arXiv:2304.08334v1 [math.CT] 17 Apr 2023

算术标量与矩阵范畴的大小

Steve
HuntsmanSTR
steve.huntsman@str.us

我们开发了相关工具，用于显式构造以生成数据为富集对象、并通过普通标量与矩阵算术运算进行复合的范畴。我们刻画了有意义的大小映射、权重和大小，它们揭示的特征类似于异常值；此前已证明，在度量空间的语境中，这些概念也能揭示此类特征。文中还提供了相关示例，展示如何将这种“异常值检测”用于分析计算机程序、神经网络、信息物理系统以及通信信道网络。

1 引言

我们考虑一个相当一般的有限状态机或类似结构模型**C**，其中与状态和/或转移相关联的矩阵数据，能够与普通标量和矩阵算术（例如数值程序中的雅可比矩阵、信号处理系统中的滤波器等）协调一致地组合，并具有良好的兼容性和可扩展性。也就是说，**C** 是一个以合适的矩阵范畴**M** 为富集范畴的范畴，[21, 15] 其底层标量是一个特殊情形（并且事实证明，也是普适情形）。标量的一个例子见图 1。

作为一个具有数学典型性的概念性启发示例，考虑由若干相互连接的分量构成的系统，它们具有各种外部输入和输出。由于工程方面的考虑，某些共址输入可能必须沿着穿过系统的不同内部路径传递，但要实现的行为仍必须完全不受所选路径影响。在 [2, 26, 38, 4], 线性时不变情形下，这实质上意味着沿不同路径分别乘以不同矩阵，但在路径汇合的任意位置都产生相同结果。[1] 这相当于将矩阵乘法（或者在傅里叶基下的逐分量乘法）作为幺半群乘积，[27, 15] 置于**M**中。

我们旨在通过类比欧几里得空间的子集来探究**C** 的几何性质；在欧几里得空间中，第 2 节回顾的权重提供了一种出色的、依赖尺度的边界／异常值检测机制 [42, 6, 19]。这种边界检测行为与势理论中的贝塞尔容量概念有关 [29]。

下面几段将对此作进一步说明。对于有限有向图$D = (V,A)$，由D确定的范畴$\langle D\rangle$ *determined by* D 是一个有向图，其对象／顶点为$V(D)$，态射／弧为$A(\langle D\rangle)$，由所有顶点上的自环以及D的传递闭包中的任意弧构成。我们希望以一种方式使用“生成”数据，从而得到一个**M**-范畴**C**，其底层范畴为 $\langle D\rangle$。这种构造可以对无记忆系统建模：在这类系统中，一系列转移所产生的效果只取决于初始状态和最终状态。第 5.3 节的 [8] 曾简要探讨过一种类似构造，其中非确定性自动机的输入带有一个标量“代价函数”。

取**M** 为$M_n(R)$，即由交换环R 上的$n \times n$ 矩阵构成的离散幺半范畴，且乘积由矩阵乘法给出，可用于研究切换线性系统 [26]。取**M** 为 **Arr(FinVect)** 或 **FinStoch**，且乘积由通常的（克罗内克）张量积给出，可用于研究量子电路 [31] 以及离散无记忆信道网络的信息论 [37]。我们

[1] 在可能存在路径依赖的情况下，一种变通方法是考虑D [18]的普遍覆盖空间中的有限子多树。这类似于编译器 [10]中的循环展开。路径无关组合性的失效意味着离散有向基底 [13, 14, 28]上的主丛具有非零曲率（物理学家称之为“场强张量”）。

提交至:
ACT 2023

© S. Huntsman 本作品采用知识共享署名许可协议授权。

依次探讨这些例子，说明如何在与计算机程序分析、神经网络、信息物理系统以及通信信道网络直接相关的结构中找出“异常值”。

我们能够以简洁的方式做到这一点，因为下文构造中涉及的半环 S 具有如下普适性：存在唯一的 S-范畴 $\mathbf{C}$，使得 $Z^{jk} = \mathbf{C}(j,k)$，而后续计算只依赖普通的矩阵算术。[2] 在我们将要讨论的情形中，$\mathbf{M}$ 由维数参数化；维数为一的情形实质上恢复了 S，因此第 2 节引入的大小映射 实质上是一个投影。从这个意义上说，除了确定大小映射之外，考察$\mathbf{M}$-范畴的大小 与S-范畴本身相比，大多并无趣味，而且它通过一个普适情形实现分解。不过，$\mathbf{M}$-范畴的构造 比S-范畴的构造更为微妙，也为应用提供了重要背景。

本文结构如下。第 2 节回顾大小的范畴论表述。第 3 节讨论以普通标量算术为富集结构的范畴的存在性、构造与分析，并给出与计算机程序分析相关的示例，以及为神经网络提供结构信息的示例。第 4 节将这些结果推广到以普通矩阵算术为富集结构的范畴，并给出一个与通信信道网络相关的示例。最后，第 5 节对算术与面向过程的数据相结合作一些推测性评论。

2 大小

设 $\mathrm{M} = (\mathrm{M}\ \otimes, 1)$ 是一个幺半范畴 [27, 15]，C 是一个有限 M-范畴。回顾一下，这意味着 C 由有限集 Ob(C) 指定；对所有 $\mathrm{j}, \mathrm{k} \in \mathrm{Ob(C)}$，都有同对象 $\mathrm{C(j,k)} \in \mathrm{M}$；对所有 $\mathrm{j} \in \mathrm{Ob(C)}$，都有恒等态射$1 \to \mathrm{C(j,j)}$；并且对所有 $\mathrm{j}, \mathrm{k}\ \ell \in \mathrm{Ob(C)}$，都有复合态射 $\mathrm{C(j,k)} \otimes \mathrm{C(k}\ \ell) \to \mathrm{C(j}\ \ell)$，所有这些都满足结合律和单位律性质 [21, 15]。

大小理论 [24, 23] 有两个主要输入。第一个输入是一个$\mathbf{M}$-范畴$\mathbf{C}$。第二个输入是一个大小映射 $\sigma\colon \mathrm{Ob}(\mathbf{M}) \to S$，其中$S$ 是一个半环。[3] 大小映射必须在同构类上为常数，并满足 $\sigma(1) = 1$ 以及 $\sigma(X \otimes Y) = \sigma(X) \odot \sigma(Y)$；其中半环的单位元和乘法如右侧所示。若$n := |\mathrm{Ob}(\mathbf{C})| < \infty$，则其相似度矩阵 $Z \in M_n(S)$ 定义为$Z_{jk} := \sigma(\mathbf{C}(j,k))$。引入（常用的）记号

$$(f[X])_{jk} := f(X_{jk})$$

作为简写，若X 是S 上的矩阵，且f 是S上的函数，则有$Z := \sigma[\mathbf{C}]$。

加权 是满足$Zw = 1$的列向量w，其中标出了半环矩阵乘法和全一列向量。余权重 是Z^T的一个加权的转置。若Z 同时具有加权和余权重，则其大小 是各分量之和（两个和相等）。

目前，大小的示例和应用几乎全部聚焦于*Lawvere* 度量空间，即以 $([0,\infty], \geq)$ 为富集结构的范畴，其中幺半群乘积是普通加法。这类空间也称为广义准伪度量空间，因为它们通过允许距离为无穷大（广义）、不对称（准）或为零（伪），推广了度量空间。据我们所知，目前仅有本文和 [9, 18] 不属于这一研究重点。

[2]回顾一下，半环S 是具有辅助加法结构的离散幺半范畴。[3]在我们的语境中，当$\mathbf{M}$ 是类似$M_n(R)$的半环时，可以——而且可能很有用——令 σ 为恒等映射，使得$S = \mathbf{M}$。

3 标量范畴

令$S \in \{\mathbb{R}, \mathbb{C}\}$，并考虑该系统

$$Z_{jk} Z_{k\ell} = Z_{j\ell} \tag{1}$$

对于 $(j,k),(k\ \ell),(j\ \ell) \in A(\langle D \rangle)$。 [4]当且仅当 Z定义了一个 S-范畴 $\mathbf{C}$，使得 $Z^{jk} = \mathbf{C}(j,k)$ 时，该系统才成立。(1) 的一个解类是

$$Z_{jk} := p_j^{-1} p_k \tag{2}$$

对于 $p : V(D) \to S_{\times}$。我们将证明，(2) 实际上就是一般的非退化情形。由于 §4.2 中将显而易见的原因，确定何时存在整数解将很有帮助。

令 $\tilde{A}(D)$ 表示 D中不是环的弧。令 $\Gamma_2(D) := \{(j,k\ \ell) : (j,k),(k\ \ell) \in \tilde{A}(\langle D \rangle); j \neq \ell\}$为 $\langle D \rangle$ 中长度为二的非退化路径集合。若 $|\Gamma_2(D)| = 0$，则我们可以直接得到 (1) 的所有解，因此不失一般性，假设 $|\Gamma_2(D)| > 0$。引理 1 给出了 (1) 的一般解（见图 1）。

引理 1. 定义一个 $|\Gamma_2(D)| \times |\tilde{A}(\langle D \rangle)|$ 矩阵 M，使其元素除 $M_{(j,k,\ell),(j,k)} = 1$, $M_{(j,k,\ell),(k\ \ell)} = 1$和 $M_{(j,k,\ell),(j\ \ell)} = -1$外均为零。那么 $m := \dim \ker M > 0$，且方程 (1) 的任意解均具有如下形式

$$Z_{jk} = \prod_{i=1}^{m} c_i^{y_{(j,k)}^{(i)}} \tag{3}$$

其中 $\{y^{(i)}\}_{i=1}^{m}$ 是 $\ker M$ 的一组基，且 $(c_i)_{i=1}^{m} \in S^m$。

证明。无妨假设 $\langle D \rangle$ 是弱连通的。若 $(j,k\ \ell) \in \Gamma_2$，则根据复合的传递性，$(j\ \ell) \in \tilde{A}(\langle D \rangle)$，类似地 $|\Gamma_2(D)| < |\tilde{A}(\langle D \rangle)|$。$My = 0$当且仅当 $y_{(j,k)} + y_{(k,\ell)} - y_{(j\ \ell)} = 0$，因此可取 $y_{(j,k)} = \log Z_{jk}$。由于 $|\Gamma_2(D)| < |\tilde{A}(\langle D \rangle)|$，我们有 $m := \dim \ker M > 0$。 □

将简化行阶梯形标准化后，我们得到 $y^{(i)} \in \mathbb{Z}^{|\tilde{A}(\langle D \rangle)|}$。若 $(c_i)_{i=1}^{m} \in \mathbb{Z}_+^m$，则 $Z_{jk} \in \mathbb{Q}_+$。**命题1**。若 D（或等价地，$\langle D \rangle$）是有向无环图（DAG），则系统 (1) 在 $\mathbb{Z}_+$ 上存在非平凡解。更一般地，对于任意环 $\gamma \in \langle D \rangle$，都有 $\prod_{(j,k) \in \gamma} Z_{jk} = 1$，因此 $\mathbb{Z}_+$ 上的任意解在环上的所有弧上都必须等于 1：特别是，若 (j,k) 位于某个环上，则 $Z_{jk} = Z_k^{-1} j$。

我们现在要为弧和顶点上的生成数据求出闭式。令U 为从箭图到无向图的函子，它会忘掉弧方向和重数。对于 $(i,i') \in V(D)^2$,，记作

$$\varepsilon(i,i') := \begin{cases} 1 & \text{if } (i,i') \in A(D) \\ -1 & \text{if } (i',i) \in A(D) \text{ and } (i,i') \notin A(D) \\ 0 & \text{otherwise} \end{cases}.$$

定理 1。设 D 为弱（= 弱连通），T 是 $U(D)$ 的一棵生成树，且 W： $E(T) \to S$ 非退化，其中与通常情况一样，$E(\cdot)$表示无向图的边。赋值

$$Z_{jk} = \prod_{(i,i') \in T[j,k]} W(i,i')^{\varepsilon(i,i')}, \tag{4}$$

其中，路径 $T\ [j,k]\ in\ T$ 中从 j 到 k 的边的乘积定义良好，并满足(1)。如果 D 是有向无环图，那么 $W : E(T) \to \mathbb{Z}_+$ 映射到 $Z_{jk} \in \mathbb{Z}_+$。最后，(1) 的任意非退化解都具有(4)的形式。

[4]若记 $d := \tau \log[Z]$ 为适当的标量 τ，则 (1) 呈现为三角等式 $d_{jk} + d_{k\ell} = d_{j\ell}$，这凸显了它与满足类似于 Menger 凸性判据的 Lawvere 度量空间“原味”大小之间的相似性 [25]。还有其他相关概念：当且仅当从 j 到 ℓ的每条路径都经过 k [22, 12, 7]时，称加权无向图为图测地（分别为割点可加）。

证明。 赋值 (4) 定义良好，因为$T[j,k]$存在；它是唯一的，因为D是弱的。该赋值满足 (1)，因为$T[j,k]$与$T[k\ \ell]$的拼接就是$T[j\ \ell]$。如果D是有向无环图，那么 $\varepsilon \equiv 1$。最后，根据传递性，(1) 的任意非退化解都具有 (4) 的形式。 □

定理 2. 若 D 是弱的，则式(1) 的任意非退化解都具有形式 $Z_{jk} = p_j^{-1} p_k$，其中 p 任意。 [5]

证明。 设Z 满足 (1)，并设P 是D的一个生成多树，即一个生成有向图，使得$T := U(P)$ 是一棵树。取$i \in V(D)$，并令$p_i = 1$，随后通过遍历P并在$A(P)$ 上求解$Z_{jk} = p_j^{-1} p_k$，将p扩展到$V(P) = V(D)$。最后，在T上取$W(j,k) := Z^{\varepsilon(j,k)}{}_{jk}$，并应用定理 1 还原$Z$。 □

例 1. 图 *1* 展示了一个例子；鉴于 *[5]* 中命题 *2.3.1* 对传递有向图的结构刻画，该例本质上具有一般性。然而，它的大小没有定义。

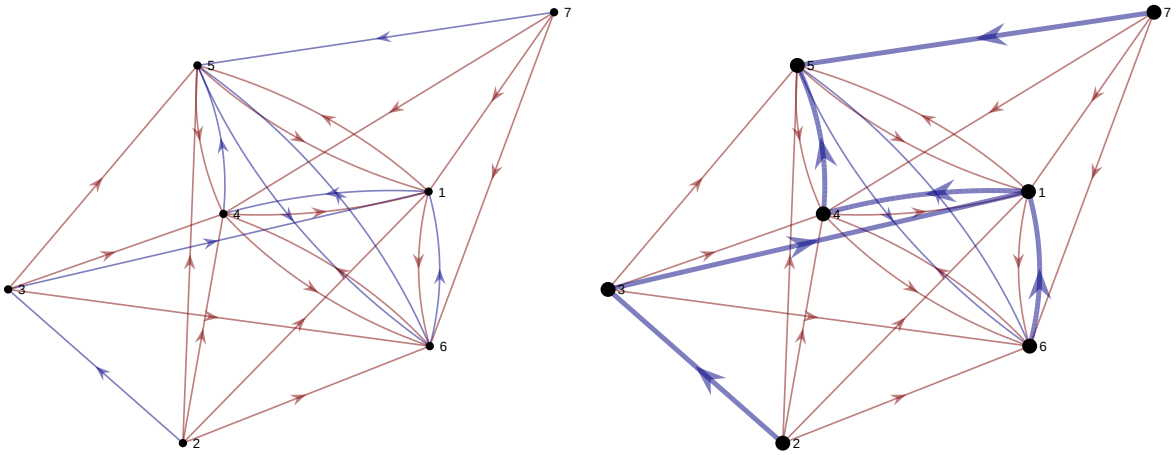

图 1：（左）在由（蓝色）底层有向图生成的传递有向图（红色）上，形式为（3）的（1）的一个符号解。（右）使用赋值（4）为**高亮的生成（多）树**生成的（1）的一个特定解。取$p \propto (14,\ 1,\ 7,\ 42,\ 462,\ 14/5,\ 462/13)$得到$Z_{jk} = p_j^{-1} p_k$。

（1）的最简单解无法呈现有趣的结构。

引理 2。 如果对于所有（j, k），$Z\,jk = Z_k^{-1} j$，那么除非 $Z_{jk} \equiv 1$，$\mathbf{C}$ 的大小没有定义。

证明。 根据假设$Z_{j\ell} = Z_{jk} Z_{k\ell} = Z_k^{-1}{}_j Z_{k\ell}$对于所有 ℓ。由于j行和k行Z互为常数倍，因此方程$Zw = 1$ 仅在$Z_{jk} = 1$时有解。 □

类似的考量也表明，如果D 含有一个环（且不是自环），并且$\mathbf{C}$ 有大小，那么其大小必定为 1。不过，加权空间仍然编码了关于$\mathbf{C}$的非平凡信息。

例 2. 考虑一个按如下方式构造的玩具程序。我们使用概率上下文无关文法 *[40]*中的产生式生成一个程序“骨架”。

$$S \to S;\ S \mid \texttt{if}\ b;\ S;\ \texttt{fi} \mid \texttt{while}\ b;\ S;\ \texttt{end}$$

[5]如果B 是入射 D 的稠密矩阵，则 log[向量化 Z] $= B^T \log[p]$, 其中，向量化会将矩阵列依次堆叠，且 p 被视为一个向量 or.

表 1：控制流图弧： $[\cdot]$:= 匹配标记的行号。

源位于j	目标(⊤)	目标(⊥)
`if b`	$j+1$	[fi]+1
当 b	$j+1$	[结束]+1
`end`	[当]	·
fi 或 S	$j+1$	·

其中 S 是行分隔符的简写：产生式概率分别为 0.6、0.1 和 0.3。词元 S 和 b 分别表示语句/子程序和布尔谓词。

接下来，我们根据表 *1* 将顶点与骨架中的行相关联，并按表中的规则确定边，从而构造出最终的控制流图 *[10]* 。我们还在程序和控制流图的开头各添加一条 *START*行/顶点/弧，并在末尾各添加一条 *HALT*行/顶点/弧。我们可以通过以下方式显式实例化一个可执行程序：*i)* 将程序第 j 行上的标记 S 替换为显式语句；*ii)* 将程序第 k 行上的标记 b 替换为显式谓词。

考虑将标量数据赋给控制流图中的弧，如下所示：

- 形如$(Sj \cdot j_{+1})$的弧被赋值为 $\delta j \in S$；
- 所有其他不形如$(if,\cdot_{[fi]+1})$或$(end, while_{[while]})$的弧，均被赋值为 $1 \in S$；
- 形如$(if,\cdot_{[fi]+1})$的弧，被赋值为连接其源与目标的任意其他路径上所赋数据的（有序）乘积；
- 形如$(end, while_{[while]})$的弧，被赋值为连接其目标与源的任意其他路径上所赋数据的（有序）乘积的逆元。

命题 **关于 2**。上述赋值定义良好，并且唯一对应一个 S-范畴 . □

考虑图*2*中所示的由 20 个上下文无关文法产生式得到的控制流图，以及对所有 j 的赋值 $\delta_j = 2$ 。所得相似度矩阵及其核分别显示在图*3*的左、右两幅图中。将向量$(1, 0, \ldots, 0)^T$加到核上，即可得到加权空间：该向量对应于 *START*顶点。余加权空间与此类似（未显示）。核中的基向量之和均为零，因此大小始终为 *1*。

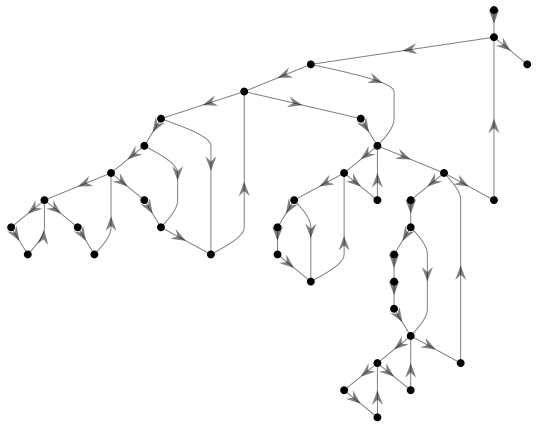

图 2：一个简单的控制流图。

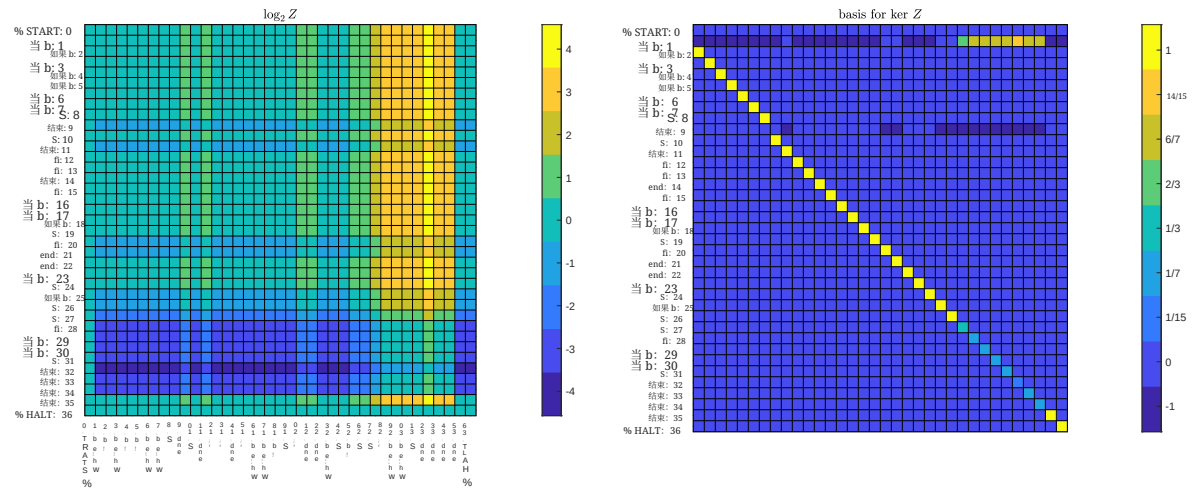

图3：左图：由赋值 $\delta_j = 2$ （对所有 j ）作用于图2所示控制流图对应的相似度矩阵 Z 的 $\log_2$ 。右图：Z 核的一个有理基，向量各分量的最大绝对值归一化为 1。注意，配色方案经过量化且为非线性。

现在假设我们将其中一个 δ_j 改为等于 8 ，而非 2。所得核（各行的最大值）显示在图*4*左侧的各幅图中；右侧的图与之类似，但其中有 $\delta_j = 1/4$。尽管大小始终为 *1*，加权空间编码了有关程序“几何”的全局上下文信息。

如果大小为正，取对数便得到一个不相似度，从而基本上可将问题进一步归约为弧加权有向无环图的普通 Lawvere 度量大小情形。事实上，我们可以利用这一点，对更一般的有向图获得一些直观认识。假设某个生成（多）树上的所有大小都满足 $\sigma = \exp(-1)$：那么，除去与可达性有关的退化情形，有 $Z = \exp[-d]$ ，其中 d 是有向图距离。类比欧几里得距离矩阵，我们因此预期，所得的加权空间会标示出那些因“位于边缘”而显得“较大”的顶点；这种特征有时会、但并非总是与度明确相关。

例 3。 令 $K^{\rightarrow}_{n_1,\ldots,n_L}$ 表示具有 $V(K^{\rightarrow}_{n_1,\ldots,n_L}) := \bigsqcup_{L}{}_{\ell=1} K_\ell$ *for* $K_\ell := [n_\ell]$ *and* $A(K^{\rightarrow}_{n_1,\ldots,n_L}) := \{(v, v') : v \in K_\ell, v' \in K_{\ell+1}; \ell \in [L-1]\}$ 的有向无环图。该有向无环图对应于具有 L 层、宽度为 $n_1, \ldots, n_L$ *[17]* 的全连接多层感知机 *(MLP)* 的架构。

现在考虑一个子有向无环图 $D \subset K^{\rightarrow}_{n_1,\ldots,n_L}$，其弧权重为 $\sim U([0,1])$，对应于一个稀疏连接的 *MLP*。固定 D，并保留 $U(D)$ 中对应于随机生成树的弧权重，[6] 便定义了 Z，从而也定义了作为随机变量的 w，如图*5*右侧所示，其中包含 $N = 500$ 个实现，对应于左侧的带权有向无环图。

w 在统计上表现良好：在我们的实验中，根据安德森-达林检验 *[34]*，w 的各个分量均满足其取自正态分布的假设，且显著性水平均达到标准计算实现所提供的最佳水平。[7] w 的分量均非平凡，唯一的例外是对应出度为零的顶点的分量，这些分量具有

[6]我们通过求取$U(D)$的最小生成树，并加入临时边权重 $\sim U([0,1])$，来生成一棵随机生成树。这种生成树通常不是均匀随机的，但可以生成此类树 [1]。[7]按照 [41] 所述方法进行多元正态性检验，计算开销极大；由于维数较高，在我们的情境下，这类检验也无法得出近乎同样确凿的结论。

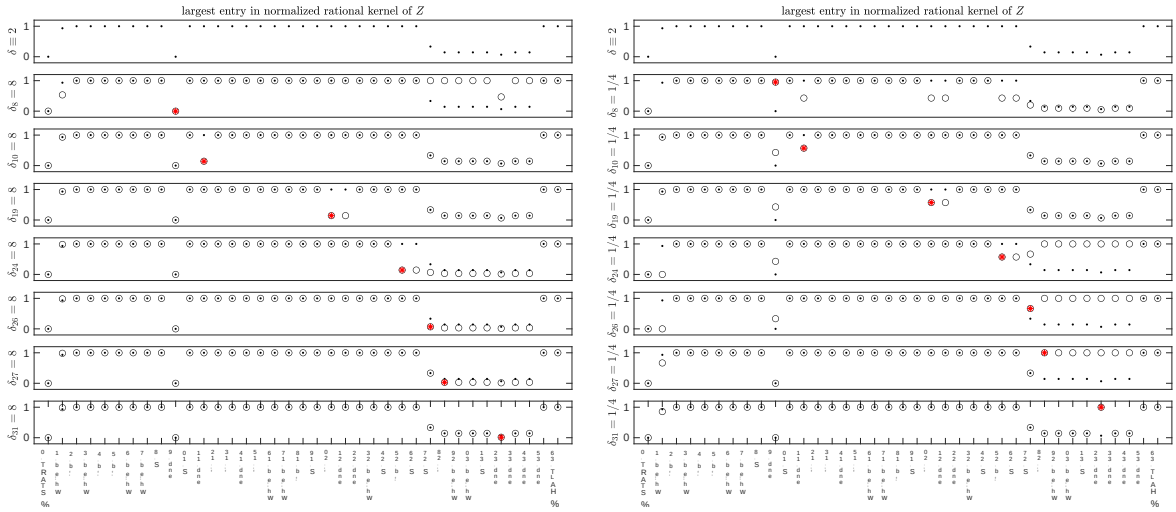

图4: （左上图）图3中 Z 的核各行的最大值。（左下图）与上图相同，但还显示（用 ○） $\delta_j = 8$ 在 $j = 8$、10、19、24、26、27、31 时的最大值。$j+1$ 对应的最大值用红色 *标出。（右侧图）与左侧图相同，但取 $\delta_j = 1/4$。

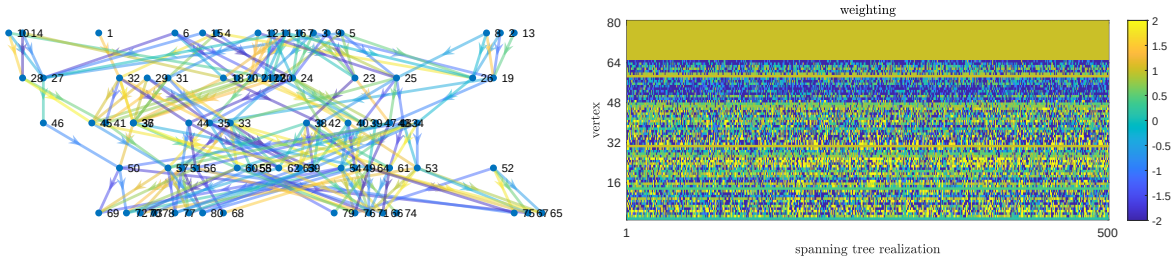

图5: （左）一个带弧权重 $\sim U([0,1])$ 的 $K_{16}^{\rightarrow}$，16，16，16，16 的加权子有向无环图，颜色采用与右图相同的（重新缩放后的）配色方案。（右）在随机生成树上生成数据后得到的 $\mathbb{R}$ 类别加权方式的 $N = 500$ 个实现。为清晰起见，颜色轴已截断。w 的统计规律性以水平条纹的形式显现。

单位值；入度为零的顶点对应的分量各自具有固定值，而其他顶点对应的分量则服从正态分布。相邻顶点往往具有符号相反的加权分量，这与欧几里得空间中的一般直觉一致： *[42, 6, 19]* 负加权分量往往出现在具有较大正加权分量的“局部边界正后方”。

在较大的网络中，推测的“近离群”顶点（其加权分量在统计上最小和最大）往往连接得很密集。具体而言，考虑一个集中趋势度量 $\mathbb{T}$ （例如均值、中位数等）以及 $t_- < t_+$，使得集合 $T_- := \{j : \mathbb{T}(w_j) < t_-\}$ 和 $T_+ := \{j : \mathbb{T}(w_j) > t_+\}$ 中各只有少数几个顶点。现在令 $T := T_- \cup T_+$，考虑集合

$$X := T \cup \big(\{i : (i,j) \in A(D); j \in T\} \cap \{k : (j',k) \in A(D); j' \in T\}\big).$$

由 X 诱导的子有向无环图通常连接非常密集，不过这个子有向无环图本身也可能存在噪声。图6 展示了一个基于图5构建的示例；我们在其他示例中也观察到了这种行为。

总的来说，这些结果表明，在神经网络的正则化和/或剪枝策略中强制要求权重具有这种组合性，或许能像随机失活一样发挥作用 *[17]*。

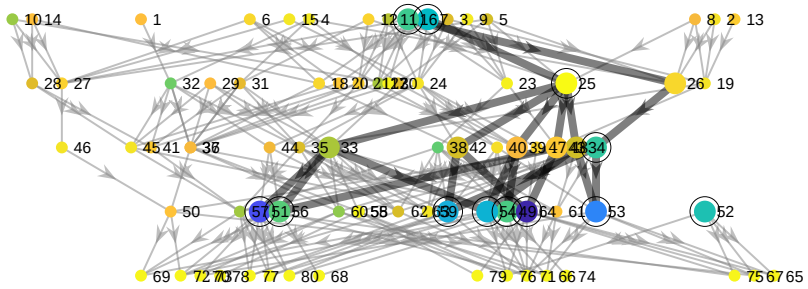

图6：图5中的有向无环图在 X for $\mathbb{T}=$ 中位数和 $(t_-,t_+)=(0.14,0.99)$所诱导的子图。顶点按中位数加权着色；X 较大，T 则以圆圈标出。

当 D 是多树时，上述行为会清晰无误地显现出来（显著子多树中有许多不连通分量的情况除外；这些分量源于明显且与此无关的结构）。图7展示了一个二叉多树上的加权，其中弧权重的一种实现为 $\sim U(\{1,2\})$。[8]

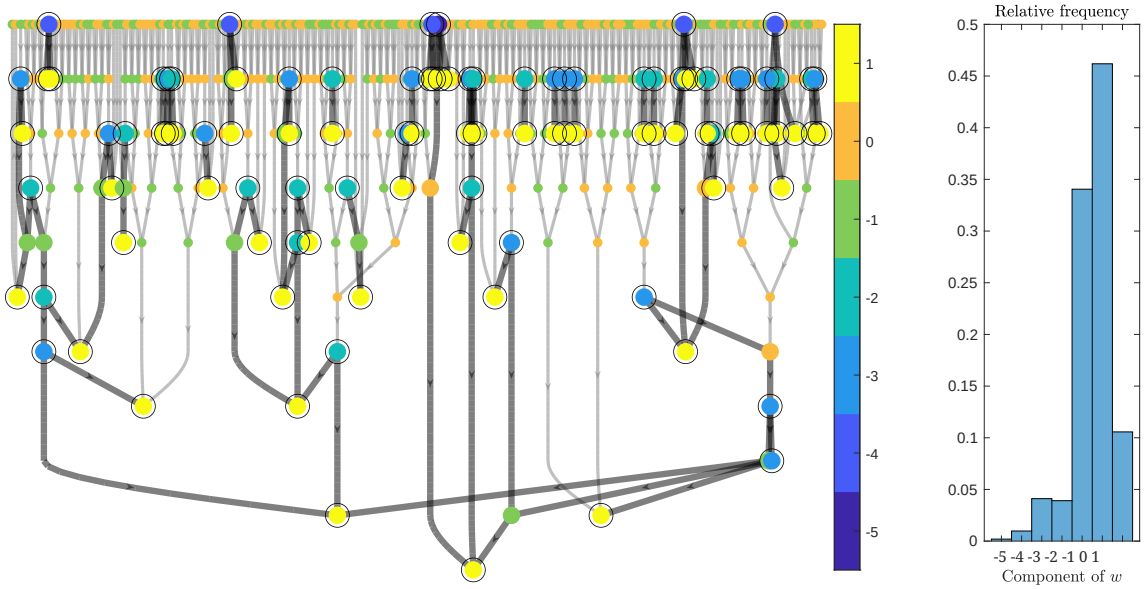

图7：（左）与图6类似，展示一个弧权重为 $\sim U(\{1,2\})$ 和 $\{j\colon w_j<-1\}$ and $\{j\colon w_j>0\}$ 的多树，其中它们分别对应于 T_- 和 T_+。（右）加权分量的相对频率。

4 矩阵范畴

矩阵范畴是一种同态集为矩阵的富范畴。具体而言，矩阵范畴是满足复合相干条件的有向图*qua* 箭图 [36] 的一个表示。[9]

[8]结果表明，只要权重是通过量化同一个伪随机数生成器的输出得到的，取弧权重 $\sim U([n])$ 时，对于任意 $n\geq 2$，所得结果经过仿射缩放后都大致相同。[9]箭图表示范畴包含一个有向图表示子范畴，而有向图又对应一个矩阵范畴范畴。这些对象的表示论本身或许就很有趣。

这种构造有几种彼此不等价的形式，不过从大小的角度来看，它们都可归结为§3中的普适构造。

4.1 矩阵乘法作为幺半群乘积

我们首先考虑一种有限状态机模型，其中线性映射与状态和／或转移相关联。该模型适用于控制理论和／或信息物理系统分析中的许多情形，例如切换线性系统 [26]。

对于 $n < \infty$，我们可以将交换环 R 上的 $n \times n$ 矩阵所构成的幺半群 $M_n(R)$ 视为一个幺半范畴，其对象为 $M_n(R)$，仅含恒等态射，幺半群乘积则为通常的矩阵乘法。经过适当的重新诠释（尤其是采用有序乘积），赋值 (4) 定义了一个 $M_n(R)$-范畴，而且任意有限的 $M_n(R)$-范畴都可以用这种方式实现：类似地，对于合适的 p，我们可以写作 $\mathbf{C}(j,k) = p_j^{-1} p$<v45>。图8展示了 $SL_2(\mathbb{Z})$ 的示例：这些示例便于书写，但从大小的角度来看并无实质内容。

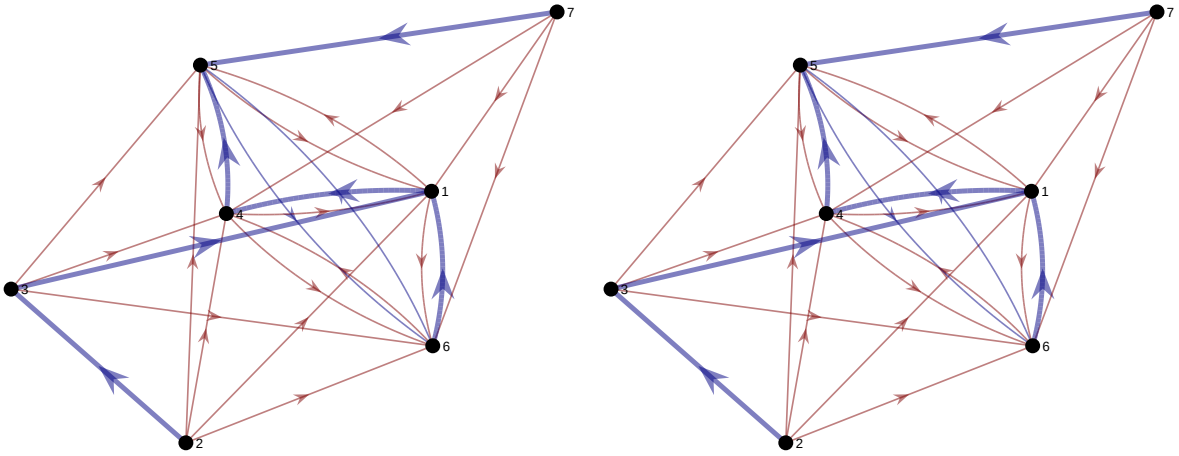

图8：（左）与图 1 右侧面板相同，但标注的是形如 $\mathrm{u(x)} := \left(\begin{smallmatrix} 1 & \mathrm{x} \\ 0 & 1 \end{smallmatrix}\right)$ 的矩阵的 (1,2) 项。请注意，这实质上将标量加法用作幺半群乘积，而且此处涉及的矩阵行列式均为单位行列式。取 $\mathrm{p} = \mathrm{u}^{\times 7}(9{,}0{,}7{,}12{,}23{,}4{,}10)$，得到 $\mathrm{C(j,k)} = \mathrm{p}_j^{-1}\mathrm{p}_k$。（右）$\mathrm{SL}_2(\mathbb{Z})$ 上一个更复杂的例子：此处取 $\mathrm{p} = \left(\left(\begin{smallmatrix} 3 & 10 \\ -1 & -3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 7 & 11 \\ -2 & -3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 19 & 35 \\ -6 & -11 \end{smallmatrix}\right), \left(\begin{smallmatrix} 384 & 457 \\ -121 & -144 \end{smallmatrix}\right), \left(\begin{smallmatrix} -11 & 29 \\ 3 & -8 \end{smallmatrix}\right), \left(\begin{smallmatrix} 165 & -587 \\ -52 & 185 \end{smallmatrix}\right)\right)$，得到 $\mathrm{C(j,k)} = \mathrm{p}_j^{-1}\mathrm{p}_k$。

与此同时，行列式提供了合适的大小映射，因此我们可以应用 §2 的构造。[10] 就相似矩阵、权重和大小而言，我们可以用它们的行列式替代矩阵，并应用 §3 的结果，而不必显式构造一个 $M_n(R)$-范畴的全部数据。也就是说，从这个角度看，考虑矩阵而非 标量并不会带来实质上的收获：$n > 1$ 因子情形可归结为$n = 1$因子情形。

[10] 更一般地，任何性质良好的大小映射都可以经由行列式分解。还要注意，行列式也以一种借鉴本节和 §4.2 的方式，为 Arr(域) 或 Arr(环) 上的富化提供依据：在这些情况下，大小映射由域扩张的范数或理想的范数给出 [20]。

4.2 克罗内克积作为幺半群乘积

此外，§4.1 还有一个变体，采用通常的（即 Kronecker 张量）幺半群乘积。最显而易见的潜在应用是量子电路 [31] [11]，不过，正如我们将看到的，大小的考量又一次归结为 §3 中的标量情形，因此我们将示例推迟到 §4.3。

回想一下，范畴 **X** 的箭头范畴 Arr(**X**) 的对象是 **X** 的态射；态射是 **X** 中的交换方块；复合则由这些交换方块的串接给出。如果 **X** 是幺半范畴，那么 Arr(**X**) 也是幺半范畴。[12] 特别地，对于 $S \in \{\mathbb{R}, \mathbb{C}\}$，Arr(**FinVect**$_S$) 的对象是 S 上的矩阵（即可以表示为矩阵的线性映射），[13] 而 Arr(**FinVect**$_S$) 关于通常的张量积 $\otimes$构成幺半范畴，即富集复合律为

$$\mathbf{C}(j,k) \otimes \mathbf{C}(k,\ell) = \mathbf{C}(j,\ell). \tag{5}$$

接下来，我们构造底层范畴由有限有向图 $D = (V, A)$ 确定的 Arr(**FinVect**$_S$)-范畴 **C** 的空间。与 §4.1 中一样，**C** 的同对象是线性映射；但由于这里的幺半群乘积是通常的张量积，这些矩阵的大小会有所不同，除非它们全都为 1×1。记 $s_{jk} := \dim \mathrm{dom}\, \mathbf{C}(j,k)$，并记 $t_{jk} := \dim \mathrm{cod}\, \mathbf{C}(j,k)$。那么 s 和 t 必须满足 (1)。

If (j、k) 和 (k ℓ) 属于同一个环，那么 (j ℓ) 也属于该环，由此得到一个简单命题 .

命题 3. 系统 (1) 在 $\mathbb{Z}_+$ 上不存在这样的解：其在能够到达 D 的强连通分量的弧上非平凡。

□

注意，除非 D 是有向无环图，否则它必然有非平凡的强连通分量。除非在退化情形下，我们可以对适当尺寸的矩阵进行“克罗内克除法”，从而得到以下推论。

推论 1. 若 D 是有限有向无环图，P 是 D 的一个生成多树（即在 U 下的像为树的有向图），且 s、t 在 $\mathbb{Z}_+$ 上满足 (1)，则一个 $Arr($**有限维向量空间**$_S)$-范畴由 $\prod_{(j,j') \in A(P)} M_{(s_{(j,j')}, t_{(j,j')})}(S)$ 的一个非退化元素确定；并且每个底层范畴由 D 确定的 $Arr($**有限维向量空间**$_S)$-范畴都以这种方式得到。

在基域 $S = \mathbb{C}$ 上的当前设定中，同时也是范数的大小映射通常称为交叉范数。矩阵在 $\mathbb{C}$ 上的置换不变大小映射恰好是 Schatten$_p$-范数 $\|\cdot\|_p$，即矩阵 [3] 的奇异值组成的向量的 ℓ_p 范数。[14]

与 §4.1 一样，将弧矩阵替换为大小后，便又一次归结为 §3 中的标量情形。

4.3 随机矩阵

§4.1 和 §4.2 中的构造可以直接推广到相关矩阵为（行）随机矩阵的情形。[15] 不存在一种信息论大小映射，能够取代行列式

[11]尽管从大小的角度来看，我们对酉矩阵的构造变得平凡，但我们可以考虑通过矩阵对数得到的厄米矩阵。这要求改为考虑由所谓的克罗内克和定义的幺半群乘积 $H \boxplus H' := H \otimes I_{H'} + I_H \otimes H'$，它满足 $\exp(H) \otimes \exp(H') = \exp(H \boxplus H')$。这里我们特别指出，所示的是矩阵指数 $\exp(\cdot)$，而不是逐分量指数 $\exp[\cdot]$。[12]例如，参见 [27] 第 165 页练习 4，其中给出了一个更一般的结果。[13]我们的一些论述适用于半环（= rig）上的半模范畴，但其中的关键部分要求在 $S \in \{\mathbb{R}, \mathbb{C}\}$ 的情况下使用 $\mathrm{FinVect}_S$。[14]对于 $p = 1$、2、∞，相应的范数分别称为迹范数/核范数、Frobenius 范数/希尔伯特-施密特范数，以及算子范数/谱范数。事实上，$\|S\|_\infty \le \|S\|_2 \le \|S\|_1$。相关理论也深受格罗滕迪克 [33] 和冯·诺依曼 [35]的影响。[15]不过请注意，非负随机矩阵的逆元必定含有负项，因此在这两种情况下，显而易见的应用对象都只有有向无环图。FinStoch 范畴的对象和态射分别是有限集和由行随机矩阵表示的随机映射，相关讨论见 [16]。

来构造 §4.1 的类似结果，正如脚注10所指出的那样。然而，存在一种有意义的信息论大小映射，可以产生 §4.2 的非平凡类似结果，即 $\mathbf{C}(j,k) \mapsto \exp C(\mathbf{C}(j,k))$，其中 C 表示信道容量 [37]。[16]

例4。离散无记忆信道（DMC）由条件概率矩阵 $W_{jk} = \mathbb{P}(y_k|x_j)$指定，其中 xj 和 y_k分别表示随机变量 X 和 Y 的实现所对应的输入和输出符号。W_{jk} 表示：如果爱丽丝发送 x_j，则鲍勃收到 y_k 的概率。输入分布 $p\,j = \mathbb{P}(x\,j)$ 可得联合概率 $V_{jk} := \mathbb{P}(x_j, y_k) = p\,jW_{\ jk}$。信道编码定理指出，当信息比特与传输比特之比小于信道容量 $C := \sup_p I(X;Y)$时，当且仅当该条件成立，才能实现可靠通信；其中 $I(X;Y)$ 表示互信息。

给定 D，其弧数据如图9所示，$W^{(i)}$ 表示具有相应信道容量 $\log c_i$的离散无记忆信道的矩阵。由此得到的相似度矩阵为

$$Z = \begin{pmatrix} 1 & 0 & c_1c_3 & c_1 & c_1c_3c_4 & c_1c_3c_5 \\ 0 & 1 & c_2c_3 & c_2 & c_2c_3c_4 & c_2c_3c_5 \\ 0 & 0 & 1 & 0 & c_4 & c_5 \\ 0 & 0 & 0 & 1 & c_3c_4 & c_3c_5 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

方程 $Zw = 1$ 可手动求解，得

$$w = (1 - [1 - c_4]c_1c_3 - [1 - c_3c_5]c_1, 1 - [1 - c_5]c_2c_3 - [1 - c_3c_4]c_2, 1 - c_4 - c_5, 1 - c_3c_4 - c_3c_5, 1, 1)^T.$$

$w_3 = 1 - c_4 - c_5$与 $w_4 = 1 - c_3c_4 - c_3c_5$在形式上存在不对称性。这反映出，尽管端到端容量与路径无关，但到达节点 3 的信道容量与到达节点 4 的信道容量不同。请注意，如果将来自顶点 2 的信道彼此交换，并将到达顶点 6 的信道彼此交换，这种不对称性就会消失。

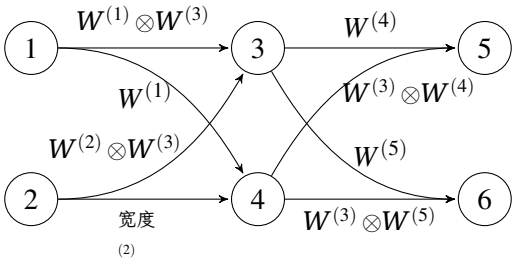

图9：一个**FinStoch**-范畴（为清晰起见，省略非平凡复合）　　).

[16]关于信息论的背景知识，参见 [11]。B. Fong 和 D. I. Spivak 在私人通信中（2019年）证明，信道容量是一个严格幺半群松弛函子（**FinStoch** $\otimes$） $\to$ ([0,∞],(∞,min) $\geq$,(+,0))；其中，我们用一个对象 $*$ 表示该松弛幺半偏序幺半群，其 Hom(∗,∗) 为 $= [0,\infty]$，恒等态射为 ∞，复合由 min 给出，局部偏序（po）结构由通常的 $\geq$给出，而幺半结构由 $+$ 和 0 给出。当将 i）**FinStoch** 上的克罗内克积/张量积 $\otimes$ 幺半结构替换为直和 $\oplus$ 幺半结构，以及 ii）将信道容量C 替换为其指数 $\exp C$ 时，这一结果同样成立。

4.3.1 关于信道容量与余权重的边栏说明

为了理解并在非常规情境中应用大小，不妨考虑线性方程 $Zw = 1$ 和/或 $vZ = 1^T$，其中差异矩阵 Z 无法对任何 Lawvere 度量 d写成 $\exp[-td]$ 的形式。例如，具有可逆信道矩阵W 的离散无记忆信道的达到容量的输入分布，可以视为这样的归一化余权重。理解这一点的关键在于 Muroga [30]提出的一个经典公式。

设 W 是一个可逆的信道矩阵，且 $\mathrm{M} := \mathrm{W}^{-1}$，并定义 $\mathrm{H}_j := -\sum_k \mathrm{W}_{jk} \log \mathrm{W}_{jk}$。此外，定义 $\mathrm{v}_j := \sum_i \mathrm{M}_{ij} \exp\left(-\sum_k \mathrm{M}_{ik}\mathrm{H}_k\right)$。Muroga 公式指出，若 $\mathrm{v} > 0$，则 $\mathrm{e}^C = \sum_j \exp\left(-\sum_k \mathrm{M}_{jk}\mathrm{H}_k\right)$，且 $\mathrm{p} := \mathrm{e}^{-C}\mathrm{v}$ 是一个达到容量的分布。现在 $\mathrm{v} = \exp(-\mathrm{MH})\mathrm{M} = 1^T \Delta(\exp(-\mathrm{MH}))\mathrm{M}$，因此令 $\mathrm{Z} := \mathrm{W}\Delta(\exp(\mathrm{MH}))$，可得 $\mathrm{vZ} = 1^T$。然而，所假设的距离 $-\mathrm{t}^{-1}[\log \mathrm{W}_{jk} - (\mathrm{MH})_k]$ 在对角线上通常为负，但对于“良好信道”$\mathrm{W} \approx \mathrm{I}$，它也近似等于某个常数乘以 $1 - \delta_{jk}$。离散无记忆信道越好，其诱导的“度量几何”（由于对角线为负，所以这里加上引号）就越接近正则单纯形的度量几何。

5 关于增添数据的“类流”图的评论

一种有趣但仍属推测的可能性是，将其更深入地应用于过程、计算机程序等领域时，可以考虑形如 $\mathbf{F} \times \mathbf{M}$ 的范畴，其中$\mathbf{F}$ 是由具有单一输入和输出的“类流”有向图构成的适当范畴，而$\mathbf{M}$ 是由数据构成的适当范畴（如果不采用非退化矩阵范畴的思路，也可以是标量）。$\mathbf{F}$ 上的幺半结构很简单：沿着 [18] 所述的方式，将一个“类流”有向图的输出与另一个的输入粘合起来。也就是说，每个“类流”图都关联着一些数据，而这些数据在串联复合下能够很好地组合。通过对双端图 [39]这一概念作少量技术性修改，很可能就能得到一种特定的$\mathbf{F}$ ，使其在算法和范畴论方面都具有良好的性质（尤其是半笛卡尔性）。

为了在此背景下得到有用的大小概念，必须具备一个大小函数（或者，对于大小同调 [25], , 需要一个从半笛卡尔对称幺半范畴到适当的阿贝尔对称幺半范畴的强对称幺半函子），以整合图数据和矩阵数据。在这方面，一个很有希望的构造方案似乎是有限马尔可夫链的 ζ 函数 [32] , 它编码了（例如）图遍历概率。

致谢

本研究由国防高级研究计划局（DARPA）资助开展。文中表达的观点、意见和/或研究结果均为作者所有，不应被视为代表美国国防部或美国政府的官方观点或政策。分发声明“A”（批准公开发布，不限分发范围）。

参考文献

[1] David J Aldous (1990)：均匀生成树和均匀标号树的随机游走构造。《*SIAM*离散数学杂志》3(4)，第 450–465 页。[2] Panos J Antsaklis & Anthony N Michel (1997)：线性系统。施普林格。[3] Guillaume Aubrun & Ion Nechita (2011)：乘法性质刻画 ℓ_p 和 L^p 范数。*Confluentes Mathematici*3(04)，第 637–647 页。

[4] Georgios Bakirtzis、Cody H Fleming 和 Christina Vasilakopoulou（2021）：信息物理系统理论的范畴语义。《ACM信息物理系统汇刊》5(3)，第1–32页。[5] Jørgen Bang-Jensen 和 Gregory Z Gutin（2008）：《有向图：理论、算法与应用》。施普林格。[6] Eric Bunch 等（2020）：度量空间大小与加权向量的实际应用。参见 https://arxiv.org/abs/2006.14063。[7] Pavel Chebotarev 和 Elena Deza（2020）：命中时间拟度量及其森林表示。《优化快报》14(2)，第291–307页。[8] Simon Cho（2019）：量度、持久性与大小同调。arXiv 预印本 arXiv:1910.02905。[9] Joseph Chuang、Alastair King 和 Tom Leinster（2016）：关于有限维代数的大小。《范畴理论与应用》31(3)，第63–72页。[10] Keith D Cooper 和 Linda Torczon（2011）：《编译器工程》。爱思唯尔。[11] Thomas M Cover 和 Joy A Thomas（2006）：《信息论基础》。Wiley-Interscience。[12] Michel Marie Deza 和 Elena Deza（2009）：《距离百科全书》。施普林格。[13] Aristophanes Dimakis 和 Folkert Muller-Hoissen（1994）：有限集上的微分演算与规范理论。《物理学杂志A：数学与通论》27(9)，第3159页。[14] Aristophanes Dimakis 和 Folkert M¨uller-Hoissen（1994）：离散微分演算：图、拓扑与规范理论。《数学物理学杂志》35(12)，第6703–6735页。[15] Brendan Fong 和 David I Spivak（2019）：《应用范畴论导论：组合性七讲》。剑桥。[16] Tobias Fritz（2020）：关于马尔可夫核、条件独立及充分统计量定理的综合方法。《数学进展》370，第107239页。[17] Ian Goodfellow、Yoshua Bengio 和 Aaron Courville（2016）：《深度学习》。MIT。[18] Steve Huntsman（2022）：有向图的大小与拓扑熵。载于：《应用范畴论会议论文集》。[19] Steve Huntsman（2023）：通过大小提升多样性。载于：《进化多目标优化会议论文集》。[20] Gerald J Janusz（1996）：《代数数域》。美国数学会。[21] Gregory Maxwell Kelly（1982）：《富化范畴论基本概念》。剑桥。[22] Douglas J Klein 和 H-Y Zhu（1998）：图的距离与体积。《数学化学杂志》23(1-2)，第179–195页。[23] Tom Leinster（2021）：《熵与多样性：公理化方法》。剑桥。[24] Tom Leinster 和 Mark W Meckes（2017）：度量空间的大小：从范畴论到几何测度论。载于 Nicola Gigli 编：《非光滑空间中的测度论》，德格鲁伊特，doi:10.1515/9783110550832-005。[25] Tom Leinster 和 Michael Shulman（2021）：富化范畴与度量空间的大小同调。《代数与几何拓扑》21(5)，第2175–2221页。[26] Daniel Liberzon（2003）：《系统与控制中的切换》。Birkh¨auser。[27] Saunders Mac Lane（2013）：《工作数学家的范畴论》。施普林格。[28] Juan Maldacena（2015）：物理定律与希格斯玻色子的对称性和简洁性。《欧洲物理学杂志》37(1)，第015802页。[29] Mark W Meckes（2015）：度量空间的大小、多样性、容量与维数。《潜在分析》42(2)，第549–572页。[30] Saburo Muroga（1953）：关于离散信道的容量。I。《日本物理学会刊》8(4)，第484–494页。[31] Michael A Nielsen 和 Isaac L Chuang（2010）：《量子计算与量子信息》。

[32] *William Parry & Robert F Williams*（*1977*）：有限马尔可夫链的分组编码与 ζ 函数。《伦敦数学会会刊》*3(3)*，第*483–495*页。[33] *Gilles Pisier*（*2012*）：格罗滕迪克定理：过去与现在。《美国数学会通报》*49(2)*，第*237–323*页。[34] *Nornadiah Mohd Razali*、*Yap Bee Wah* 等（*2011*）：*Shapiro-Wilk*、*Kolmogorov-Smirnov*、*Lilliefors* 和安德森-达林检验的功效比较。《统计建模与分析杂志》*2(1)*，第*21–33*页。[35] *Robert Schatten &* 约翰·冯·诺伊曼（*1948*）：线性变换的交叉空间。*III*。《数学年刊》，第*557–582*页。[36] *Ralf Schiffler*（*2014*）：《箭图表示》。施普林格。[37] 克劳德·*E*· 香农（*1957*）：噪声信道编码理论的一些结果。《信息与控制》*1(1)*，第*6–25*页。[38] *David I Spivak*（*2015*）：耦合动力系统的稳态遵循矩阵算术进行组合。*arXiv* 预印本 *arXiv:1512.00802*。[39] *Jussi Vanhatalo*、*Hagen V¨olzer & Jana Koehler*（*2008*）：改进的流程结构树。载于：《业务流程管理：第六届国际会议，*BPM 2008*，意大利米兰，*2008*年*9*月*2*日至*4*日。论文集》第*6*卷，施普林格，第*100–115*页。[40] *Nikita Visnevski*、*Vikram Krishnamurthy*、*Alex Wang & Simon Haykin*（*2007*）：多功能雷达的句法建模与信号处理：一种随机上下文无关文法方法。《*IEEE*会刊》*95(5)*，第*1000–1025*页。[41] *Chun-Chao Wang*（*2015*）：用于多元正态性检验的 *MATLAB* 软件包。《统计计算与模拟杂志》*85(1)*，第*166–188*页。[42] *Simon Willerton*（*2009*）：度量空间大小的启发式与计算机计算。*arXiv* 预印本 *arXiv:0910.5500*。