arXiv:2602.08147v1 [math.DS] 8 Feb 2026

稀疏耦合线性余循环的李雅普诺夫指数

Reza Rastegar*

2026年2月10日

摘要

本文研究实矩阵的结构化乘积。通过将动力学约化为上三角矩阵的一种适用范围更广的推广形式，可以求得其最大李雅普诺夫指数。利用预先给定的零模式（包括分块三角结构和稀疏分解，并可借助有向稀疏图方便地编码），我们结合确定性的三角控制与适用于分块三角乘积的 Furstenberg–Kifer 引理的一种恰当改进，得到了明确且可计算的界，并在条件有利时给出公式γ_1 [17] 。这些估计适用于温和的（可能是确定性的）序列，也适用于满足标准可积性条件的平稳遍历随机余循环。我们还讨论了其在线性系统扰动模型中的应用，包括低秩更新；在这些模型中，该约化可将问题转化为较低维的余循环或标量余循环。

MSC2010：37H15, 37M25.

关键词：李雅普诺夫指数、随机三角矩阵、随机分块矩阵、零元素模式、图分解、矩阵扰动。

1 引言

李雅普诺夫指数在随机和混沌动力系统的分析中发挥着核心作用：它们量化了系统对初始条件的敏感性；在线性余循环情形下，则刻画了矩阵乘积的长期增长率；例如参见 [5, 6, 8, 16, 37]。尽管李雅普诺夫指数十分重要，显式公式却仅适用于数量相对有限的可解系综——经典研究中的许多计算集中于 2×2 乘积或高度对称的分布（包括若干高斯型模型）；例如参见[4, 10, 11, 13, 21, 23, 24, 26–30, 34]和 [1–3, 9, 14, 15, 22, 32]。本文旨在从 [4, 17, 24, 34] 中拓展并重整若干思路，形成一套小型工具箱，用于获得结构化矩阵乘积的李雅普诺夫指数的可计算界（在条件有利时，也可求得显式值）。从矩阵分析的角度看，本文的指导原则是：三角或块三角结构——以及更一般地，具有有限反馈的预设稀疏性／零模式——

*华盛顿州伍丁维尔；电子邮箱：reza.j.rastegar@gmail.com

会将指数增长局限于低维对角动力学，而其余元素至多带来一个可明确控制的组合因子。具体而言，我们的界以对角块乘积和描述允许的非对角耦合的有向图为依据，因此只需少量低维范数估计即可轻松计算。该框架既适用于平稳遍历矩阵序列，也适用于满足温和性条件 (1) 的确定性序列，因此可视为一套矩阵分析工具箱，用于结构化或稀疏耦合线性系统的增长率界。

对于整数 $d \geq 2$，令 $\mathcal{M}_d$ 表示实数值元素的 $d \times d$ 矩阵空间。全文中，我们用 $\|\cdot\|$ 表示 $\mathbb{R}^d$ 上的一般范数（具体选取何种范数无关紧要），也表示 $\mathcal{M}_d$ 上诱导的算子范数。令 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$ 为 $\mathcal{M}_d$ 中的一个序列，使得

$$\limsup_{n\to\infty} \frac{1}{n} \log \|X_n\| < \infty \qquad \text{where} \qquad X_n := A_n \cdots A_1. \tag{1}$$

我们为 $\mathcal{A}$ 关联一个（上）李雅普诺夫指数函数 $\gamma\colon \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$，定义为

$$\gamma(v) := \limsup_{n\to\infty} \frac{1}{n} \log \|X_n v\|. \tag{2}$$

尤其 地，我们将序列 $\mathcal{A}$b 的最大（上）李雅普诺夫指数定义为 y

$$\gamma_1(\mathcal{A}) := \limsup_{n\to\infty} \frac{1}{n} \log \|X_n\| = \sup_{\|v\|=1} \gamma(v). \tag{3}$$

这种方向增长率是我们旨在控制的基本对象，所考虑的是下文中的结构化情形。

当 $\mathcal{A}$ 是随机、平稳且遍历的，并满足标准可积性条件时，会呈现一种特别简洁的结构

$$\mathbb{E}\left(\log^+ \|A_1\|\right) < \infty, \tag{4}$$

其中，对于 $x > 0$，$\log^+ x := \max(0,\ \log\ x)$，并约定 $\log^+ 0 := 0$。根据奥塞莱德茨乘法遍历定理，可得确定的李雅普诺夫指数

$$\gamma_1 > \cdots > \gamma_\ell, \qquad \ell \leq d,$$

以及相应的子空间滤过

$$\mathbb{R}^d = E_1 \supset E_2 \supset \cdots \supset E_\ell \supset E_{\ell+1} := \{0\},$$

使得式 (2) 中的上极限实际上是极限，并且 $\gamma(v) \in \{\gamma_1, \ldots, \gamma_\ell\}$ 几乎处处成立，其值由 v 在滤过中的位置决定；参见 [7, 第 2.1] 节和 [5, 6, 37]。特别地，指数和子空间以概率 1 与具体实现无关。

然而，即使在这一遍历框架下，指数 γ_k 也很少能以闭式形式求得。三角余循环是一个值得注意的例外，其中对角元素决定了渐近增长。特别地，对于上三角矩阵，我们有：

命题 1.1。 *[34]* 设 $\mathcal{A} = (A_k)_{k\in\mathbb{N}} \in \mathcal{M}_d^{\mathbb{N}}$ 是由 $d \times d$ 可逆矩阵组成的随机、平稳且遍历的序列，其中 $A_k(i,j) = 0$ 对于 $i > j$ 成立，并且满足 (4)。那么，

$$\gamma_1 = \max_{i\in[d]} \mathbb{E}\left(\log |A_1(i,i)|\right),$$

其中 $[d] := \{1,\dots,d\}$.

事实上，在命题的条件下， $\mathcal{A}$ 的所有李雅普诺夫指数都可以识别为多重集 $\{\mathbb{E}(\log |A_1(i,i)|)\}_{i=1}^{d}$ 中的不同元素（参见 [19]中的推论 1，以及 [5, 第 3.2]节第 130 页的条目 (iv)）。

Furstenberg 和 Kifer 的一个引理为命题 1.1 提供了概念上的解释 [17] （此处引作引理 2.3）：它将分块三角乘积的最大李雅普诺夫指数计算化简为对角块的计算；例如，参见[17, 25, 12, 18, 31]。相关扩展涉及多个方向，包括切换系统[33, 35, 36] 和丛上的线性余循环 [25, 引理 3.6]；另见 [12, 命题 1]，[18, 定理 1.1]和 [31, 引理 4.9]。本文的出发点是，同样的分块三角视角可以整理成一套实用工具，即使超越经典的平稳随机情形，仍然有用。具体而言，所得界限可由对角块增长率和零模式的简单组合参数明确表示，因此适用于矩阵分析应用中的计算。

我们的贡献围绕四种结构化情形展开。第2节首先讨论上三角矩阵乘积的确定性控制：对于一般的温和序列（不一定是随机的），命题 2.1 给出了 γ_1 关于对角增长率 $\alpha_i^{\pm}$ 的双边估计；当相关对角平均值收敛时，推论2.2则得到熟知的“对角指数的最大值”这一等式。同一节中，我们还在对角块满足李雅普诺夫正则性假设的条件下，为分块三角乘积建立了确定性的 Furstenberg–Kifer 型约化（命题 2.4），提供了一个模块化步骤，将γ_1 约化为对角块数据加上一个显式误差项。

第3节展示了这些约化如何将具体的扰动模型（包括秩一更新）转化为可计算李雅普诺夫指数的低维余循环或标量余循环，并阐明扰动何时只会改变顶端指数。最后，第4节针对支撑不相交且反馈有限的稀疏分解引入形状图，证明了一个形式为 $\gamma_1(\mathcal{A}) \le \beta + \log k$（或在 $O_d \in \mathcal{V}$时为 $\beta + \log k_*$）的能量–熵界（定理 4.4），并在第4.1节提供了一份简明的“操作指南”及实例，其中包括一个使其机制一目了然的传递矩阵/DAG 模型。

2 分块三角矩阵

我们首先给出关于给定上三角矩阵序列的最大李雅普诺夫指数的如下不等式（参见 [6, 引理 3.1.4]）。

命题 2.1。 设 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$ 是一个实可逆上三角 $d \times d$矩阵序列（即 $A_n(j,k) = 0$ 对于 $j > k$ 以及 $A_n(j,j) \neq 0$, $j,k \in [d]$），满足

$$\limsup_{n\to\infty} \frac{1}{n} \log \|A_n\| = 0. \tag{5}$$

那么，

$$\max_{j\in[d]} \alpha_j^+ \le \gamma_1 \le \max_{j\in[d]} \Big(\alpha_j^+ + \sum_{r=1}^{j-1} \big(\alpha_r^+ - \alpha_r^-\big)\Big), \tag{6}$$

其中 α_i^- := 下极限$_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \log |A_k(i,i)|$ 且 α_i^+ := 上极限$_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \log |A_k(i,i)|$.

证明。 上界可直接由 [6]中的引理 3.1.4 得出，该引理表明，在满足命题条件时，

$$\limsup_{n\to\infty} \frac{1}{n} \log |X_n(i,j)| \le \alpha_j^+ + \sum_{r=1}^{j-1} \big(\alpha_r^+ - \alpha_r^-\big), \qquad \forall i,j \in [d].$$

下面给出的下界证明利用了奇异值的增长率与 X_n 的李雅普诺夫指数之间的联系。对于 $A \in \mathcal{M}_d$，令 $\sigma_1(A) \ge \cdots \ge \sigma_d(A)$ 为 A 的奇异值，，即 $(A^TA)^{1/2}$ 的特征值，，其中 A^T 是 A 的转置。由于 $\mathbb{R}^d$ 上的所有范数都等价，最大李雅普诺夫指数 $\gamma_1(\mathcal{A})$ 不依赖于算子范数的具体选择；尤其是在讨论奇异值时，我们可以使用欧几里得范数。设

$$\ell_i^{\sup} := \limsup_{n\to\infty} \frac{1}{n} \log \sigma_i(X_n) \qquad \text{and} \qquad \ell_i^{\inf} := \liminf_{n\to\infty} \frac{1}{n} \log \sigma_i(X_n). \tag{7}$$

关于这些增长率的一般理论，以及在 $\mathcal{A} = (A_n)_{n\in\mathbb{N}}$ 为非随机时它们与李雅普诺夫指数之间的关系，参见 [5, 7]。如果 $\mathcal{A}$ 是一个随机、平稳且遍历的序列，并且（4）成立，那么（例如参见 [5] 中的定理 3.4.1），

$$\ell_i^{\sup} = \ell_i^{\inf} = \gamma_i, \qquad \forall i \in [d], \quad \text{a. s.}$$

对于非随机序列 $\mathcal{A}$， [7] 已证明

$$\ell_1^{\sup} = \gamma_1 \qquad \text{and} \qquad \ell_i^{\sup} \le \gamma_i \quad \text{for } i > 1, \tag{8}$$

此外，上述不等式可能是严格的。然而，事实证明 [6, 7] ，如果（5）成立（此类序列在 [6, 7] 中称为温和的），并且 $\mathcal{A}$ 中的矩阵是可逆的，那么以下性质等价:

$$\text{(a) } (X_n^TX_n)^{\frac{1}{2n}} \text{ converges as } n\to\infty \quad \text{and} \quad \text{(b) } \gamma_i = \ell_i^{\inf} = \ell_i^{\sup} \text{ for all } i \in [d]. \tag{9}$$

满足这些条件的序列称为李雅普诺夫正则序列，参见 [6] ，其中给出了其他一些等价条件（参见下文的注 2.5）。

下界的证明：令 $\lambda_1(X_n)$ 为 X_n. 的按绝对值计最大的特征值。根据魏尔不等式 [20], $|\lambda_1(X_n)| \le \sigma_1(X_n)$。由于 $X_n(j,\ j) = \prod_{i=1}^n A_i(j,\ j)$,

$$\log |\lambda_1(X_n)| = \max_{j\in[d]} \sum_{i=1}^n \log |A_i(j,j)| \le \log \sigma_1(X_n),$$

由此，利用 (7) 和 (8) 中的恒等式，可得 (6) 中的下界。 □

命题 2.1 将一个常见的启发式观点精确化：对于上三角乘积，对角线控制指数增长，而非对角项只能造成暂时性放大。量 $\alpha_i^{\pm}$ 记录了沿各个坐标的对角余循环的渐近增长率。当对角平均值收敛时（即 $\alpha_i^+ = \alpha_i^-$ ），推论2.2 给出简洁的恒等式 $\gamma_1 = \max_i \alpha_i$。若这些平均值不收敛，额外的修正项$\sum_{r<j}(\alpha_r^+ - \alpha_r^-)$则衡量较早坐标上 limsup 与 liminf 之间的最大“跨度”；这恰好是控制非对角耦合放大第 j 个对角增长的频率所需的余量。在平稳遍历情形下，该跨度几乎必然消失，因此可以将此命题视为经典结果周围的一个确定性包络。

现在可以陈述命题 1.1 的一个确定性类比：

推论2.2。 设命题 *2.1* 的条件对 $\mathcal{A} = (A_n)_{n\in\mathbb{N}}$. 成立。另假设对所有 $i \in [d].$，都有 $\alpha_i^+ = \alpha_i^-$。那么， $\gamma_1 = \max^{j\in[d]} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \log |A_i(j, j)|.$ $^-$

接下来我们陈述 命题 1.1 关于分块三角矩阵的推广，其形式如下

$$A_n = \begin{bmatrix} B_n(1,1) & B_n(1,2) & \dots & B_n(1,m) \\ O(2,1) & B_n(2,2) & \dots & B_n(2,m) \\ \vdots & \vdots & \ddots & \vdots \\ O(m,1) & O(m,2) & \dots & B_n(m,m) \end{bmatrix}, \tag{10}$$

其中 $B_n(i, j)$ 是由正整数序列 $(s_1, \dots, s_m)$ 确定的 $s_i \times s_j$ 实矩阵，满足$s_1 + \cdots + s_m = d$，且 $O(i, j)$ 是零 $s_i \times s_j$ 矩阵。

引理 2.3。 *[17,* 引理 *3.6]* 给定 $m \in \mathbb{N}$ 和 $(s_1, \dots, s_m) \in \mathbb{N}^m$，设 $(A_n)_{n\in\mathbb{N}}$ 是一个随机的平稳遍历矩阵序列，每个矩阵的形状均为式 (10) 中引入的形状。另假设对所有 $i \in [m]$, $B_1(i, i)$ 以概率 *1* 可逆，且对所有$1 \le i \le j \le m$,

$$\mathbb{E}\big(\log^+ \|B_1(i,j)\|\big) < \infty \qquad and \qquad \mathbb{E}\big(\log^+ \|B_1^{-1}(i,i)\|\big) < \infty.$$

对于 $i \in [m]$, 令

$$\beta_i = \lim_{n\to\infty} \frac{1}{n} \log \big\|B_n(i,i)B_{n-1}(i,i)\cdots B_1(i,i)\big\|, \qquad \text{a. s.}$$

那么， $\gamma_1 =$ 最大值$_{i\in[m]}\, \beta_i.$

在本文的其余部分中，对于 $\mathcal{A} \in \mathcal{M}_d^{\mathbb{N}}.$ 的李雅普诺夫指数，我们有时写作 $\gamma_1(\mathcal{A})$，而不是 γ_1。接下来我们将得到引理 2.3 的以下确定性版本。

命题 *2.4.* 对于给定的 $m \in \mathbb{N}$ 以及 $(s_1, \dots, s_m) \in \mathbb{N}^m$，设 $\mathcal{A} = (A_n)_{n\in\mathbb{N}}$ 为一个缓增矩阵序列（即 *(5)* 成立），其中每个矩阵均具有 *(10)* 中引入的形状。对于 $i \in [m]$, 设 $\mathcal{B}_i := \big(B_n(i, i)\big)_{n\in\mathbb{N}}$，并假设所有 $B_n(i, i)$ 均可逆，且所有 $\mathcal{B}_i$ 均为李雅普诺夫正则（即对 $\mathcal{B}_i$ 成立与 *(9)* 类似的条件）。则 $\gamma_1(\mathcal{A}) = \max^{j\in[m]} \gamma_1(\mathcal{B}_j)$。

注 2.5。

(a) 对于可逆余循环 $\mathcal{A}=(A_n)_{n\in\mathbb{N}}$，*Barreira* 引入Grobman 系数 （参见例如 *[7]* 中的第 *2.2* 节），将其作为一个非负量，用以度量对 *Lyapunov* 正则性的偏离；特别地，该余循环为李雅普诺夫正则当且仅当其 *Grobman* 系数为零点。在命题 *2.4* 的假设下，每个对角分块余循环 $\mathcal{B}_j$ 均为李雅普诺夫正则，因而其 *Grobman* 系数（在 *[7]* 的意义下）为零:

$$\mathrm{Grob}(\mathcal{B}_j)=0,\qquad \forall\ j\in[m], \tag{11}$$

其中 $\mathrm{Grob}(\cdot)$表示 *Grobman* 系数。

(b) 在引理 *2.3* 的假设下（特别是 $B_1(i,i)^{-1}$ 的可积性），对角分块余循环 $\mathcal{B}_i$ 的最小 *Lyapunov* 指数可通过逆乘积表示（例如参见 *[5]* 中的定理 *3.3.10*）:

$$\lambda_{s_i}(\mathcal{B}_i)=-\lim_{n\to\infty}\frac{1}{n}\log\left\|B_1^{-1}(i,i)\cdots B_n^{-1}(i,i)\right\|,\qquad \text{a. s.,} \tag{12}$$

对所有 $i\in[m]$，其中 $\lambda_{s_i}(\mathcal{B}_i)$ 表示 $\mathcal{B}_i$ 的最小 *Lyapunov* 指数。

命题 *2.4* 的证明。 注意到 X_n 与 A_n 矩阵具有相同的形状。记

$$X_n=\begin{bmatrix} C_n(1,1) & C_n(1,2) & \dots & C_n(1,m)\\ O & C_n(2,2) & \dots & C_n(2,m)\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \dots & C_n(m,m)\end{bmatrix}\quad H_n:=\begin{bmatrix} C_n(1,1) & O & \dots & O\\ O & C_n(2,2) & \dots & O\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \dots & C_n(m,m)\end{bmatrix},$$

其中 $C_n(i,j)$ 是 $s_i\times s_j$ 矩阵。于是,

$$\|X_n\|_F^2=\|H_n\|_F^2+\|X_n-H_n\|_F^2+2\langle H_n,X_n-H_n\rangle_F,$$

其中 $\langle A,B\rangle_F=\sum_{i,j}A(i,j)B(i,j)$ 是两个矩阵 A和$B,$的 Frobenius 内积，前提是它们具有相同的形状。由于对于任意一对 $(i,j),\ \in[m]^2,\ H_n(i,j)$ 或 $(X_n-H_n)(i,j)$ 为零，$\langle H_n,X_n-H_n\rangle_F=0.$这意味着 $\|X_n\|_F\geq\|H_n\|_F$。由于 H_n 是分块对角的,

$$\|H_n\|_F^2=\sum_{j=1}^{m}\|C_n(j,j)\|_F^2.$$

利用范数等价 $\|M\|\leq\|M\|_F\leq\sqrt{d}\,\|M\|$（对所有$\bar{}d\times d$矩阵均成立），无论用 $\|\cdot\|$ 还是 $\|\cdot\|_F$ 度量，每个块的指数增长率都相同。因此,

$$\lim_{n\to\infty}\frac{1}{n}\log\|H_n\|_F=\max_{j\in[m]}\gamma_1(\mathcal{B}_j)=:\beta,$$

对每个 $i\in[m]$，令 $\beta_i:=\gamma_1(\mathcal{B}_i)$，使得 $\beta_i\leq\beta$。

并取 $\frac{1}{n}\log$，在 $\|X_n\|_F\geq\|H_n\|_F$ 中得到 $\beta\leq\gamma_1(\mathcal{A})$。现在转而证明反向不等式 $\beta\geq\gamma_1(\mathcal{A})$。令 $d_0=0$和 $d_j=s_1+\cdots+s_j$，其中 $j\in[m].$。对于 $u\in\mathbb{R}^d$，令 $u_j,\ j=1,$ ....., m为 $\mathbb{R}^{s_j}$中的一个向量，其

$$u_j(i)=u(d_{j-1}+i),\qquad i=1,\dots,s_j.$$

用 $\|\cdot\|_1$ 表示 $\mathbb{R}^d$ 中的 ℓ^1-范数及相应的矩阵范数。因此， $\|u\|_1 = \sum_{i=1}^d |u(i)|$，并且 $\|A\|_1 = \sup_{u\in\mathbb{R}^d\setminus\{0\}} \frac{\|Au\|_1}{\|u\|_1} = \max_{j\in[d]} \sum_{i=1}^d |A(i,\ j)|$，其中 $A \in \mathcal{M}_d$.。那么，对于 $u \in \mathbb{R}^d$,

$$\begin{aligned}\|X_n u\|_1 &= \sum_{i=1}^m \Big\| \sum_{j=i}^m C_n(i,j)u_j \Big\|_1 \le \sum_{i=1}^m \sum_{j=i}^m \|C_n(i,j)\|_1 \cdot \|u_j\|_1 \\ &\le \sum_{i=1}^m \sum_{j=i}^m \|C_n(i,j)\|_1 \cdot \|u\|_1.\end{aligned}$$

因此,

$$\|X_n\|_1 \le \sum_{i=1}^m \sum_{j=i}^m \|C_n(i,j)\|_1. \tag{13}$$

鉴于 (13)，只需证明对于所有满足 $i,j \in [m]$ 且 $i \le j$,的

$$\limsup_{n\to\infty} \frac{1}{n} \log \|C_n(i,j)\|_1 \le \beta, \qquad \text{a. s.} \tag{14}$$

为证明 (14)，我们改编引理 3.1.4 的证明中采用的归纳论证，以便得到适用于上三角矩阵的这一估计版本。对于 $j > i,$，按照 X_0 表示单位矩阵的约定，我们有（参见 [7] 中的 (3.10) 和 (3.11)）：

$$\begin{aligned}C_n(i,j) &= \sum_{t=i+1}^j B_n(i,t)C_{n-1}(t,j) + B_n(i,i)C_{n-1}(i,j) \\ &= \sum_{t=i+1}^j B_n(i,t)C_{n-1}(t,j) + B_n(i,i)\sum_{t=i+1}^k B_{n-1}(i,t)C_{n-2}(t,j) \\ &\qquad + B_n(i,i)B_{n-1}(i,i)C_{n-2}(i,j) = \cdots \\ &= \sum_{r=0}^{n-1} B_n(i,i)\cdots B_{n-r+1}(i,i)\sum_{t=i+1}^j B_{n-r}(i,t)C_{n-1-r}(t,j).\end{aligned} \tag{15}$$

固定某个 $i^*, j^* \in [m]$，使得 $i^* < j^*$. 对于 $i = j = j^*$.，(14) 中的不等式显然成立。现在假设，对于所有满足 $j = j^*$ 的 (i,j)，(14) 均已得到证明，并且 $i > i^*$. 。为了完成该定理的证明，接下来我们将使用反向归纳，证明它对 $(i,j) = (i^*, j^*)$. 也成立。根据 (5) 和归纳假设，对于所有 $\varepsilon > 0$，存在一个常数 D_ε（它取决于 $\mathcal{A}$），使得对于所有 $n \in \mathbb{N},$

$$\|B_n(i,j)\|_1 \le D_\varepsilon e^{n\varepsilon}, \qquad i,j \in [m],$$

and

$$\|C_n(i,j^*)\|_1 \le D_\varepsilon e^{n(\beta+\varepsilon)}, \qquad i^* < i \le j^*.$$

此外，由 (11) 可知，对于所有 $\varepsilon > 0$，存在一个常数 G_ε（它取决于 $\mathcal{A}$），使得对于任意 $i \in [m]$ 和 $k, n \in \mathbb{N}$，满足 $k < n$,

$$\begin{aligned} \left\|B_n(i,i)\cdots B_k(i,i)\right\|_1 &= \left\|B_n(i,i)\cdots B_1(i,i)\cdot B_1^{-1}(i,i)\cdots B_{k-1}^{-1}(i,i)\right\|_1 \\ &\le \left\|B_n(i,i)\cdots B_1(i,i)\right\|_1 \cdot \left\|B_1^{-1}(i,i)\cdots B_{k-1}^{-1}(i,i)\right\|_1 \\ &\le \left(G_\varepsilon e^{(\beta_i+\varepsilon)n}\right)\cdot\left(G_\varepsilon e^{(-\beta_i+\varepsilon)(k-1)}\right) \le G_\varepsilon^2 e^{(n-k+1)\beta_i} e^{2n\varepsilon}. \end{aligned} \tag{16}$$

因此，根据 (15),

$$\begin{aligned} \|C_n(i^*,j^*)\|_1 &\le \sum_{r=0}^{n-1} G_\varepsilon^2 e^{\beta_i r} e^{2n\varepsilon} \sum_{t=i^*+1}^{j^*} D_\varepsilon^2 e^{n\varepsilon} e^{(n-r+1)(\beta+\varepsilon)} \\ &\le nj^* G_\varepsilon^2 D_\varepsilon^2 e^{(4n+1)\varepsilon} e^{(n+1)\beta}, \end{aligned}$$

由于 $\varepsilon > 0$ 是任意的，这意味着 $(i,\ j) = (i^*,\ j^*)$ 满足 (14)。 □

3 矩阵扰动的应用

接下来，我们将第 2 节的结果应用于线性系统的扰动。这些扰动代表了一类重要的矩阵变换，在科学和工程领域有着广泛的应用。

我们的第一个结果涉及特定形式的随机秩一扰动的乘积。这个命题也可以在某种程度上视为 [4], 中定理 5.1 的推广，并对同一篇论文中备注 2.4 的讨论有所补充。

命题 3.1. 设 $(u_n, \eta_n)_{n\in\mathbb{N}}$ 是一个随机平稳遍历序列，其中 $u_n \in \mathbb{R}^d$ 和 $\eta_n \in \mathbb{R}$。假设 $v \in \mathbb{R}^d$ 是一个固定向量，使得

$$\mathbb{E}\big(\log|\eta_1|\big) \le \mathbb{E}\big(\log|\eta_1 + v^T u_1|\big) < \infty. \tag{17}$$

(a) 令 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$，其中 $A_n := \eta_n I_d + u_n v^T$ 满足（4）。则

$$\gamma_1(\mathcal{A}) = \mathbb{E}\big(\log|\eta_1 + v^T u_1|\big) \quad and \quad \gamma_r(\mathcal{A}) = \mathbb{E}\big(\log|\eta_1|\big), \quad r > 1. \tag{18}$$

(b) 假设 $A \in \mathcal{M}_d$ 是一个非奇异矩阵。令 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$，其中 $A_n := \eta_n A + u_n v^T$满足条件（4），且

$$P\big(Au_1 v^T = u_1 v^T A\big) = 1. \tag{19}$$

则，

$$-\log\rho(A^{-1}) \le \gamma_1(\mathcal{A}) - \mathbb{E}\big(\log|\eta_1 + v^T A^{-1} u_1|\big) \le \log\rho(A),$$

其中 $\rho(A)$ 表示矩阵 A 的谱半径。

我们注意到，在序列 $(\eta_n,\ u_n)_{n\in\mathbb{N}}$ 满足较弱条件时，可以放宽不等式 (17)，因为该结果可直接应用于逆余循环 $\mathcal{A}^{-1} := (A_n^{-1})_{n\in\mathbb{N}}$，并相应调整 (a) 部分。

注 3.2（右向量变化：两种可延拓情形）。考虑 $A_n = \eta_n I + u_n v_n^\top$。

(i) 若对于固定的 $v \neq 0$，$v_n = c_n v$，则 $A_n = \eta_n I + \tilde{u}_n v^\top$，且 $\tilde{u}_n := c_n u_n$。因此，命题 *3.1* 可直接适用（在满足相应的 $\tilde{u}_n$ 可积性条件下），得到 $d-1$ 个李雅普诺夫指数等于 $\mathbb{E}\log|\eta_1|$，其余指数为 $\mathbb{E}\log|\eta_1 + c_1 v^\top u_1|$。

(ii) 更一般地，若对于固定的、维数为 r 的确定性子空间 $V \subset \mathbb{R}^d$，几乎处处有 $v_n \in V$，则 $H := V^\perp$ 是一个公共不变子空间，且 $A_n|_H = \eta_n \,\mathrm{Id}$。因此，$\mathbb{E}\log|\eta_1|$ 是一个重数至少为 $d-r$ 的李雅普诺夫指数。其余 r 个指数是商空间 $\mathbb{R}^d/H$ 上诱导余循环的指数；一般而言，如果 v_n 在 V 中旋转，它们就无法显式表示。尽管如此，行列式恒等式表明

$$\sum_{i=1}^{d} \gamma_i = (d-1)\mathbb{E}\log|\eta_1| + \mathbb{E}\log|\eta_1 + v_1^\top u_1|.$$

为证明命题 3.1，我们首先给出一个预备结果，说明如何利用引理 2.3 提取分块上三角矩阵的前几个李雅普诺夫指数。这推广了 [19] 中的推论 1（参见本讲义中命题 1.1 后的段落）。

命题 3.3. 固定 d、$m \in \mathbb{N}$，并设 $\mathcal{N} := (N_n)_{n\in\mathbb{N}}$ 为一个随机平稳遍历矩阵序列

$$N_n := \begin{bmatrix} A_n & * \\ O & \eta_n I_m \end{bmatrix},$$

其中 $\eta_n \in \mathbb{R}$ 且 $A_n \in \mathcal{M}_d$。假设 $\mathcal{N}$ 满足 (4)（即 $\mathbb{E}(\log^+\|N_1\|) < \infty$），$A_1$ 几乎必然可逆，且 $\mathbb{E}(\log^+\|A_1^{-1}\|) < \infty$，并且 $\eta_1 \neq 0$ 几乎必然成立，且$\mathbb{E}(\log^+|\eta_1|^{-1}) < \infty$。令 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$。那么

$$\gamma_1(\mathcal{N}) = \max\big(\mathbb{E}(\log|\eta_1|),\, \gamma_1(\mathcal{A})\big),$$

对于任意 $2 \le r \le \min(m, d)$，

$$\gamma_r(\mathcal{N}) = \max_{0\le\ell\le r}\left(\sum_{s=1}^{\ell}\gamma_s(\mathcal{A}) + (r-\ell)\mathbb{E}(\log|\eta_1|)\right) - \max_{0\le\ell\le r-1}\left(\sum_{s=1}^{\ell}\gamma_s(\mathcal{A}) + (r-\ell)\mathbb{E}(\log|\eta_1|)\right).$$

证明。回想一下，大小为 r 的子式与长度为 r 的有序指标序列 J 和 L 相关联，其定义为由指标为 J 的行和指标为 L 的列构成的子矩阵的行列式。记作 $[N] := \{1, \dots, N\}$；对于 $0 \le r \le N$，令 $[N]_r := \{(i_1, \dots, i_r) : 1 \le i_1 < \cdots < i_r \le N\}$；我们将这类递增的 r 元组与其对应的 r 元子集等同起来。对于 $B \subset [m]$，记作 $B+d := \{b+d : b \in B\} \subset \{d+1, \dots, d+m\}$；对于 $J_2, L_2 \subset \{d+1, \dots, d+m\}$，我们将 $I_m[J_2, L_2] := I_m[J_2 - d,\, L_2 - d]$ 理解为。对于任意矩阵 A，令 $H_r(A)$ 表示 A 的第 r 阶复合矩阵，其元素均为大小为 r 的子式。令 $B_n^{(r)} := H_r(N_n)$，并令 $\mathcal{B}_r := (B_n^{(r)})_{n\in\mathbb{N}}$ 表示 $\mathcal{N}$ 中第 r 阶复合矩阵构成的序列。

回想一下，根据 [5, 第 3.2], 节第 130 页的讨论，我们记作

$$\gamma_1(\mathcal{B}_r)=\sum_{i=1}^{r}\gamma_i(\mathcal{N})\quad\text{or equivalently}\quad\gamma_r(\mathcal{N})=\begin{cases}\gamma_1(\mathcal{B}_r)-\gamma_1(\mathcal{B}_{r-1}) & r\geq 2,\\ \gamma_1(\mathcal{B}_1) & r=1.\end{cases}\tag{20}$$

注意到， $B_n^{(r)}$也是一个分块上三角矩阵，而 N_n 的任意“对角线”子式都可以写成 $\eta_n I_m$ 的一个子式与 A_n的一个子式的乘积。为使其形式化，我们定义 $H_0(A) = I_1$。对于任意两个大小相同的子集 $J, L \subset [m]$， $\eta_n I_m$ 的第(J, L)个子式在 $J \neq L$时为零点，否则等于 $\eta_n^{|J|}$。设 $J, L \in [m+d]_r$ 满足 J 和 L 分别可写为 $J_1 \cup J_2$ 和 $L_1 \cup L_2$，其中 J_1, $L_1 \in [d]_\alpha$ 且$J_2, L_2 \in \{B + d \mid B \in [m]_\beta\}$，对某个 $\alpha + \beta = r$成立。那么，

$$\det N_n[J, L] = \eta_n^{\beta} \det A_n[J_1, L_1] \det I_m[J_2 - d, L_2 - d].$$

那么，观察可知，对于 $2 \leq r \leq \min(m, d)$，

$$B_n^{(r)}=\begin{bmatrix} H_r(\eta_n I_m)\otimes H_0(A_n) & * & * & \cdots & * \\ O & H_{r-1}(\eta_n I_m)\otimes H_1(A_n) & * & \cdots & * \\ O & O & \cdots & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & O & H_0(\eta_n I_m)\otimes H_r(A_n) \end{bmatrix},$$

其中 $\otimes$ 表示张量积。因此，引理 2.3 给出

$$\gamma_1(\mathcal{B}_r)=\max_{0\leq\ell\leq r}\gamma_1\big(\big(\eta_n^{r-\ell}H_\ell(A_n)\big)_{n\in\mathbb{N}}\big)=\max_{0\leq\ell\leq r}\Big(\sum_{s=1}^{\ell}\gamma_s(\mathcal{A})+(r-\ell)\mathbb{E}(\log|\eta_1|)\Big).$$

因此，由 (20)，证明随之得出。 □

现在，我们给出

命题 *3.1* 的证明。 我们从 (a) 的证明开始。若 $v = 0$，则 $A_n = \eta_n I_d$，结论立即可得。设 $v \neq 0$。选取满足 $v^T w = 1$ 的 $w \in \mathbb{R}^d$，并将 H 的一组基：$= \ker(v^T)$ 扩充为 $\mathbb{R}^d$ 的一组基，方法是把 w 添加为最后一个向量。设 P为相应的换基矩阵，从而在新坐标下 $v^T x$ 是 x 的最后一个坐标。直接计算可得

$$\tilde{A}_n := P^{-1} A_n P = \eta_n I_d + \tilde{u}_n e_d^T, \qquad \text{where } \tilde{u}_n := P^{-1} u_n \text{ and } e_d = (0, \ldots, 0, 1)^T.$$

˜特别地，每个 A_n 都是上三角的，其对角元素为 η_n（重数为 $d-1$）和 $\eta_n + e_d^T \tilde{u}_n = \eta_n + v^T u_n$。由于李雅普诺夫指数在共轭下不变，$(A_n)$ 的 ˜李雅普诺夫谱等于 (A_n) 的李雅普诺夫谱。将命题 2.1（连同 (4) 下的乘法遍历定理）应用于该三角余循环，可得

$$\gamma_1(\mathcal{A}) = \mathbb{E}\big(\log|\eta_1 + v^T u_1|\big) \quad \text{and} \quad \gamma_r(\mathcal{A}) = \mathbb{E}\big(\log|\eta_1|\big), \quad r > 1,$$

这就是（18）。

现在证明（b）部分。定义 $\tilde{\mathcal{A}} := (\tilde{A}_n)_{n\in\mathbb{N}}$，其中 $\tilde{A}_n := \eta_n I_d + A^{-1}u_n v^T$。注意，$A_n = AA_n$。反复应用（19）可得

$$X_n = A(\eta_n I_d + A^{-1}u_n v^T)\cdots A(\eta_1 I_d + A^{-1}u_1 v^T) = A^n Y_n, \tag{21}$$

其中 $Y_n := \prod_{i=1}^n A_i$。由于乘积的范数至多等于各范数之积，我们得到一个上界：

$$\begin{aligned}\gamma_1(\mathcal{A}) &\le \lim \frac{1}{n}\log\|A^n\| + \gamma_1(\tilde{\mathcal{A}}) \\ &= \log\rho(A) + \mathbb{E}\big(\log|\eta_1 + v^T A^{-1}u_1|\big),\end{aligned} \tag{22}$$

其中，最后一个等式利用了将 (a) 部分应用于 $\mathcal{A}$ 以及谱半径的盖尔范德公式。

接下来，由 (21) 得，

$$\|Y_n\| \le \|X_n\|\,\|A^{-n}\|.$$

对前一个不等式取对数，除以 n，并令 $n\to\infty$（对余循环 $(\tilde{A}_n)$ 应用 (a) 部分以及对 $\rho(A^{-1})$ 应用谱半径的盖尔范德公式），得到下界：

$$\gamma_1(\mathcal{A}) + \log\rho(A^{-1}) \ge \mathbb{E}\big(\log|\eta_1 + v^T A^{-1}u_1|\big). \tag{23}$$

结合 (23) 和 (22)，即可完成 (b) 部分的证明。 □

命题 3.1 阐明了一种一般机制：低秩扰动通常会保留一个较大的不变子空间，而在相应商空间上诱导的余循环则刻画其余的李雅普诺夫指数。在我们的情形中，超平面 $H = \ker(v^T)$ 是不变的，且 $A_n|_H = \eta_n I$；而在 $\mathbb{R}^d/H$ 上诱导的一维余循环就是乘以$\eta_n + v^T u_n$（如恒等式 $v^T A_n = (\eta_n + v^T u_n)v^T$ 所示）。等价地，经过一个将 $v^T x$ 变换到最后一个坐标的固定基变换后，每个 A_n 都变为分块上三角矩阵，其对角块为 $\eta_n I_{d-1}$ 和 $\eta_n + v^T u_n$。

这一观点并不限于秩一情形：当余循环保持一个子空间滤过时，李雅普诺夫谱由连续商空间上诱导的余循环所控制（参见引理 2.3 和命题 3.3）。

最后，我们通过恰当地将一种标准的 2×2 技术应用于当前情形，将这一结果推广为以下简单引理（例如参见 [27, 29]）。

引理 3.4。 给定 $m\in\mathbb{N}$ 和 $V\in\mathcal{M}_{d\times m}$，设 $(U_n,\eta_n)_{n\in\mathbb{N}}$ 是一个随机、平稳且遍历的序列，其中 $U_n\in\mathcal{M}_{d\times m}$ 且 $\eta_n\in\mathbb{R}$。定义 $\mathcal{A} := (A_n)_{n\in\mathbb{N}}$ 和 $\tilde{\mathcal{A}} := (\tilde{A}_n)_{n\in\mathbb{N}}$，其中

$$A_n := \eta_n I_d + U_n V^T \quad \textit{and} \quad \tilde{A}_n := \eta_n I_m + V^T U_n.$$

Then

$$\sum_{i=1}^{d}\gamma_i(\mathcal{A}) - \sum_{i=1}^{m}\gamma_i(\tilde{\mathcal{A}}) = (d-m)\mathbb{E}(\log|\eta_1|). \tag{24}$$

此外，若 $\gamma_1(\mathcal{A}) \ge \mathbb{E}(\log|\eta_1|)$，则 $\gamma_1(\mathcal{A}) = \gamma_1(\mathcal{A})$。类似地，若对于某个 $2\le r\le\min(m,d)$，有 $\gamma_r(\mathcal{A}) \ge \mathbb{E}(\log|\eta_1|)$，则 $\gamma_r(\mathcal{A}) = \gamma_r(\mathcal{A})$。

证明。 考虑以下分块方阵:

$$R := \begin{bmatrix} I_d & O \\ V^T & I_m \end{bmatrix}, \quad R^{-1} = \begin{bmatrix} I_d & O \\ -V^T & I_m \end{bmatrix}, \qquad \text{and} \qquad N_n = \begin{bmatrix} A_n & U_n \\ O & \eta_n I_m \end{bmatrix}.$$

引理的证明依赖于以下恒等式:

$$\tilde{N}_n := RN_nR^{-1} = \begin{bmatrix} \eta_n I_d & V \\ O & \tilde{A}_n \end{bmatrix}. \tag{25}$$

回顾 (1) 并利用 (25)，可写为

$$\prod_{i=1}^{n} N_i = \begin{bmatrix} X_n & * \\ O & \prod_{i=1}^{n} \eta_i I_m \end{bmatrix} = R^{-1} \begin{bmatrix} \prod_{i=1}^{n} \eta_i I_d & * \\ 0 & \tilde{X}_n \end{bmatrix} R = R^{-1} \Big(\prod_{i=1}^{n} \tilde{N}_i\Big) R,$$

其中 $*$ 表示未指定的块。令 $\mathcal{N} := (N_n)_{n\in\mathbb{N}}$，且 $\tilde{\mathcal{N}} := (\tilde{N}_n)_{n\in\mathbb{N}}$。注意，由奥塞列德茨定理，$\gamma_i(\mathcal{N})$ 和 $\gamma_i(\tilde{\mathcal{N}})$s 存在，且

$$\begin{aligned} \sum_{i=1}^{d} \gamma_i(\mathcal{A}) + m\mathbb{E}(\log|\eta_1|) &= \mathbb{E}\big(\log|\det(A_1)|\big) + m\mathbb{E}(\log|\eta_1|) \\ &= \mathbb{E}\big(\log|\det(N_1)|\big) = \sum_{i=1}^{d+m} \gamma_i(\mathcal{N}) = \sum_{i=1}^{d+m} \gamma_i(\tilde{\mathcal{N}}) \\ &= \mathbb{E}\big(\log|\det(\tilde{N}_1)|\big) = \mathbb{E}\big(\log|\det(\tilde{A}_1)|\big) + d\mathbb{E}(\log|\eta_1|) \\ &= \sum_{i=1}^{m} \gamma_i(\tilde{\mathcal{A}}) + d\mathbb{E}(\log|\eta_1|). \end{aligned}$$

重新排列可得 (24)。

接下来，$\gamma_1(\mathcal{N})$ 和 $\gamma_1(\tilde{\mathcal{N}})$ 均存在且相等。因此，由引理 2.3,

$$\max\big(\gamma_1(\mathcal{A}), \mathbb{E}(\log|\eta_1|)\big) = \max\big(\gamma_1(\tilde{\mathcal{A}}), \mathbb{E}(\log|\eta_1|)\big). \tag{26}$$

如果 $\gamma_1(\mathcal{A}) \geq \mathbb{E}(\log|\eta_1|)$，则由 (26) 可得 $\gamma_1(\mathcal{A}) = \gamma_1(\tilde{\mathcal{A}})$。

类似地，根据命题 3.3，对于任意 $2 \leq r \leq \min(m, d)$，可以归纳得出

$$\max_{0\leq\ell\leq r}\Big(\sum_{s=1}^{\ell} \gamma_s(\mathcal{A}) + (r-\ell)\mathbb{E}(\log|\eta_1|)\Big) = \max_{0\leq\ell\leq r}\Big(\sum_{s=1}^{\ell} \gamma_s(\tilde{\mathcal{A}}) + (r-\ell)\mathbb{E}(\log|\eta_1|)\Big).$$

在假设 $\gamma_1(\mathcal{A}) \geq \cdots \geq \gamma_r(\mathcal{A}) \geq \mathbb{E}(\log|\eta_1|)$ 的条件下，即可得到该结果。 □

下面给出该结果对应于命题 3.1 中 (b) 部分的结论。

命题 3.5。 取定 $m, d \in \mathbb{N}$、一个秩为 m 的矩阵 $V \in \mathcal{M}_{d\times m}(\mathbb{R})$ 以及一个可逆矩阵$A \in \mathcal{M}_{d\times d}(\mathbb{R})$。设 $(U_n, \eta_n)_{n\geq 1}$ 是一个随机平稳遍历序列，满足 $U_n \in \mathcal{M}_{d\times m}(\mathbb{R})$和 $\eta_n \in \mathbb{R}$，并定义

$$A_n := \eta_n A + U_n V^T, \qquad \widetilde{A}_n := \eta_n I_m + V^T A^{-1} U_n.$$

假设

$$\mathbb{P}(AU_1V^T = U_1V^TA) = 1.$$

~ ~接着，记 $\mathcal{A} = (A_n)$ 和 $\mathcal{A} = (A_n)$，

$$-\log\rho(A^{-1}) + \max\Big(\mathbb{E}\log|\eta_1|,\, \gamma_1(\widetilde{\mathcal{A}})\Big) \;\le\; \gamma_1(\mathcal{A}) \;\le\; \log\rho(A) + \max\Big(\mathbb{E}\log|\eta_1|,\, \gamma_1(\widetilde{\mathcal{A}})\Big).$$

证明。定义

$$\widehat{A}_n := \eta_n I_d + A^{-1}U_nV^T.$$

By ^根据交换假设， A 与 A_n 对每个 n都交换，因此 e

$$X_n := A_n\cdots A_1 = A^n\widehat{X}_n, \qquad \text{where } \widehat{X}_n := \widehat{A}_n\cdots\widehat{A}_1.$$

因此，根据次乘性，

$$\|X_n\| \le \|A^n\|\,\|\widehat{X}_n\| \quad \text{and} \quad \|\widehat{X}_n\| = \|A^{-n}X_n\| \le \|A^{-n}\|\,\|X_n\|.$$

取$\frac{1}{n}$的对数并令 $n\to\infty$，再利用盖尔范德公式 $\lim_{n\to\infty}\frac{1}{n}\log\|A^{\pm n}\| = \log\rho(A^{\pm1})$，得到

$$-\log\rho(A^{-1}) + \gamma_1(\widehat{\mathcal{A}}) \le \gamma_1(\mathcal{A}) \le \log\rho(A) + \gamma_1(\widehat{\mathcal{A}}).$$

还需要确定 $\gamma_1(\widehat{\mathcal{A}})$。令 $K := \ker(V^T)$，其维数为 $d-m$ （因为 $\operatorname{rank}(V) = m$）。对于 $x\in K$， $V^Tx = 0$，从而 $\widehat{A}_nx = \eta_nx$，因此 K 中的向量以速率$\mathbb{E}\log|\eta_1|$增长。

在商空间 $\mathbb{R}^d/K$上，诱导作用通过满射$q(x) = V^Tx$计算：

$$q(\widehat{A}_nx) = V^T(\eta_nx + A^{-1}U_nV^Tx) = \big(\eta_nI_m + V^TA^{-1}U_n\big)\,q(x) = \widetilde{A}_n\,q(x).$$

~因此，在商空间上诱导的余循环恰好是 $\mathcal{A}$。根据分块/扩张 Furstenberg– Kifer 引理（引理 2.3），

$$\gamma_1(\widehat{\mathcal{A}}) = \max\big(\mathbb{E}\log|\eta_1|, \gamma_1(\widetilde{\mathcal{A}})\big).$$

代入先前的不等式即可得到所声称的界。 □

4 形状图

我们以介绍引理 2.3 的一个推广来结束本笔记，该推广部分受到 [34] 中定理 A 证明的启发。作为讨论的铺垫，我们先给出引理 2.3 中上界的另一种证明，其依据是这样一个观察： A_n 中特定元素中的零元使我们能够根据结构更简单的矩阵所对应的李雅普诺夫指数来计算李雅普诺夫指数。我们指出，先前给出的定理证明对各块的增长行为提供了更精确的洞见，尤其是可以加以改编，用来证明命题 2.1。

在本节中，设 $\mathcal{A}:=(A_n)_{n\in\mathbb{N}}$ 是一个随机平稳遍历矩阵序列，如引理 2.3 的陈述所述，并将每个矩阵 A_n 写作

$$A_n := A_{n,1} + A_{n,2},$$

其中，$A_{n,1}$ 表示分块对角部分，而 $A_{n,2}$ 表示 A_n 的其余严格上三角部分。

范数约定。 在整个第4节中，我们取 $\|\cdot\|$ 为由向量 ℓ^1-范数（最大绝对列和）诱导的算子范数；在有助于明确时记作 $\|\cdot\|_1$。由于有限维空间 $\mathcal{M}_d$上的范数等价，这一选择不会改变李雅普诺夫指数，并且 **(4)** 对 $\|\cdot\|$ 成立当且仅当其对 $\|\cdot\|_1$成立。于是

$$X_n = A_{n,1}\cdots A_{1,1} + \sum_{j=1}^{n} \sum_{1\le i_1<\cdots<i_j\le n} A_{n,1}\cdots A_{i_j+1,1}\,A_{i_j,2}\,A_{i_j-1,1}\cdots A_{i_{j-1}+1,1}\,A_{i_{j-1},2}\,\cdots .$$

容易验证，将分块对角矩阵与严格分块上三角矩阵（具有相同的分块大小）相乘，会得到严格分块上三角矩阵。此外，将两个严格上三角矩阵相乘，会得到另一个非零超对角线更少的严格上三角矩阵——这一性质让人联想到严格上三角矩阵的幂零性。因此，所有包含 $j > d$ 的乘积都为零，从而

$$X_n = A_{n,1}\cdots A_{1,1} + \sum_{j=1}^{d} \sum_{1\le i_1<\cdots<i_j\le n} A_{n,1}\cdots A_{i_j+1,1}\,A_{i_j,2}\,A_{i_j-1,1}\cdots A_{i_{j-1}+1,1}\,A_{i_{j-1},2}\,\cdots .$$

固定 $\varepsilon > 0$。我们将使用以下标准的 Borel–Cantelli 估计（应用于$Y_n = \log^+ \|A_{n,2}\|_1$）。

引理 4.1。 设 $(Y_n)_{n\ge1}$ 是一个非负随机变量的平稳序列，且$\mathbb{E}(Y_1) < \infty$。那么，对每个 $\varepsilon > 0$,

$$\sum_{n\ge1} \mathbb{P}(Y_n > \varepsilon n) < \infty, \qquad \textit{and hence} \qquad \mathbb{P}(Y_n > \varepsilon n \text{ i.o.}) = 0.$$

证明。 由平稳性， $\mathbb{P}(Y_n > \varepsilon n) = \mathbb{P}(Y_1 > \varepsilon n)$。利用尾积分恒等式

$$\mathbb{E}(Y_1) = \int_0^\infty \mathbb{P}(Y_1 > t)\,dt$$

并将积分区间划分为长度为 ε 的子区间，可得

$$\sum_{n\ge1} \mathbb{P}(Y_1 > \varepsilon n) \le \frac{1}{\varepsilon}\int_0^\infty \mathbb{P}(Y_1 > t)\,dt = \frac{1}{\varepsilon}\mathbb{E}(Y_1) < \infty.$$

Borel–Cantelli 引理得出该结论。 □

应用引理 4.1，其中 $Y_n := \log^+ \|A_{n,\,2}\|_1$。由于 $A_{1,\,2}$ 是通过将 A_1 中的某些项置为零得到的，诱导的 ℓ^1 算子范数具有单调性，因此 $\|A_{1,\,2}\|_1 \le \|A_1\|_1$，从而 $\mathbb{E}\log^+ \|A_{1,\,2}\|_1 < \infty$ 由 (4) 得出（对 $\|\cdot\|_1$ 同样成立）。因此，

$$\mathbb{P}\big(\log^+ \|A_{n,2}\|_1 > \varepsilon n \text{ i.o.}\big) = 0.$$

特别地，几乎必然存在一个（随机的）$N_\varepsilon < \infty$，使得 $\|A_{n,\,2}\|_1 \le e^{\varepsilon n}$ 对所有 $n \ge N_\varepsilon$ 都成立。令

$$C_\varepsilon := \max\Big\{1,\ \max_{1\le n\le N_\varepsilon} \|A_{n,2}\|_1 e^{-\varepsilon n}\Big\},$$

我们得到一致界

$$\|A_{n,2}\|_1 \le C_\varepsilon e^{\varepsilon n} \qquad \text{for all } n \ge 1.$$

因此，利用上面的 X_n 展开式，以及任一非零单项式至多包含 $d-1$ 个来自严格上三角部分的因子这一事实（$A_{k,\,2}$），我们得到：对所有充分大的 n，几乎必然有

$$\|X_n\|_1 \le C_\varepsilon^{d-1}\, n^d\, e^{(d-1)\varepsilon n} \max_{1\le j<d} \max_{1\le i_1<\cdots<i_j\le n} \prod_{r=0}^{j} \big\|A_{i_r+1,1}\cdots A_{i_{r+1}-1,1}\big\|_1, \tag{27}$$

按约定，$i_0 := 0$ 且 $i_{j+1} := n+1$，空乘积解释为单位矩阵。（我们在展开中对求和项的数量使用了粗略上界 $1 + \sum_{j=1}^{d-1} \binom{n}{j} \le n^d$。）

令 $\mathcal{A}_1 := (A_{n,\,1})_{n\in\mathbb{N}}$。注意，$A_{n,\,1}$ 是对角线的（因此也是分块对角的），所以对任意 $1 \le i \le j$，

$$\|A_{j,1}\cdots A_{i,1}\|_1 = \max_{p\in[d]} \prod_{t=i}^{j} \big|(A_{t,1})_{pp}\big|.$$

对每个 $p \in [d]$，令

$$\beta_p := \lim_{n\to\infty} \frac{1}{n} \log \prod_{t=1}^{n} \big|(A_{t,1})_{pp}\big| \qquad \text{a.s.,} \qquad \text{so that} \qquad \gamma_1(\mathcal{A}_1) = \max_{p\in[d]} \beta_p.$$

固定 $\varepsilon > 0$。由于 $\frac{1}{n}\sum_{t=1}^{n} \log|(A_{t,1})_{pp}| \to \beta_p$ 几乎必然成立，随机变量

$$G^+_{\varepsilon,p} := \sup_{n\ge1} \exp\Big(\sum_{t=1}^{n} \log|(A_{t,1})_{pp}| - (\beta_p+\varepsilon)n\Big), \qquad G^-_{\varepsilon,p} := \sup_{n\ge1} \exp\Big(-\sum_{t=1}^{n} \log|(A_{t,1})_{pp}| - (-\beta_p+\varepsilon)n\Big)$$

几乎必然有限，因此取 $G_{\varepsilon,p} := \max(G^+_{\varepsilon,p}, G^-_{\varepsilon,p})$，对所有 $n \ge 1$，我们有

$$\prod_{t=1}^{n} \big|(A_{t,1})_{pp}\big| \le G_{\varepsilon,p}\, e^{(\beta_p+\varepsilon)n}, \qquad \prod_{t=1}^{n} \big|(A_{t,1})_{pp}\big|^{-1} \le G_{\varepsilon,p}\, e^{(-\beta_p+\varepsilon)n}.$$

因此，对于所有 $1 \le i \le j$，

$$\begin{aligned}\prod_{t=i}^{j} \big|(A_{t,1})_{pp}\big| &= \Big(\prod_{t=1}^{j} \big|(A_{t,1})_{pp}\big|\Big)\Big(\prod_{t=1}^{i-1} \big|(A_{t,1})_{pp}\big|\Big)^{-1} \\ &\le G^2_{\varepsilon,p}\, e^{\beta_p(j-i+1)}\, e^{\varepsilon(j+i-1)} \le G^2_{\varepsilon,p}\, e^{\beta_p(j-i+1)}\, e^{2\varepsilon j}.\end{aligned}$$

对 $p \in [d]$ 取最大值，并令 $G_\varepsilon := \max_{p\in[d]} G_{\varepsilon,\ p}$，即可得到一致估计

$$\|A_{j,1}\cdots A_{i,1}\|_1 \le G_\varepsilon^2\, e^{\gamma_1(\mathcal{A}_1)(j-i+1)}\, e^{2\varepsilon j} \qquad \text{for all } 1\le i\le j. \tag{28}$$

将 (28) 应用于 (27) 中的每个因子，并利用每个端点至多为 n 且$j \le d-1$，得到

$$\|X_n\|_1 \le \widetilde{C}_\varepsilon\, n^d\, \exp\big((\gamma_1(\mathcal{A}_1) + C\,\varepsilon)n\big),$$

其中某个确定性常数为 $C = C(d)$，且某个几乎必然有限的（随机）常数为$\sim C_\varepsilon$。由于 $\varepsilon > 0$ 是任意的，我们得出 $\gamma_1(\mathcal{A}) \le \gamma_1(\mathcal{A}_1)$。最后，对于对角（更一般地说，块对角）余循环，其最大李雅普诺夫指数等于其对角块的最大李雅普诺夫指数中的最大值，从而完成证明。

现在我们引入若干定义，以将前述论证抽象化，并将其推广到块三角情形之外的更广泛矩阵类别。

令 $\wedge$ 表示逐项“与”运算。固定 $d, k \in \mathbb{N}$，使得 $k \le d^2$。一个阶数为 d的形状集，记作 $\mathcal{L}$，是由大小为 $d \times d$的非零零一矩阵组成的集合，

$$\mathcal{L} := \{L_j \mid j \in [k]\},$$

使得对于所有互不相同的 $i, j \in [k]$，都有 $L_i \wedge L_j = O_d$。对于任意矩阵 $A \in \mathcal{M}_d$，我们将其形状记作 形状(A)，定义为一个 $d \times d$ 零一矩阵，其第 (i, j) 个元素在 $A(i, j) \neq 0$ 时为 $\mathbf{1}$，否则为 $\mathbf{0}$。

接下来，我们引入 $\mathcal{L}$-形状图 $(\mathcal{V}, \mathcal{T})_\mathcal{L}$的概念，它由其顶点集 $\mathcal{V}$与转移映射 $\mathcal{T}$定义：$\mathcal{V} \times \mathcal{L} \to \mathcal{V}$。给定形状集 $\mathcal{L}$，将顶点集定义为

$$\mathcal{V} := \Big\{ \mathit{Shape}(L_n \cdots L_1) \,\Big|\, n \in \mathbb{N},\ (L_1, \ldots, L_n) \in \mathcal{L}^n \Big\}. \tag{29}$$

每个顶点在必要时都可视为一个零一矩阵。两个顶点 $v_1, v_2 \in \mathcal{V}$相连，记作 $v_1 \to v_2$，当且仅当存在某个 $L \in \mathcal{L}$，使得

$$v_2 = \mathcal{T}(v_1, L) := \mathit{Shape}(Lv_1).$$

最后，对任意矩阵序列 $\mathcal{A}$:= $(A_n)_{n\ge 0}$，我们用形状集 $\mathcal{L}$ 将每个 A_n 分解为 k 个“互不重叠”的矩阵 $\{A_{n,1}, \ldots, A_{n,k}\}$ 之和，使得

$$A_n = \sum_{j=1}^{k} A_{n,j}, \quad \text{where} \quad \mathit{Shape}(A_{n,j}) \vee L_j = L_j. \tag{30}$$

为澄清这些定义，我们以若干例子结束本节，其中一例说明了适用于三角矩阵的形状图。

例 4.2. 设 $\mathcal{A}$ 是一列 3×3上三角矩阵。将每个 A_i 写成其对角部分与严格上三角部分之和。于是，一个适当的形状集为

$$\mathcal{L} := \left\{ L_1 := I_3,\ L_2 := \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right\}. \tag{31}$$

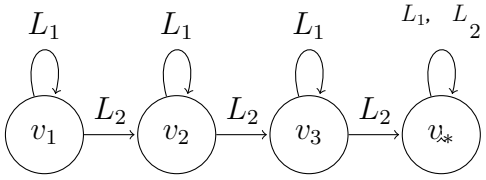
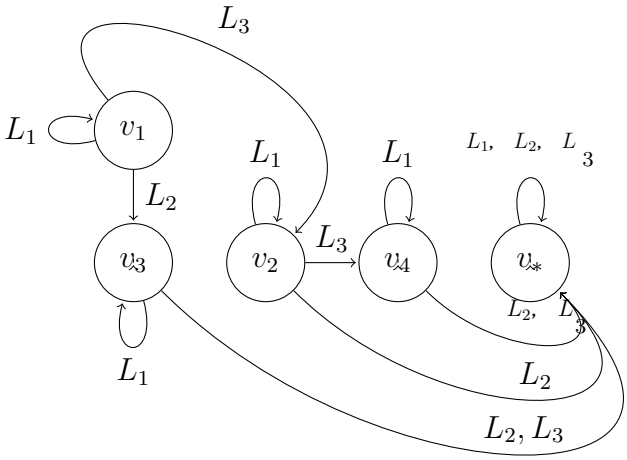
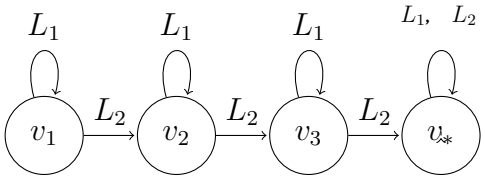


图1：(31) 和 (32) 的形状图。



(a) 参见公式 (34) 和 (35)



（b）参见式（36）和（37）

图2：由（33）所描述的矩阵对应的两种不同图

根据 (29)， $\mathcal{L}$-形状图的顶点集为

$$\mathcal{V} := \left\{ v_1 := I_3,\ v_2 := \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ v_3 := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ v_* := O_3 \right\}. \tag{32}$$

相应的形状图如图*1*所示。

下一个例子讨论矩阵并非都为上三角矩阵的情形。

例4.3。 令 $\mathcal{A}$ 为从该集合中采样得到的矩阵序列

$$\left\{ \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix},\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 5 & 0 & 0 & 4 \end{bmatrix} \right\}. \tag{33}$$

在这种情况下，形状集 $\mathcal{L}$ 的适当选择为

$$\mathcal{L} = \left\{ L_1 := I_4,\ L_2 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ L_3 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}. \tag{34}$$

不难验证，取 $\mathcal{L}$ 时，顶点集为

$$\mathcal{V} := \left\{ v_1 := L_1,\ v_2 := L_2,\ v_3 := L_3,\ v_4 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ v_* := O_4 \right\}. \tag{35}$$

相应的形状图如图 *2(a)* 所示。需要指出的是，形状集$\mathcal{L}$ 并非唯一，因此得到的图也不唯一。例如，在本例中， $\mathcal{L}$ 还可以有另一种取法:

$$\mathcal{L} = \left\{ L_1 := I_4,\ L_2 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \right\}. \tag{36}$$

此时，顶点集变为

$$\mathcal{V} := \left\{ v_1 := L_1,\ v_2 := L_2,\ v_3 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ v_* := O_4 \right\}. \tag{37}$$

相应的形状图如图 *2(b)* 所示。

在确立形状图的概念后，我们的策略是利用其结构性质，分析 $\mathcal{A}$ 中的矩阵乘积，从而得到关于 $\mathcal{A}$ 的李雅普诺夫指数的有意义的界。本节的主要结果如下定理所述。

定理 4.4. 设 $\mathcal{L}$ 是一个阶为 d、含有 k 个元素的形状集，且 $(\mathcal{V}, \mathcal{T})_{\mathcal{L}}$ 表示其对应的形状图。假设该形状图仅包含长度为一的环（自环），且每个非零顶点至多有一个自环。设 $\mathcal{A} = (A_n)_{n\in\mathbb{N}}$ 是一个随机平稳遍历序列，由 $d \times d$ 矩阵组成，并且每个矩阵都可按照形状集 $\mathcal{L}$进行如 (30) 所示的分解。进一步假设:

(i) 对于每一个 $1 \le s \le k$,

$$\mathbb{E}\big(\log^+ \|A_{1,s}\|\big) < \infty.$$

(ii) 令 $\mathcal{H} \subset \mathcal{V} \setminus \{O_d\}$ 为带有自环的顶点集合，并令 $\mathcal{L}_{\mathcal{H}}$ 为与这些自环对应的标签子集。对于每个 $s \in \mathcal{L}_{\mathcal{H}}$，矩阵 $A_{1,\ s}$ 几乎必然可逆，并且

$$\mathbb{E}\big(\log^+ \|A_{1,s}^{-1}\|\big) < \infty.$$

定义

$$\beta_s := \lim_{n\to\infty} \frac{1}{n} \log \|A_{n,s}A_{n-1,s}\cdots A_{1,s}\|, \quad a.s.,$$

并设

$$\beta := \max_{s\in\mathcal{L}_{\mathcal{H}}} \beta_s.$$

那么：

(a) 最大李雅普诺夫指数满足

$$\gamma_1(\mathcal{A}) \le \beta + \log k.$$

此外，若 $O_d \in \mathcal{V}$，定义

$$k_* := \max_{v\in\mathcal{V}\setminus\{O_d\}} \#\big\{\, s\in[k] : \mathcal{T}(v, L_s) \neq O_d \,\big\},$$

使得 $k_* \le k$ 和 k_* 是能阻止一个非零顶点一步转移到 O_d 的标签的最大数量。那么

$$\gamma_1(\mathcal{A}) \le \beta + \log k_*.$$

（特别地，若每个非零顶点至少有一个出边标签能转移到 O_d，那么 $k_* \le k-1$。）

(b) 令 $\mathcal{W} \subset \mathcal{H}\setminus\{O_d\}$ 为自环顶点的集合 w，满足形状图中任何以 w 结尾的有向路径都避开除 w 以外的环顶点（等价地说，它不会经过 $\mathcal{H}\setminus\{w\}$ 中的任何顶点）。此外，假设对每个 $w\in\mathcal{W}$，若 s 表示其自环的标签，则存在 $r = r(w) \ge 1$，使得对所有 $n \ge r$，$\mathit{Shape}(L_s^n) = w$。若对每个 $w\in\mathcal{W}$ 和每个 $v\in\mathcal{V}\setminus\{w\}$，我们有

$$v \wedge w = O_d,$$

并且还假设所有矩阵 $A_{n,s}$ 几乎必然逐项非负（因此，形状相同的单项式不会逐项相消）。那么

$$\max_{s\in\mathcal{L}_{\mathcal{W}}} \beta_s \le \gamma_1(\mathcal{A}),$$

其中 $\mathcal{L}_{\mathcal{W}}$ 表示与 $\mathcal{W}$ 中各顶点的自环对应的标签集合。

4.1 定理 4.4 的实际应用

定理 4.4 旨在为随机矩阵可进行稀疏分解且具有不相交支撑的余循环提供一种“即插即用”的估计。在许多情况下，验证过程可归结为对相应形状图进行有限的组合检验。下面列出一份简短的检查清单。

1. 选择一个形状集。 选择一个形状集 $\mathcal{L} = \{L_1, \dots, L_k\}$（不相交支撑： $L_i \wedge L_j = O_d$ 对于 $i \neq j$，其中 O_d 是 $d\times d$ 零矩阵），使其与 A_n中存在的稀疏类型相匹配。

2. 构建形状图。 计算（或推理得出）(29) 中的顶点集 $\mathcal{V}$ 以及转移映射 $\mathcal{T}(v,\ L)\ =$ 形状(vL)。实际上， $\mathcal{V}$ 是通过从标签开始迭代转移而生成的。

3. 检查无环性条件。验证有向图除自环外不含有向环，且每个非零顶点至多有一个自环标签。（在许多常见构造中， $L_1 = I_d$ 是唯一的自环标签。）

4. 识别环标号并计算 β。令 $\mathcal{H}$ 为环顶点， $\mathcal{L}_{\mathcal{H}}$ 为其自环标签。则 $\beta\ =\ \max_{s\in\mathcal{L}_{\mathcal{H}}}\beta_s$, 其中 β_s 是子余循环（$A_{n,\ s}$）$_{n\geq 1}$的最大李雅普诺夫指数。

5. 应用该界。定理 4.4(a) 给出 $\gamma_1(\mathcal{A})\ \leq \beta + \log k$，并且若 $O_d \in \mathcal{V}$，则更精确的$\gamma_1(\mathcal{A})$ $\leq \beta + \log k_*$。

我们现在给出两个演算示例：第一个使用上文引入的非三角例子，第二个描述一个一般的“有向无环图支撑”族，其中 $\log k$项具有清晰的组合意义。

例4.5 （非三角例子（33）的一个详细界）。 设（A_n）$_{n\geq 1}$是独立同分布的，并以概率 $p \in$ （0、1） 取（*33*）中的第一个矩阵，以概率 $1-p$取第二个矩阵。使用来自（34）的形状集 $\mathcal{L} = \{L_1, L_2, L_3\}$，于是 $k = 3$；注意由图 *2(a)*可知， $O_4 \in \mathcal{V}$，且形状图除自环外没有其他有向环。

唯一的自环标签是 $L_1 = I_4$，因此 $\mathcal{L}_{\mathcal{H}} = \{1\}$ 和 $\beta = \beta_1$。将分解（30）中的对角分量写作 $D_n :\ = A_{n,\ 1}$， 则 D_n 是对角的，其对角元素的分布为

$$(D_n(1,1), D_n(2,2), D_n(3,3), D_n(4,4)) = \begin{cases} (2,1,-1,-2), & with\ prob.\ p, \\ (1,2,3,4), & with\ prob.\ 1-p. \end{cases}$$

对于对角余循环， β_1 是显式的：

$$\beta_1 = \max_{1\leq i\leq 4}\ \mathbb{E}\big(\log|D_1(i,i)|\big) = \max\Big\{p\log 2,\ (1-p)\log 2,\ (1-p)\log 3,\ p\log 2+(1-p)\log 4\Big\}.$$

因此，定理 *4.4(a)*给出了具体估计

$$\gamma_1(\mathcal{A}) \leq \beta_1 + \log k_* = \beta_1 + \log 2 \qquad (here\ k_* = 2).$$

当对角增长 β_1 较为有限，而非对角元素导致较大的瞬态放大时，这个界在相应区域内是非平凡的；形状图估计将真正的指数增长（由 β_1捕获）与单项式的纯组合增殖（由 $\log 2$捕获）区分开来。

例4.6 (DAG支撑稀疏模式). 将这一类模型视为定义在有向无环网络上的随机转移矩阵（或随机动力学网络）模型，是一种便捷的理解方式。设 $x_n \in \mathbb{R}^d$ 为一个非负“质量”（或部分配分函数）向量，并按照以下方式演化：

$$x_n = A_n x_{n-1}, \qquad n \geq 1.$$

那么， $\gamma_1(\mathcal{A})$ 就是 $\|x_n\|$ 的几乎必然指数增长率，也可以将其解释为淬火自由能密度。一种具体的参数化方式是取（对于逆温度 $\theta > 0$）

$$D_n = Diag\left(e^{-\theta E_{n,1}}, \ldots, e^{-\theta E_{n,d}}\right), \qquad B_{n,e} = e^{-\theta U_{n,e}} E_{ij} \quad for\ e = (i,j) \in E(G),$$

其中 $(E_{n,\,i})$ 和 $(U_{n,\,e})$ 是平稳遍历能量/边权场。在这种情况下，对角子余循环是显式的，并且

$$\beta_1 = \max_{1 \le i \le d} \mathbb{E}\big(\log D_1(i,i)\big) = -\theta \min_{1 \le i \le d} \mathbb{E}(E_{1,i}),$$

因此，定理 ***4.4(a)*** 给出的淬火自由能上界是一个简单的“能量–熵”上界（能量来自 β_1，熵来自对数 $|E(G)|$ 项）。

为了让这一解释更加具体，请注意，将 $X_n = A_n \cdots A_1$ 展开为单项式，自然会枚举状态空间 $[d]$中的时间有序路径。事实上，如果我们在所有元素均非负的基下进行运算（或者直接逐元素取绝对值），那么对于每个坐标 j，都可以将$(X_n\mathbf{1})^j$ 视为对一系列

$$i_0 \to i_1 \to \cdots \to i_n = j$$

其中每一步要么停留原处（从 D_t中选择对角项），要么沿着 G 的一条允许的有向边前进（从某个$B_{t,\,e}$中选择）。路径的权重是沿路径相应随机因子的乘积，因此总质量是随机路径权重的一个类似配分函数的和。最大李雅普诺夫指数 $\gamma_1(\mathcal{A})$ 便是这个配分函数的淬火指数增长率。

定理 ***4.4(a)***可解读如下：在无环环境中，长期的指数增长由“最佳”循环（对角余循环，贡献 β_1）控制，而非对角移动的贡献仅来自允许的局部选择数量。粗略的熵项 $\log|E(G)|$ 源于将允许的标签序列数量上界为 $|E(G)|^n$；在许多应用中，可以用更小的量（例如最大出度或相关转移矩阵的谱半径）替换 $|E(G)|$，从而改进这一上界。

固定 $d \ge 2$，并设 G 是顶点集为 $[d]$ 的有向无环图（***DAG***）。对于每条有向边 $e = (i,j) \in E(G)$，令 $L_e := E_{ij}$ 为仅在条目 $1\,(i,j)$ 处取值为 ***1***、其余位置为零的 $d \times d$ 矩阵，并令 $L_1 := I_d$。于是

$$\mathcal{L} := \{L_1\} \cup \{L_e : e \in E(G)\}$$

这是一个形状集（各支撑互不相交），且 $k = 1 + |E(G)|$。考虑任意形如以下形式的平稳遍历余循环$(A_n)^{n \ge 1}$

$$A_n = D_n + \sum_{e \in E(G)} B_{n,e}, \qquad Shape(D_n) \vee L_1 = L_1, \quad Shape(B_{n,e}) \vee L_e = L_e,$$

其中 D_n 为对角矩阵，且每个 $B_{n,\,e}$ 的支撑仅包含由 e规定的单个元素。如果 D_n 几乎处处可逆且 $\mathbb{E}\log^+\|D_1^{\pm 1}\| < \infty$，并且对于所有 e，都有 $\mathbb{E}\log^+\|B_{1,\,e}\| < \infty$，那么定理 ***4.4*** 适用。

DAG 假设意味着 $\mathcal{L}$-形状图中不存在有向环，只有由 L_1 诱导的自环（边标签的乘积对应有向路径，而有向路径在 *DAG* 中不可能回到起点）。因此， $\mathcal{L}_{\mathcal{H}} = \{1\}$ 和 $\beta = \beta_1$，其中

$$\beta_1 = \lim_{n\to\infty} \frac{1}{n} \log \|D_n D_{n-1} \cdots D_1\| = \max_{1\le i\le d} \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \log |D_t(i,i)| \qquad a.s.$$

此外，如果 $O_d \in \mathcal{V}$ （例如，当 G 不存在长度为 d 的有向路径时，这一条件成立），那么定理 *4.4(a)* 给出

$$\gamma_1(\mathcal{A}) \le \beta_1 + \log k_* = \beta_1 + \log |E(G)| \qquad (here\ k_* = |E(G)|).$$

在这一族中，加性项 $\log |E(G)|$ 可视为一种粗略的“分支熵”，它计入了每个乘法步骤中允许的稀疏类型数量，而 β_1则反映了来自对角子余循环的实际指数增长。

注 4.7. 定理 *4.4(a)* 的证明将 X_n 展开为由 $\{1,\dots,k\}^n$ 中的标签序列索引的单项式，然后对 k^n 项应用并集界。加性项 $\log k$ （或 $\log k_*$）正是由此产生的。在应用中，这一步通常可以进一步改进：

- 如果经过几步之后实际上只有一部分标签可达（等价地说，形状图迅速收缩），则可以用实际有贡献的标签数量来替代 k 。

- 在“类图”构造中，可以用辅助转移图中的有向游走数量（等价地，用形状图中的有向路径数量）来界定部分 *(a)* 的展开中允许的标签序列数量。若M 是该转移图的邻接矩阵，则长度为-n 的游走数量至多为 $\mathbf{1}^\top M^{n-1}\mathbf{1} \le C\,\rho(M)^n$，其中 $\rho(M)$是谱半径。因此，加性熵项通常可从 $\log\ k$改进为 $\log\ \rho(M)$ （或者较粗略地改进为 $\log\ \Delta^+$，其中 Δ^+ 是最大出度）。特别是，如果转移图无环，则该数量呈次指数增长，极限中的“熵”贡献便会消失。

我们在此不探讨这些改进，但形状图的视角使它们显得顺理成章。

4.2 定理 4.4 的证明

无环假设意味着，每个时间有序单项式要么 (i) 长时间停留在某个自环顶点内（因此继承相应环余循环的指数增长率 β ），要么 (ii) 最终因到达 O_d 而归零。由于非自环转移不可能无限重复，其贡献至多呈次指数增长（利用$\mathbb{E}\log^+ \|A_{1,s}\|$ 假设来控制）。剩下唯一的指数代价是不同非零单项式的数量，其上界为 k_{eff}^n；当 O_d 可达时，可按定理 4.4 中的方式取 $k_{\text{eff}} = k_*$，从而得到相应的加性 $\log k_{\text{eff}}$ 项。

现在记录证明 (a) 部分所需的关键一致估计。

引理 4.8 (非零单项式的一致界)。 假设定理 *4.4* 的条件成立。那么对于每个 $\varepsilon > 0$ ，存在一个几乎必然有限的随机常数 K_ε ，使得对所有 $n \geq 1$ 以及每个多重指标 $(i_1, \ldots, i_n) \in [k]^n$ ，只要相应的单项式

$$A_{1,n}(i_1, \ldots, i_n) := A_{n,i_n} A_{n-1,i_{n-1}} \cdots A_{1,i_1}$$

不为零，我们有

$$\|A_{1,n}(i_1, \ldots, i_n)\|_1 \leq K_\varepsilon \, \exp\big((\beta + \varepsilon)n\big).$$

证明。 固定 $\varepsilon > 0$ 并令 $m := |\mathcal{V}|$。 令 $\delta := \varepsilon/(4m)$。

定义平稳序列

$$Y_n := \max_{1 \leq j \leq k} \log^+ \|A_{n,j}\|_1.$$

假设 (i) 蕴含 $\mathbb{E}(Y_1) < \infty$。由引理 4.1（应用于 (Y_n)），存在一个几乎必然有限的随机常数 C_δ ，使得对所有 $n \geq 1$ 以及所有 $1 \leq j \leq k$,

$$\|A_{n,j}\|_1 \leq C_\delta \, e^{\delta n}.$$

（事实上，上述界对所有充分大的 n 都成立，而剩下有限多个 n 可并入 C_δ。）

固定一个自环标签 $s \in \mathcal{L}_\mathcal{H}$。根据假设 (ii)，序列 $(A_{n,\,s})_{n \geq 1}$ 平稳且遍历，并且 $A_{n,\,s}$ 几乎必然可逆，且 $\mathbb{E}\log^+ \|A_{1,\,s}\|_1 + \mathbb{E}\log^+ \|A_{1,\,s}^{-1}\|_1 < \infty$。将此平稳过程实现为遍历保测变换上的线性余循环$\theta$ （使得 $A_{n,\,s}(\omega) = A_{1,\,s}(\theta^{n-1}\omega)$）。

根据可积线性余循环的奥塞莱德茨乘法遍历定理（例如参见 [5, 8]），~对于每个 $\delta > 0$ ，都存在一个可测的温和的随机变量 $K_{s,\,\delta}(\omega) < \infty$ ，使得对所有 $\ell \geq 1$,

$$\big\|A_{\ell,s}(\omega) \cdots A_{1,s}(\omega)\big\|_1 \leq \widetilde{K}_{s,\delta}(\omega) \, e^{(\beta_s + \delta)\ell}.$$

这里，“温和的”意味着 $\lim_{n\to\infty} \frac{1}{n} \log \widetilde{K}_{s,\,\delta}(\theta^n \omega) = 0$ 几乎处处成立。特别地，$\sup_{t \geq 1} \widetilde{K}_{s,\,\delta}(\theta^{t-1}\omega) e^{-\delta t} < \infty$ 几乎处处成立；定义

$$K_{s,\delta}(\omega) := \sup_{t \geq 1} \widetilde{K}_{s,\delta}(\theta^{t-1}\omega) e^{-\delta t}.$$

那么对所有 $t \geq 1$ 和所有 $\ell \geq 1$,

$$\Big\|A_{t+\ell-1,s} A_{t+\ell-2,s} \cdots A_{t,s}\Big\|_1 \leq K_{s,\delta}(\omega) \, e^{\delta t} \, e^{(\beta_s + \delta)\ell}. \tag{38}$$

令 $K_\delta := \max_{s \in \mathcal{L}_\mathcal{H}} K_{s,\,\delta}$，由于 $\mathcal{L}_\mathcal{H}$ 有限，因此它几乎处处有限。

固定 n 以及一个多重指标 $(i_1, \ldots, i_n)$，其中 $A_{1,n}(i_1, \ldots, i_n) \neq 0$。形状图上的相应游走至多经过 m 个顶点，因为该图除了自环外没有其他有向环。因此，指标集 $\{1, \ldots, n\}$ 可划分为至多 m 个连续的块，其中每个块要么 (i) 在当前顶点处由一个自环标签重复构成，持续整个块的长度；要么 (ii) 仅包含一次非自环转移（即游走改变顶点的一次时间步）。

设 $\ell_1, \dots, \ell_r$ 为自环块的长度（因此 $\sum_{q=1}^{r} \ell_q \le n$ 和 $r \le m$），并设 t_q 为第 q 个自环块的起始时间指标（因此 $t_q \le n$）。对每个自环块使用 (38)，并对至多 m 个转移步使用上述界，我们得到

$$\|A_{1,n}(i_1, \dots, i_n)\|_1 \le (C_\delta)^m (K_\delta)^m \exp\Big(\delta m n\Big) \exp\Big(\sum_{q=1}^{r} (\beta_{s_q} + \delta)\ell_q\Big) \exp\Big(\delta m n\Big),$$

其中 s_q 表示区块 q 中使用的自环标签。由于 $\beta_{s_q} \le \beta$ 和 $\sum_{q=1}^{r} \ell_q \le n$，中间的指数项可由 $\exp((\beta + \delta)n)$ 控制。吸收常数后，我们得到

$$\|A_{1,n}(i_1, \dots, i_n)\|_1 \le (C_\delta K_\delta)^m \exp\big((\beta + \delta + 2\delta m)n\big).$$

采用我们选择的 $\delta = \varepsilon/(4m)$，有 $\delta + 2\delta m \le \varepsilon$，因此令 $K_\varepsilon := (C_\delta K_\delta)^m$，便得到

$$\|A_{1,n}(i_1, \dots, i_n)\|_1 \le K_\varepsilon \, e^{(\beta+\varepsilon)n}.$$

此上界对所有长度为 n 的非零单项式以及所有 $n \ge 1$ 均一致成立。 □

(a) 的证明。 利用 X_n 中的分解 (30) 并展开乘积，我们写成

$$X_n = \sum_{i_1=1}^{k} \cdots \sum_{i_n=1}^{k} A_{1,n}(i_1, \dots, i_n), \quad \text{where} \quad A_{1,n}(i_1, \dots, i_n) := A_{n,i_n} A_{n-1,i_{n-1}} \cdots A_{1,i_1}. \tag{39}$$

每个多重指标 $(i_1, \dots, i_n)$ 都通过依次出现的形状 形状(L_{i_1})，形状$(L_{i_2} L_{i_1})$ $\dots$，形状$(L_{i_n} \cdots L_{i_1})$ 对应于形状图上的一条游走。如果 $O_d \in \mathcal{V}$ 且这条游走曾进入顶点 O_d，那么对应的单项式等于零矩阵（其支撑变为空集），因此对 (39) 没有贡献。因此，和式 (39) 至多包含 k_{eff}^n 个非零单项式，其中我们令

$$k_{\text{eff}} := \begin{cases} k, & O_d \notin \mathcal{V}, \\ k_*, & O_d \in \mathcal{V}, \end{cases}$$

其中 k_* 与定理陈述中的相同。

固定 $\varepsilon > 0$。由引理 4.8，存在一个几乎处处有限的随机常数 K_ε，使得对所有 $n \ge 1$ 以及出现在 (39) 中的每个非零单项式 $A_{1,\,n}(i_1,\ \dots,\ i_n)$，

$$\|A_{1,n}(i_1, \dots, i_n)\|_1 \le K_\varepsilon \, e^{(\beta+\varepsilon)n}.$$

因此，

$$\|X_n\|_1 \le k_{\text{eff}}^n \, K_\varepsilon \, e^{(\beta+\varepsilon)n}.$$

取对数，除以 n，并令 $n \to \infty$，得到

$$\gamma_1(\mathcal{A}) \le \beta + \log k_{\text{eff}} + \varepsilon.$$

由于 $\varepsilon > 0$ 是任意的，由此得到 $\gamma_1(\mathcal{A}) \le \beta + \log\ k_{\text{eff}}$，当 $O_d \notin \mathcal{V}$ 时它为 $\beta + \log\ k$，而当 $O_d \in \mathcal{V}$ 时它为 $\beta + \log k_*$。 □

对于 $v \in \mathcal{V}$，将最终形状为 v 的单项式的部分和定义为

$$Z_{n,v} := \sum_{(i_1,\dots,i_n)\in[k]^n:\ \mathit{Shape}\left(L_{i_n}\cdots L_{i_1}\right)=v} A_{1,n}(i_1,\dots,i_n).$$

那么 $X_n = \sum_{v\in\mathcal{V}} Z_{n,\ v}$。

（*b*）的证明。固定 $w \in \mathcal{W}$，令 $s \in \mathcal{L}_{\mathcal{W}}$ 为 w处自环的标签。根据（b）中的附加假设，存在 $r = r(w) \geq 1$，使得对所有 $n \geq r$，形状$(L_s^n) = w$。对于$n \geq r$，令 $A_{1\,n}^{(s)} := A_{n,\ s}A_{n-1,\ s}\cdots A_{1,\ s}$；则 $A_{1\,n}^{(s)}$ 是对 $Z_{n,\ w}$ 有贡献的单项式之一。

根据假设，对所有 $v \in \mathcal{V} \setminus \{w\}$，$Z_{n,\ w}$ 和 $X_n - Z_{n,\ w}$的支撑互不相交，因此

$$\langle Z_{n,w},\, X_n - Z_{n,w}\rangle_F = 0 \quad \text{and hence} \quad \|X_n\|_F \geq \|Z_{n,w}\|_F.$$

在矩阵 $A_{n,\ s}$ 几乎处处逐项非负这一附加假设下，定义 $Z_{n,\ w}$ 的和逐项非负，且不小于其中的每一项。特别地，$Z_{n,\ w} \geq A_{1\,n}^{(s)}$，逐项成立，因此

$$\|Z_{n,w}\|_F \geq \|A_{1,n}^{(s)}\|_F.$$

由于 $\mathcal{M}_d$ 上的所有范数都等价，因此李雅普诺夫指数 β_s 无论用 $\|\cdot\|$ 还是 $\|\cdot\|_F$ 定义都相同，故

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_{1,n}^{(s)}\|_F = \beta_s.$$

合并前面的不等式可得 $\beta_s \leq \gamma_1(\mathcal{A})$。对 $s \in \mathcal{L}_{\mathcal{W}}$ 取最大值即可得到所述结论。

□

5 讨论与展望

最后，我们简要讨论形状图界限应如何解读，以及可以从哪些方面寻求改进。

能量–熵结构与不同情形。 在定理 4.4(a) 所涵盖的无环情形中，上界

$$\gamma_1(\mathcal{A}) \leq \beta + \log k \qquad (\text{or } \beta + \log k_* \text{ when } O_d \in \mathcal{V})$$

可分为两个概念上截然不同的部分。项 β 是一种“能量”贡献：它是主导环余循环的最大李雅普诺夫指数，刻画了沿着能够无限期持续存在的方向所发生的真正指数增长。对数项则是一种“熵”贡献，用于衡量乘积展开中可容许单项式的数量。用转移矩阵语言来说（见上例），β 对应于长期内最佳的局部权重累积，而熵项对应于可容许时间有序路径的增殖。

在许多模型中，熵项有意设置得较为粗略：如果非对角分量较小或较少出现，那么 $\gamma_1(\mathcal{A})$ 预计会接近 β，而定理量化说明，两者的差异不会超过对数组合修正。反过来，当许多非对角分量同时活跃，且其大小与环分量相当时，熵项表明，单纯的组合增殖可能会使顶端指数偏移一个 $O(1)$ 量级。

熵项的可能改进。 正如注 4.7 所强调的，$\log k$项源于对所有 标签序列的计数。一种自然的改进是用与形状图本身相关的有效增长因子（例如谱半径，或施加可达性约束后的最大分支因子）替代 k 。这类改进能使该界反映稀疏模式的几何结构，而不仅仅取决于标签数量。

超越无环形状图。 定理 4.4 中的仅自环假设适用于非对角转移所产生的反馈在结构上不可能出现的情形（例如有向无环图支撑的稀疏性）。允许有向环将带来全新的行为：反复出现的非对角转移可能产生额外的环余循环，并可能以单个 β无法刻画的方式改变指数。一种可能的研究方向是分析形状图的强连通分量，并将本文的方法与诱导分量图上的次可加／金曼型论证相结合。

重叠支撑与抵消。第二个局限是形状集定义中的支撑互不相交要求，这会排除不同分量之间的抵消。放宽这一假设需要谨慎：抵消可能减缓增长，也可能使基于展开的计数论证过于保守。若能找出一些中间类别（例如具有定号元素的矩阵、单调锥或非负矩阵），使其允许支撑重叠，同时组合结构仍然易于处理，将会很有意义。

实际应用。从使用者的角度来看，形状图方法最适合作为一种诊断性上界：对于给定的结构化模型，它能给出淬火增长率的显式、可检验上界，并指出哪些环分量决定了该指数。即使上界并不紧，它仍可用于分析参数标度（例如转移矩阵参数化中对温度 θ 的依赖），并区分增长的结构性来源与概率性来源。

致谢

作者感谢 Alex Roitershtein 富有成效的讨论，以及他对早期稿件的评论，尤其感谢他对命题 2.4 证明所作的贡献。

参考文献

[1] F. C. Adams、A. M. Bloch 和 J. C. Lagarias，《随机 Hill 方程、随机游走与随机矩阵乘积》，载于：A. Johann、H. P. Kruse、F. Rupp 和 S. Schmitz（编），《动力系统最新趋势》，Springer 数学与统计学论文集，第 35 卷，Springer，2013。[2] A. Ahn 和 R. Van Peski，《截断酉矩阵与 Ginibre 矩阵的李雅普诺夫指数》，亨利·庞加莱研究所年刊（B）概率与统计，59（2023），1029–1039）。[3] G. Akemann、Z. Burda 和 M. Kieburg，《Ginibre 矩阵乘积的李雅普诺夫指数普适分布》，《物理学杂志 A：数学与理论》47（2014），395202。[4] J. M. Altschuler 和 P. A. Parrilo，《秩一矩阵的李雅普诺夫指数：遍历公式与最优分布的不可近似性》，《工业与应用数学学会控制与优化杂志》58（2020），510– 528。[5] L. Arnold，《随机动力系统》，Springer，1998。[6] L. Barreira，《李雅普诺夫指数》，Springer，2017。[7] L. Barreira 和 C. Valls，《通过奇异值刻画 Lyapunov 正则性》，Trans. Am. Math. Soc. 369（2017），8409–8436。[8] P. Bougerol 和 J. Lacroix，《随机矩阵乘积及其在 Schr ¨odinger 算子中的应用》，Birkh¨auser，1985。[9] J. E. Cohen 和 C. M. Newman，《大型随机矩阵及其乘积的稳定性》，Ann. Probab. 12（1984），283–310。[10] A. Comtet、C. Texier 和 Y. Tourigny，《随机矩阵乘积与广义量子点散射体》，J. Stat. Phys. 140（2010），427–466。[11] A. Comtet、J. M. Luck、C. Texier 和 Y. Tourigny，《接近单位矩阵的随机矩阵乘积的李雅普诺夫指数》，J. Stat. Phys. 150（2013），13–65。[12] D. Dragi˘cevi´c 和 K. J. Palmer，《通过可容许性刻画三角系统的二分性》，即将发表于《动力系统与微分方程杂志》（2024）。[13] M. Embree 和 L. N. Trefethen，《随机斐波那契数列的增长与衰减》，《皇家学会会刊 A 辑》455（1987），2471–2485。[14] P. J. Forrester，《复高斯随机矩阵乘积的李雅普诺夫指数》，J. Stat. Phys. 151（2013），796–808。[15] P. J. Forrester 和 J. Zhang，《某些各向同性随机矩阵系综的李雅普诺夫指数》，J. Stat. Phys 180（2020），558–575。[16] H. Furstenberg 和 H. Kesten，《随机矩阵的乘积》，《数学统计年刊》31（1960），457–469。

[17] *H. Furstenberg* 和 *Y. Kifer*，随机矩阵乘积与射影空间上的测度，*Israel J. Math. 46*（*1983*），*12–32*。[18] *L. Gerencs´er*、*G. Michaletzky* 和 *Z. Orlovits*，块三角平稳随机矩阵的稳定性，*Syst. Control Lett. 57*（*2008*），*620–625*。[19] *H. Hennion*，独立随机矩阵的不可约乘积的大数定律与扰动，*Z. Wahrscheinlichkeitstheorie verw. Gebiete 67*（*1984*），*265–278*。[20] *R. G. Horn* 和 *C. R. Johnson*，《矩阵分析》，第 *2* 版，剑桥大学出版社，*2013*。[21] *´E. Janvresse*、*B. Rittaud* 和 *T. de la Rue*，广义随机斐波那契数列的几乎必然增长率，*Ann. Inst. H. Poincar´e Probab. Statist. 46*（*2010*），*135–158*。[22] *V. Kargin*，关于高斯矩阵乘积的最大李雅普诺夫指数，*J. Sta.t Phys. 157*（*2014*），*70–83*。[23] *E. Key*，可计算的最大李雅普诺夫指数示例。*Probab. Th. Rel. Fields 75*（*1987*），*97–107*。[24] *E. S. Key*，具有不变子空间的矩阵的李雅普诺夫指数。*Ann. Probab. 16*（*1988*），*1721–1728*。[25] *Y. Kifer*，《随机变换遍历论》，*Birkh¨auser*，*1986*。[26] *G. Letac* 和 *V. Seshadri*，通过连分数刻画广义逆高斯分布，*Z. Wahrsch. verw. Gebiete 62*（*1983*），*485–489*。[27] *R. Lima* 和 *M. Rahibe*，随机矩阵的无穷乘积的精确李雅普诺夫指数，*J. Phys. A: Math. Gen. 27*（*1994*），*3427*。[28] *R. Majumdar*、*P. Mariano*、*H. Panzo*、*L. Peng* 和 *A. Sisti*，与随机斐波那契数列相关的随机矩阵乘积的中心极限定理中的李雅普诺夫指数与方差，*Discrete Contin. Dyn. Syst. Ser A 25*（*2020*），*4779–4799*。[29] *D. Mannion*，随机矩阵的乘积，*Ann. Appl. Probab. 3*（*1993*），*1189–1218*。[30] *J. Marklof*、*Y. Tourigny* 和 *L. Wo-lowski*，随机矩阵乘积的显式不变测度，*Trans. Amer. Math. Soc. 360*（*2008*），*3391–3427*。[31] *T. Mikosch* 和 *D. Straumann*，鞅变换的稳定极限及其在 *GARCH* 参数估计中的应用，*Ann. Statist. 34*（*2006*），*493–522*。[32] *C. M. Newman*，李雅普诺夫指数分布：随机矩阵的精确结果，*Commun. Math. Phys. 103*（*1986*），*121–126*。[33] *N. Otsuka* 和 *T. Shimizu*，任意切换下切换块三角系统的指数稳定性，*Linear Multilinear Algebra 72*（*2024*），*655–677*。[34] *S. Pincus*，随机矩阵乘积的强大数定律，*Trans. Am. Math. Soc. 287*（*1985*），*65–89*。

[35] P.RajandD.Pal，切换微分代数方程稳定性的李代数判据，IFAC-PapersOnLine **53**（2020），2004–2009。[36] E. Strickler，在公共不变面上共享一个零点的随机切换向量场，Stoch. Dynam. **21** （2021），2150007。[37] M. Viana，李雅普诺夫指数讲义，剑桥大学出版社，2014。