arXiv:2603.17188v1 [math.DS] 17 Mar 2026

有理语言的序列密度

Alexi Block Gorman 和 Dominique Perrin

2026年3月19日

摘要

我们引入了相对于概率测度序列的有理语言密度这一概念。我们证明，如果 (μ_n) 是一个收敛于正伯努利测度 μ 的伯努利测度序列，则序列密度是相对于 μ的通常密度。我们还证明，如果 (μ_n) 是一个以强意义收敛于不变概率测度 μ 的不变概率测度序列，则每个有理语言相对于该序列的序列密度都存在。

1 引言

关于有理语言相对于伯努利测度的密度，人们已经进行了广泛研究（参见 [3]中的参考文献）。最近有人证明，在不变概率测度这一更一般的情形下，该密度仍然存在 [4]。在 [7], 中，研究了伯努利测度情形的另一种推广，即有理语言相对于伯努利测度序列的密度。 [7]证明了该密度总是存在。

在本文中，我们引入有理语言相对于概率测度序列的序列密度这一概念，从而考虑了一个比 [4] 和 [7]都更一般的情形。我们证明了如下结果：当伯努利测度序列收敛于一个正测度时，序列密度具有明确的描述（定理 4.1）。这也为该情形下的Lynch的结果提供了一个直接证明。

我们还证明，每个有理语-言相对于一个强收敛于不变概率测度的不变概率测度序列 μ 的序列密度都存在（定理 5.1）。这推广了 [7]的主要结果，在该结果中只考虑了伯努利测度。证明使用了 [4]的主要结果。

本文的组织结构如下。第2节介绍本文所用的符号动力系统和概率论基本概念。第3节定义序列密度。我们在第4节证明定理 4.1，并在第5节证明定理 5.1。第6节给出一些例子，说明在哪些情况下这种密度可能不存在，或者虽然存在，却与该序列收敛到的极限测度的密度不一致。

最后，我们在第7节考察组合密度这一概念，它是序列密度的一种特殊情形。

2 预备知识

我们简要介绍符号动力系统中的一些关键概念（更多详情请参见 [6] ）。

2.1 移位空间

设 A 为有限集。我们用 A^* 表示字母表 A 上的词集，用 A^+ 表示非空词集。前缀码是 A^* 的一个子集，其中不包含其任何元素的真前缀。类似地，后缀码是 A^* 的一个子集，其中不包含其任何元素的真后缀。

我们将 $A^{\mathbb{Z}}$ （由 A 中的元素构成的双向无限序列集）视为一个拓扑空间，其拓扑由距离函数$d(x,\ y) = 2^{-r(x,y)}$诱导，其中 $r(x,\ y) = \min\{|n| \mid n \in \mathbb{Z}, x_n \neq y_n\}$。对于 $x \in A^{\mathbb{Z}}$和$i < j$，我们记$x_{[i,j)} = x_i \cdots x_{j-1}$。对于 $w \in A^*$，我们用 $[w] = \{x \in A^{\mathbb{Z}} \mid x_{[0,|w|)} = w\}$ 表示以 w为基的柱集。

在 A 上的一个移位空间是 $A^{\mathbb{Z}}$的一个闭且移位不变的子集。等价地，X 是一个移位空间，当且仅当存在 A上的一组词 W ，称为禁块集合，使得 X 是由不含 W中任何词的序列 x 构成的集合。

X的语言记作 $\mathcal{L}(X)$，是指在 X的某个元素中出现的词 w 的集合。我们用 $\mathcal{L}_k(X)$表示 $\mathcal{L}(X)$中长度为 k 的词的集合。有限型移位是由有限个禁块定义的移位空间。若对每个 $u, w \in \mathcal{L}(X)$，都存在 w 使得 $uwv \in \mathcal{L}(X)$，则称移位空间 X 是不可约的。

设 X 是一个移位空间， $k \geq 1$ 是一个整数。设 $u \mapsto \langle u \rangle$ 是从集合 $\mathcal{L}_k(X)$到字母表 A_k的双射。令 $\varphi_k\colon\ X \to A_k^{\mathbb{Z}}$ 为如下定义的映射：若对每个 $n \in \mathbb{Z}$都有$y_n = x_{[n,n+k)}$ ，则 $y = \varphi_k(x)$。 X 在 φ_k 下的像 $\varphi_k(X)$是 A_k 上的一个移位空间，称为 X的 k块表示。

2.2 有理语言

字母表 A上的语言是 A^*的一个子集。字母表 A 上的有限自动机 $\mathcal{A} = (Q,\ i,\ T)$ 由一个映射 $(q,\ a) \mapsto q \cdot a$ 给出，该映射从有限集 $Q \times A$映射到 Q。这个映射以自然的方式扩展为 A在 Q上的右作用。自动机 $\mathcal{A}$ 所识别的语言是集合

$$L = \{w \in A^* \mid i \cdot w \in T\}.$$

有理语言是指能够被某个有限自动机识别的语言。

语言 L 若存在一个同态 $\varphi\colon\ A^* \to M$ 从 A^* 到幺半群M，使得 $L = \varphi^{-1}(N)$ 对某个 $N \subset M$成立，则称其被该同态识别。一个语言可识别，当且仅当它能被映射到有限幺半群上的同态识别。

有限型移位的语言是有理语言。作为一种等价定义，如果一个语言可以通过有限次并集、乘积（连接）和（Kleene）星运算，由A^* 的有限子集得到，那么它就是有理语言，其中语言 L 的星运算定义为

$$L^* = \{w_1 w_2 \cdots w_n \mid w_i \in L, n \geq 0\}.$$

例如，字母表 $\{a, b\}$ 上以 ab 开头的词所构成的语言 $L = ab\{a, b\}^*$ 是有理语言。

有限幺半群 M 是无周期的，如果存在整数 $n \geq 1$，使得对每个 $m \in M$，都有 $m^n = m^{n+1}$。如果一个有理语言可以通过到某个有限无周期幺半群的同态识别，那么它就是无周期的。例如，语言 $(a^2)^*$是有理语言，但可以证明它不是无周期的。

有理语言可以用带加法的整数的单子二阶逻辑定义，而无周期语言则可以用相应的一阶逻辑定义（见 [8]）。

2.3 概率测度

我们介绍移位空间上的概率测度（更完整的介绍请参见 [5] ）。

在 $A^{\mathbb{Z}}$ 上的概率测度 是一个映射 μ，它为每个博雷尔子集 $U \subseteq A^{\mathbb{Z}}$ 赋予一个实数 $\mu(U)$，并且满足 σ可加性 （也就是说，对于任意一列 (U_n) 两两不相交的博雷尔集 U_n，都有 $\mu(\cup_{n\geq 0} U_n) = \sum_{n\geq 0} \mu(U_n)$），且 $\mu(X) = 1$。

由 σ-可加性可知，测度是单调的；也就是说，$U \subset V$ 蕴含 $\mu(U) \leq \mu(V)$，并且若 $U_1 \subset U_2 \subset \ldots$，则 $\mu(\cup_{i\geq 1} U_i) = \lim_{i\to\infty} \mu(U_i)$。对于 $w \in A^*$，我们记 $\mu(w) = \mu([w])$。

对于每个 $n \geq 0$，由 $\mu_n(w) = \mu(w)$ 定义的映射 $\mu_n\colon\ A^n \to [0, 1]$ 是 A^n 上的概率测度，称为 μ 的第 n 个边缘分布。对于集合$W \subset A^*$，我们用 $\mu(W)$ 表示总和 $\sum_{w\in W} \mu(w)$，它是一个实数或$+\infty$。

概率测度 μ 在 $A^{\mathbb{Z}}$ 上是不变的，如果对于每个博雷尔集 U，都有 $\mu(S^{-1}U) = \mu(U)$。若一个不变概率测度对于每个在移位下不变的博雷尔集U，都有 $\mu(U) = 0$或 $\mu(U) = 1$，则称其为遍历的。根据伯克霍夫遍历定理，不变测度 μ 是遍历的，当且仅当

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(U \cap S^{-i}V) = \mu(U)\mu(V) \tag{2.1}$$

对于每一对 U、V 博雷尔集。

两个不同的遍历测度 μ, ν 在 $A^{\mathbb{Z}}$ 上互相奇异，也就是说，存在一个博雷尔集 U，使得 $\mu(U) = \nu(X \setminus U) = 0$。事实上，设 f是一个可测函数，使得 $\int f d\mu \neq \int f d\nu$。令 U 为满足 $\frac{1}{n}\sum_{i=0}^{n-1} f(x)$ 收敛到 $\int f d\nu$ 的 $x \in A^{\mathbb{Z}}$ 的集合。则 $\mu(U) = \nu(X \setminus U) = 0$。

如果不取平均值时式（2.1）成立，即如果

$$\lim_{n\to\infty}\mu(U\cap S^{-n}V)=\mu(U)\mu(V) \tag{2.2}$$

对每一对博雷尔集 U， V。

我们将 $A^{\mathbb{Z}}$ 上的概率测度空间视为一个拓扑空间，并赋予弱$*$拓扑；在该拓扑下， $\mu_n\to\mu$若对每个博雷尔集 U，都有 $\mu_n(U)\to\mu(U)$。根据巴拿赫-阿劳格鲁定理，这个空间是紧的。不变测度序列的极限仍是不变测度，而遍历测度序列的极限仍是遍历测度。

若概率测度序列$(\mu_n)_{n\geq1}$ 定义在 $A^{\mathbb{Z}}$ 上以强意义收敛到测度 μ，则对每个 $\varepsilon>0$，都存在 $N\geq1$，使得对每个博雷尔集 U 以及每个 $n\geq N$，都有 $|\mu_n(U)-\mu(U)|\leq\varepsilon$。

伯努利测度是定义在 $A^{\mathbb{Z}}$ 上的概率测度 μ，满足对每个 u， $v\in A^*$，都有 $\mu(uv)=\mu(u)\mu(v)$。伯努利测度是遍历的。若伯努利测度序列$(\mu_n)_{n\geq1}$ 收敛，则其极限也是伯努利测度。事实上，对每个 u， $v\in A^*$，都有 $\mu(uv)=\lim_{n\to\infty}\mu_n(uv)=\lim_{n\to\infty}\mu_n(u)\mu_n(v)=\mu(u)\mu(v)$。

马尔可夫测度是由一个随机 $A\times A$ 矩阵 P（称为其转移矩阵）和一个向量 π（称为其初始向量）定义的测度，其定义为

$$\mu(a_0\cdots a_{n-1})=\pi_{a_0}P_{a_0a_1}\cdots P_{a_{n-2}a_{n-1}}$$

对于 a_0、.....、$a_{n-1}\in A$。如果 $\pi P=\pi$，则该测度是不变的。由不可约矩阵 P 定义的不变马尔可夫测度是遍历的。

定义在 $A^{\mathbb{Z}}$ 上的 k步马尔可夫测度是满足以下条件的测度 μ：对于每个博雷尔集 U，都有 $\mu(U)=\nu(\varphi_k(U))$；并且存在定义在 k 块表示 $\varphi_k(A^{\mathbb{Z}})$ 上的某个马尔可夫测度 ν。

定义在 $A^{\mathbb{Z}}$ 上的概率测度 μ 的支撑是由满足 $\mu(w)>0$ 的词 w 构成的集合。对于每个不可约有限型移位，都存在一个支撑在 X 上的遍历测度，称为其最大熵测度。

设 M 是一个不可约矩阵，其对应于主特征值 λ 的左特征向量为 v，右特征向量为 w，且满足归一化条件 $v\cdot w=1$。由 M 定义的有限型移位上的最大熵测度 μ，是初始向量为 $\pi_i=v_iw_i$、转移矩阵为 $P_{ij}=(w_j/\lambda w_i)M_{i,\,j}$ 的马尔可夫测度（见第 13] 章 [6, ）。

例2.1 令 X 为由矩阵定义的有限型移位

$$M=\begin{bmatrix}1&0&1\\1&1&1\\1&1&1\end{bmatrix}.$$

其主特征值为 $\lambda=(3+\sqrt{5})/2$。对应的左、右特征向量 v 和 w 经 $v\cdot w=1$ 归一化后为

$$v=\frac{1}{2\lambda-3}\begin{bmatrix}1&\lambda-2&1\end{bmatrix},\quad w=\begin{bmatrix}\lambda-2\\1\\1\end{bmatrix}.$$

因此， μ_2 是由以下方式定义的马尔可夫测度

$$\pi = \frac{1}{2\lambda - 3}\begin{bmatrix}\lambda - 2 & \lambda - 2 & 1\end{bmatrix}, \quad P = \begin{bmatrix}\frac{1}{\lambda} & 0 & \frac{1}{\lambda-1} \\ \frac{\lambda-2}{\lambda} & \frac{1}{\lambda} & \frac{1}{\lambda} \\ \frac{\lambda-2}{\lambda} & \frac{1}{\lambda} & \frac{1}{\lambda}\end{bmatrix}.$$

3 序列密度

设 $\mu = (\mu_n)_{n\geq 0}$ 为 $A^{\mathbb{Z}}$上的一个概率测度序列。给定字母表 A上的语言L ， L 相对于 μ 的序列密度 为

$$\delta_\mu(L) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mu_i(L \cap A^i)$$

当极限存在时。若极限 $\lim_{n\to\infty} \mu_n(L \cap A^n)$ 存在，而不仅仅是其平均值存在，我们就称 L在强意义下具有密度。当序列（μ_n）为常数时，我们得到通常的密度概念（参见 [3]）。

下面列举序列密度的一些基本性质，它们实际上与普通密度的性质相同。若 L 相对于 μ具有序列密度，则 $0 \leq \delta_\mu(L) \leq 1$。

序列密度是有限可加的；也就是说，若 L, L' 是互不相交且具有序列密度（分别在强意义下具有序列密度）的语言，则 $L \cup L'$ 具有序列密度（分别在强意义下具有序列密度）。

序列密度是单调的，也就是说，若 L 和 L' 具有序列密度（分别在强意义下具有序列密度），且 $L \subset L'$，则 $\delta_\mu(L) \leq \delta_\mu(L')$。

以下结果来自 [7]。

定理 3.1 若 μ 是定义在 $A^{\mathbb{Z}}$上的伯努利测度的收敛序列，则每个有理语言的序列密度都存在。若 L是无周期的，则该密度在强意义下存在。

这一来自 [7] 的结果是针对 $A = \{a, b\}$陈述的。但这实际上并不构成限制，因为每个字母表都可以用两个字母进行编码。此外，假设 $\mu_n(a)$ 不收敛于 1。交换 a 和 b 后，我们总能满足这一假设。最后，还对序列 $\mu_n(a)$ 的收敛规定了各种假设。各种情形（将在下面的例子中出现）涵盖了所有可能性，因此无需在陈述中提及这些情形。

我们首先给出一个说明Lynch定理的例子。对于两个正实函数 f 和 g，若对每个 $\epsilon > 0$ 都存在 $n_0 \in \mathbb{N}$，使得对每个 $n \in \mathbb{N}_{\geq n_0}$都有$f(n) \leq \epsilon g(n)$，我们记作 $f \ll g$。若 $f(n)/g(n)$ 收敛于 1，我们记作 $f \sim g$。

例 3.2 令 $A = \{a, b\}$ 和 $L = A^*aA^*$。我们有

$$\delta_\mu(L) = \begin{cases} 0 & \text{if } \mu_n(a) \ll n^{-1}, \\ 1 - e^{-c} & \text{if } \mu_n(a) \sim cn^{-1}, \\ 1 & \text{if } n^{-1} \ll \mu_n(a). \end{cases} \tag{3.1}$$

我们将在下面的例4.2中得到这些值。请注意，在第二种情况下，序列密度存在，但不等于 δ_μ。事实上，在这种情况下，$\mu(a) = 0$，因此 $\delta_\mu(L) \neq \delta_\mu(L) = 0$.

下面的例子说明极限仅在平均意义下存在的情形。

例3.3 令 $L = (A^2)^*$ 为偶数长度的词的集合。那么，

$$\mu_i(L \cap A^i) = \begin{cases} 1 & \text{if } i \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

因此，序列 $\mu_i(L \cap A^i)$ 不收敛。当然，对于任意概率测度序列 μ，都有 $\delta_\mu(L) = 1/2$。

4 序列密度的一个公式

下面的结果涵盖了 μ收敛到正测度时的定理 3.1。它还证明了一个更强的性质，因为我们证明序列密度不仅存在，而且等于相对于极限测度 μ的密度（我们已在例 3.2 中看到，如果 μ 不是正测度，这一结论可能不成立）。我们给出一个直接证明。

定理 4.1 设 $\mu = (\mu_n)_{n \geq 0}$ 是 $A^{\mathbb{Z}}$ 上的伯努利测度序列。若μ 收敛到正伯努利测度 μ，则对于每个有理语言$L \subset A^*$，L 的序列密度等于其相对于 μ的密度。对于无周期语言 L，强意义下的密度也同样成立。

证明。设 $\mathcal{A} = (Q, i, T)$ 是一个识别 L 的确定性自动机。由于 L 是由 $(Q, i, \{t\})$ 在 $t \in T$ 下识别的语言的有限并，我们可以假设 $T = \{t\}$。不失一般性，我们还可以假设每个 $q \in Q$ 都位于一条从 i 到 t 的路径上。对于 $A^{\mathbb{Z}}$ 上的伯努利测度 ν，令 $P(\nu)$ 为由下式定义的随机 $Q \times Q$ 矩阵：

$$P(\nu)_{p,q} = \sum_{a \in A, p \cdot a = q} \nu(a). \tag{4.1}$$

由于 μ 为正，矩阵 $P(\mu)$ 不可约。由于 μ 收敛到 μ，我们可以假设所有矩阵 $P_n = P(\mu_n)$ 的周期 p 都等于 $P(\mu)$ 的周期。我们可以假设 $p = 1$，因此每个 P_n 都是本原的。

令 $\lambda_{n,k}$, $1 \leq k \leq d$ 为 P_n 的不同特征值，并令 $\gamma_{n,k}$ 为 $\lambda_{n,k}$ 在 P_n 的最小多项式中的重数。我们可以假设 $\lambda_{n,1} = 1$。由于 P_n 是本原的，因此有 $\gamma_{n,1} = 1$。

利用 P_n 的约当标准形，我们可以写成

$$P_n = U \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & J_{n,2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J_{n,d} \end{bmatrix} U^{-1}$$

其中，对于 $2 \le k \le d$, $J_{n,\,k}$ 是 P_n 对应于 $\lambda_{n,\,k}$ 的若尔当块。矩阵 U 的列是 P_n 的广义特征向量。特别地，第一列是向量 v，其所有分量均为 1；第一行是随机向量 w_n，满足 $w_n P_n = w_n$。由于 P_n 收敛到 $P(\mu)$，向量 w_n 趋于随机向量 w，满足 $wP(\mu) = w$。

由于 $P(\overline{\mu})$ 是本原的，存在 $\epsilon > 0$，使得所有特征值 $\lambda_{n,k}$ 对于 $k \ge 2$都满足 $|\lambda_{n,k}| < 1 - \epsilon$。因此，每个 $J^n_{n,k}$ 都趋于 0，从而，

$$\lim_{n\to\infty} P_n^n = \lim_{n\to\infty} vw_n = \lim P(\overline{\mu})^n.$$

这意味着，对于每个有理语言 L，都有 $\delta_\mu(L) = \delta_\mu(L)$。 ∎

例4.2 设 $A = \{a, b\}$ 和 $L = A^* a A^*$。自动机 $\mathcal{A}$ 如图4.1所示。矩阵 P_n 为

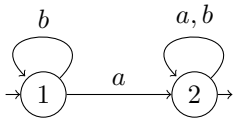

图4.1：自动机 $\mathcal{A}$。

$$P_n = \begin{bmatrix} q_n & p_n \\ 0 & 1 \end{bmatrix}$$

并且我们有

$$P_n^n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q_n \end{bmatrix}^n \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + q_n^n \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

因此 $\delta_\mu(L) = 1 - \lim_{n\to\infty} q_n^n$。

这给出了公式（3.1）中所示的 $\delta_\mu(L)$ 的值。

对于单个伯努利测度，以下结果归功于 Sch¨utzenberger[9]（另见 [2, 定理 13.4.5]）。

该表述用到了有限幺半群的基本概念（见 [2]）。特别地，有限幺半群有唯一的极小理想，它是 M 的极小右理想之并，也是极小左理想之并。对于每个极小右理想 R 和每个极小左理想 L，它们的交集$R \cap L$都是一个有限群。这些群彼此同构。如果 M 是无周期的，那么G是平凡的。

推论 4.3 设 $\mu = (\mu_n)_{n\ge 0}$ 是定义在 $A^{\mathbb{Z}}$ 上、收敛于正伯努利测度 μ 的一列伯努利测度。设 $\varphi\colon A^* \to M$ 是一个到有限幺半群的满同态。设 K 是 M 的极小理想。令 $\nu = \delta_\mu \circ \varphi^{-1}$。那么，对每个 $m \in M$，都有以下结论。

1. $\nu(m) > 0$ 当且仅当 $m \in K$。

2. 对于每个 $m \in K$，都有

$$\nu(m) = \frac{\nu(mM)\nu(Mm)}{d}, \tag{4.2}$$

其中 $d =$ 基数$(mM \cap Mm)$。

如果 M 是无周期的，那么 $d = 1$ 和 $\delta_\mu(L)$ 在强意义下存在。

证明。根据定理 4.1，我们有 $\delta_\mu = \delta_\mu$。由于上述结果对正伯努利测度成立，故结论得证。 ∎

例 4.4 我们再次考虑例 3.2 和 4.2 中的 $L = A^*aA^*$。当 μ 为正时，便会出现 $n^{-1} \ll p_n$。数值 $\delta_\mu(L) = 1$ 与推论 4.3 一致。事实上，在这种情况下，幺半群 M 有两个元素 0,1，其中 $\varphi(a) = 0$ 且 $\varphi(b) = 1$。我们有 $K = \{0\}$ 且 $\delta(L) = \nu(K) = 1$。

In the cases $\mu_n(a) \ll n^{-1}$ or $\mu_n(a) \sim cn^{-1}$, we have $\nu(K) < 1$.
μ

定理 4.1 可以推广到一列马尔可夫测度收敛到一个正马尔可夫测度的情形。事实上，设 ν 是由初始向量 v 和随机 $A \times A$-矩阵 M定义的马尔可夫测度。那么，我们将方程 4.1 中的矩阵 $P(\nu)$替换为 $(Q \times A) \times (Q \times A)$矩阵

$$P(\nu)_{(p,a),(p\cdot a,b)} = M_{a,b}.$$

5 强意义下的收敛

定理 5.1 设 $\mu = (\mu_n)^{n\geq 0}$ 是一个不变概率测度序列，并在强意义下收敛于一个不变概率测度μ。那么，对于每个有理语言 L，L 的序列密度存在，且 $\underline{\delta_\mu}(L) = \overline{\delta_\mu}(L)$。如果 L是无周期的，则序列密度在强意义下存在。

我们使用以下结果，该结果已在 [4]中得到证明。

定理 5.2 设 μ 是不变概率测度。每个有理语言 L 相对于 μ都有密度。如果 L 是无周期的，则该密度在强意义下存在。

我们还将使用以下实分析中的基本结果（见 [1, 第 8.20]节）。若对每个 $\varepsilon > 0$，都存在 $N \in \mathbb{N}$ 使得当 $m, n \geq N$时有$|u_{m,n} - a| \leq \varepsilon$，则称实数序列 $(u_{m,n})$ 收敛于 a。如果该序列收敛于 a，那么极限 $\lim_{m\to\infty} \lim_{n\to\infty} u_{m,n}$ 和 $\lim_{n\to\infty} \lim_{m\to\infty} u_{m,n}$ 也都收敛于 a。若再加上一致收敛这一条件，逆命题也成立。

引理 5.3 设 $(u_{m,n})_{m,n\geq 0}$ 是一个实数序列。如果

(i) $\lim_{m\to\infty}\lim_{n\to\infty} u_{m,n} = a$, (ii) $\lim_{n\to\infty} u_{m,n}$ 在 m 上一致存在，

然后 $(u_{m,\,n})$ 收敛到 a。

定理 5.1 的证明。令

$$u_{m,n} = \frac{1}{m}\sum_{i=0}^{m-1} \mu_n(L \cap A^i).$$

根据定理 5.2，我们有

$$\lim_{m\to\infty}\lim_{n\to\infty} u_{m,n} = \lim_{m\to\infty}\frac{1}{m}\sum_{i=0}^{m-1}\overline{\mu}(L\cap A^i) = \delta_{\overline{\mu}}(L).$$

另一方面，

$$\lim_{n\to\infty} u_{m,n} = \lim_{n\to\infty}\frac{1}{m}\sum_{i=0}^{m-1}\mu_n(L\cap A^i) = \frac{1}{m}\sum_{i=0}^{m-1}\overline{\mu}(L\cap A^i),$$

由于序列 (μ_n) 强收敛，因此在 m 中一致成立。因此，根据引理 5.3，序列 $(u_{m,\,n})$ 收敛到 $\delta_\mu(L)$。 –

如果 L 是无周期的，令 $v_{m,\,n} = \mu_n(L \cap A^m)$。根据定理 5.2，我们有

$$\lim_{m\to\infty}\lim_{n\to\infty} v_{m,n} = \lim_{m\to\infty}\overline{\mu}(L\cap A^m) = \delta_{\overline{\mu}}(L),$$

and

$$\lim_{n\to\infty} v_{m,n} = \lim_{m\to\infty}\overline{\mu}(L\cap A^m)$$

关于 m 一致。由引理 5.3，序列 $v_{m,\,n}$ 收敛到 $\delta_\mu(L)$。 – ∎

然而，定理 5.1 的适用范围有限。事实上，遍历测度序列若要在强意义下收敛，就必定最终为常值。确实，设 μ_n 是一个遍历测度序列，并在强意义下收敛到 μ。若对所有 $n \geq 0$ 都有 $\mu_n \neq \mu$，由于测度 μ_n 和 μ 互相奇异，序列 (μ_n) 不可能在强意义下收敛到 μ。 –

6 更一般情形下的密度

我们给出两个有趣的例子。先看一个序列收敛、但不收敛到正测度的例子。

考虑 $A = \{a, b, c\}$ 上的有限型移位，其禁块集合为

$$W_n = \{aA^i a : i \leq n\}$$

并将这个移位称为每个 $n \geq 2$ 对应的 X_n。注意，X_n 是 $A^{\mathbb{Z}}$ 中这样的元素：a 的每次出现之间至少相隔 n 位数字。我们还注意到，$X_n \supsetneq X_{n+1}$

对于每个 n，且每个都是有限型移位。因此，我们可以将 μ_n 定义为 X_n 的支撑上的最大熵测度，并且显然， μ_n 是马尔可夫测度。

移位 $X = \cap_{n\in\mathbb{N}} X_n$ 由至多出现一次字母 a的序列构成。我们考虑有理语言 $L = \{b,c\}^* a \{b,c\}^*$。集合 $L \cap A^n$ 由 A 上长度为 n 且只含一个a的 n 序列组成。因此， $\mu_n(L \cap A^n)$ 趋于 0。

序列 $(\mu_n)_{n\in\mathbb{N}}$ 收敛于$\{b,c\}^{\mathbb{Z}} \bigcup \{b,c\}^{-\mathbb{N}} a \{b,c\}^{\mathbb{N}}$上的最大熵测度。将此测度记为 $\bar{\mu}$。显然， $\bar{\mu}$ 是由 $\bar{\mu}(a) = 0$, $\bar{\mu}(b) = \bar{\mu}(c) = \frac{1}{2}$ 给出的伯努利测度。我们看到 $\delta_\mu(L) = 0$。因此， L的序列密度等于其相对于 $\bar{\mu}$的密度。另一方面，如果我们考虑有理语言 $L' = \mathcal{L}(X) = L \cup \{b,c\}^*$，那么 $0 = \delta_\mu(L') \neq \delta_\mu(L') = 1$。

我们给出第二个例子，说明 Lynch定理对于非马尔可夫测度的概率测度可能不成立。

对于 $n \geq 2$，令 X_n 为定义在 $A = \{a,b,c\}$ 上、具有以下禁块集的有限型移位：

$$W_n = \begin{cases} aA^{n-2}a & \text{if } \lfloor \log_2(n) \rfloor \text{ is even,} \\ aA^{n-2}b & \text{otherwise.} \end{cases}$$

移位 X_n 在图6.1中表示为 $n = 2$。

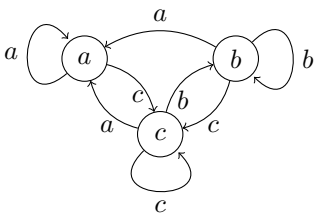

图6.1：移位 X_2。

每个移位 X_n 都是不可约的。事实上， X_n 是图 Γ_n 上的边移位，其顶点集为 A^{n-1}，边则是长度为 n 且不属于 W_n 的词。顶点 a^{n-1} 与 c^{n-1} 相连，而 c^{n-1} 又与所有长度为 $n-1$ 的词相连。

令 μ_n 为 X_n 上的最大熵测度。那么，(μ_n) 收敛到全移位 $\{a,b,c\}^{\mathbb{Z}}$ 上的均匀伯努利分布。测度 μ_2 是例2.1中计算出的测度。

我们声称，序列 (μ_n) 是一列遍历测度，并且收敛到均匀伯努利测度；语言 $L = aA^* \cap A^*a$ 的序列密度不存在。

事实上，每个 μ_n 都是不可约平稳马尔可夫链，因此它是遍历的。由于

序列 (X_n) 收敛到全移位 $A^{\mathbb{Z}}$。接下来，我们有

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu_i(L\cap A^i) < \frac{1}{2}$$

如果 $n = 2^{2k+1}$ 且

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu_i(L\cap A^i) > \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$$

否则。因此，该极限不存在。

7 组合密度

设 X 是字母表 A上的一个移位空间。语言 $L \subset A^*$ 的组合密度为以下极限

$$\delta_X(L) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\frac{\mathrm{Card}(L\cap\mathcal{L}_i(X))}{\mathrm{Card}(\mathcal{L}_i(X)}$$

前提是该极限存在。

组合密度是序列密度的一个特例。对于每个 $n \geq 0$，设 μ_n 是 $A^{\mathbb{Z}}$ 上的一个概率测度，使得其关于 $w \in A^n$ 的 n^{th} 边缘分布如下:

$$\mu_n(w) = \begin{cases} \frac{1}{\mathrm{Card}(\mathcal{L}_n(X)} & \text{if } w \in \mathcal{L}_n(X), \\ 0 & \text{otherwise.} \end{cases}$$

那么，

$$\mu_n(L) = \frac{\mathrm{Card}(L\cap\mathcal{L}_n(X)}{\mathrm{Card}(\mathcal{L}_n(X)}.$$

因此， L 的组合密度等于它相对于序列 (μ_n) 的序列密度。

下文将说明，当右理想的组合密度存在时，它等于相对于 X 上某个概率测度的密度。

命题 7.1 设 X 是一个移位空间，使得对每个 $u \in\mathcal{L}(X)$， $\delta_X(uA^*)$ 都存在。存在一个定义在 X 上的概率测度 μ ，使得对每个 $u \in \mathcal{L}(X)$ 都有

$$\delta_X(uA^*) = \mu(u).$$

证明。 对每个 $u \in A^*$，都有

$$\begin{aligned}\sum_{a\in A}\delta_X(uaA^*) &= \sum_{a\in A}\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\frac{\mathrm{Card}(uaA^*\cap\mathcal{L}_i(X))}{\mathrm{Card}(\mathcal{L}_i(X))}\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\frac{\sum_{a\in A}\mathrm{Card}(uaA^*\cap\mathcal{L}_i(X))}{\mathrm{Card}(\mathcal{L}_i(X))}\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\frac{\mathrm{Card}(uA^*\cap A^i)}{\mathrm{Card}(\mathcal{L}_i(X))}\\ &= \delta_X(uA^*).\end{aligned}$$

由于 $\delta_X(A^*) = 1$，这表明在 X 上存在唯一的博雷尔概率测度，使得对于每个 $u \in \mathcal{L}(X)$，都有 $\mu([u]) = \delta_X(uA^*)$。 ∎

命题 7.2 设 X 是 $A = \{0, 1\}$ 上的一个斯图姆序列移位。对于每个 $u \in \mathcal{L}(X)$，有

$$\delta_X(uA^*) = \mu(u),$$

其中 μ 是 X 上唯一的不变概率测度。

证明。 设 $x \in A^{-\mathbb{N}}$ 是 X^- 中唯一的右特殊序列。整数 $\mathrm{Card}(uA^*\cap\mathcal{L}_n(X)) - 1$ 等于以 u 开头且长度至多为 n 的右特殊词的数量。由于所有右特殊词都是 x 的后缀，它也等于 u 在 $x_{(-n,0]}$ 中出现的次数 $f_n(x, u)$。因此

$$\delta_X(uA^*) = \lim_{n\to\infty}\frac{1 + f_n(x, u)}{n + 1} = \mu(u).$$

∎

7.1 致谢

第一作者获得了大巴黎数学联合资助计划以及玛丽·斯克沃多夫斯卡-居里行动（H2020-MSCA-COFUND-2020-GA101034255）的资助。

参考文献

[1] Tom M. Apostol。《数学分析》。Addison-Wesley Publishing Co.，马萨诸塞州雷丁—伦敦—安大略省唐米尔斯，第二版，1974年。

[2] Jean Berstel、Dominique Perrin 和 Christophe Reutenauer。《编码与自动机》。剑桥大学出版社，2009年。

[3] Val´erieBerth´e、Herman Goulet-Ouellet、Carl-Fredrik NybergBrodda、Do- minique Perrin 和 Karl Petersen。《移位空间中群语言的稠密性》，2024。

[4] Val´erie Berth´e、Herman Goulet-Ouellet 和 Dominique Perrin。移位不变测度下有理语言的密度。载于*ICALP 2025*。Dagstuhl，2025。[5] Mike Boyle 和 Karl Petersen。符号动力系统背景下的隐马尔可夫过程。载于隐马尔可夫过程的熵及其与动力系统的联系：班夫国际研究站研讨会论文集。伦敦数学学会讲义系列，剑桥大学出版社，2011。[6] Douglas Lind 和 Brian Marcus。符号动力系统与编码导论。剑桥大学出版社，剑桥，1995。第二版，2021。[7] James F. Lynch。随机词的收敛律。澳大拉西亚组合学杂志，7：145–156，1993。[8] Dominique Perrin 和 Jean ´Eric Pin。无限词：自动机、半群、逻辑与博弈。爱思唯尔，2004。[9] M. P. Sch¨utzenberger。论某些自由子幺半群。法国数学会通报，93：209–223，1965。