而把它当成通常图, 又如, 给出一个 3-图连同其一个顶点, 则我们可以看其链接 (link), 它又是一个通常图. 此类构造可通过一种解释的逻辑概念来把握, 而在旗代数语言中, 它们则对不同理论导出 A^{σ} 间的代数同态. 这使得我们可在定理 (即形如 $f \geq 0$ 的陈述) 之间作内容上的合适转换.

空间 $\operatorname{Hom}^+(\mathcal{A}^0,\mathbb{R})$ 是紧的,由此可知每个极值问题有一个 特有的 最优解. 可以从 多种途径来探知其极值性: 例如,可以根据某些直观改变来写出泛函导数 (的类似物),并 使其等于零而得到新的,且并非任一 $\phi \in \operatorname{Hom}^+(\mathcal{A}^0,\mathbb{R})$ 都具有的有用关系.

密度 $\phi \in \text{Hom}^+(A^0,\mathbb{R})$ 的多元组编码了关于研究对象的巨量信息. 例如,设我们随机 "捡取"一个基顶点 (或更一般地,一个更复杂的类型 σ 的复本),则"应当"引出在 $\phi \in \text{Hom}^+(A^0,\mathbb{R})$ 上的一个概率分布. 结果是这个信息已经可以从 ϕ 中获取,即使我们并不确实了解我们的样本空间. 例如,我们可以只用 ϕ 来直接确定图中相对度至少为 $\frac{1}{3}$ 的那部分顶点 (譬如说,可不必使用 Szemerédi (斯泽梅雷迪) 的正则性引理).

像在其他领域那样进行公理化处理,使得可以使用 Cauchy-Schwarz (柯西 - 施瓦茨) 不等式的想法令人神往. 其困难是更技术性和实际的: 对称群的内涵不如其他情况那样丰富,使得我们很难找出不但在理论上可行,而且又能得到新的具体结果的运算. 不过在这方面仍然已经有一些工作 (见 [1] §4.2).

最后,但肯定不是最不重要的一点是,我们常常不仅要关注极限密度 ϕ 本身的性质,还要关注这些密度所对应的 实际 极限对象 (或者更合理地说,还要关注所结合的 模型理论 (model theory)). 这导致深刻和优美的 图极限 的理论,它和其他学科有很多关联,对此我们高度推荐 Lovaśz (罗瓦兹) 的近著 [3] 作为此理论的导引.我们强烈地感到,强调对同一类对象所进行的句法学上 (旗代数) 和语义学上 (图极限) 两种研究方法之间的关联,将对两方面都大有裨益.

参考文献

- [1] A. Razborov, Flag algebra: an interim report, 2013. To appear in the Erdős Centennial Volume, preliminary text available at http://people.cs.uchicago.edu/~razborov/surveys.html/flag_survey.pdf.
- [2] —, Flag algebra. Journal of Symbolic Logic, 72 (4):1239–1282, 2007.
- [3] L. Lovász, Large Networks and Graph Limits. American Mathematical Society, 2012.

(李乔 译 陆柱家 校)

(上接 271 页)

- [7] B. Vallée, Euclidean dynamics, Discrete Contin. Dyn. Syst. 15 (2006) 281–352.
- [8] H. C. Williams, Solving the Pell equation, in Number Theory for the Millenuium, III, M. A. Bennett et al., eds., A K Peters, Natick, MA, 2002, 397–435.

(陆柱家 译 陈凌宇 校)