

多个体系统理论课程复习提纲

刘志新

1 引言

多个体系统理论不是一个成熟的学科，目前没有统一的研究框架，也没有一般的分析方法和工具。它涉及到多个学科，包括计算机，物理，化学，进化论，生物，经济学，社会学等等。只能根据我的了解介绍多个体系统研究中的一些典型的例子。

课程主线见图 1：

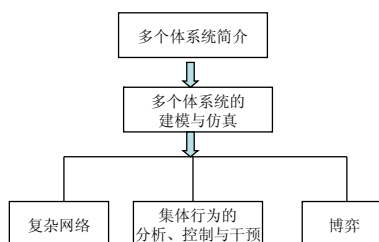


图 1

我希望通过本课程学习掌握：

1. 多个体系统的一些基本概念，包括 agent 的定义、性质，多 agent 系统的定义。
2. 多个体系统建模常用的两种方法：平均场方法和基于 agent 的建模。掌握这两种建模的特点，平均场建模方法会用相平面图法分析系统的稳态；在基于 agent 的建模中，会用元胞自动机建模。
3. 掌握度量复杂网络拓扑结构的一些概念，包括平均路径长度，聚类系数，度分布，介数等。对于简单的网络，能够计算出平均路径长度，聚类系数，度分布，介数等。对于几种基本网络，包括随机图、小世界网络、无标度网络，掌握它们的生成过程以及相应的网络拓扑结构性质。对 BA 模型，掌握平均场或者主方程的证明方法。
4. 在集体行为研究中，重点掌握同步的分析与控制。在静态网络情形，对连续系统和离散系统，系统同步需要什么条件，同步的收敛速度如何刻画。对动态网络，掌握离散系统与混杂系统（系统连续，控制律离散）同步的条件，证明方法以及同步速率的估计。对于一些简单的多个体系统，能够设计牵制控制策略使得系统同步到期望的状态。掌握复杂网络中结构能控的定义，对于给定的网络控制系统（A，B 结构给定），能够判断系统是否结构能控；给定一网络结构（A 结构给定），能找出使系统结构能控所需的最少驱动节点数目。
5. 对于博弈部分，掌握策略式博弈的表述方法、纳什平衡与进化稳定策略的定义，对于简单的博弈，能找到纳什平衡解和进化稳定策略（如果存在的话）。

2 多个体系统基本概念

要定义多个体系统，我们先定义个体 (agent)，这里我们引用 Wooldridge 与 Jennings (1995) 的定义：Agent 是处在某个环境中的计算机系统，该系统有能力在这个环境中自主行动以实现某设计目标。

一般来说，每个 agent 都有自己的目标，Agent 在追求目标时产生动作输出，该输出往往会环境影响，Agent 不能完全控制环境，只能影响环境。

自主体 / 个体 (agent) 具有两个重要特性：

- 自主性: capable of autonomous action – of deciding for themselves what they need to do in order to satisfy their design objectives ; 决策能力、自利。
- 交互性: capable of interacting with other agents - the kind of social activity that we all engage in every day of our lives: cooperation, competition, negotiation, and the like.

什么是多个体系统？简言之，多个体系统是由大量 agent 组成，agent 之间存在相互作用，还包含外部环境。多 agent 系统可以通过图 2 来描述：

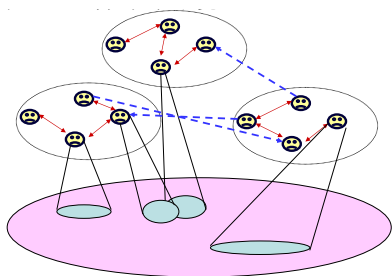


图 2

3 多自主体系统的建模与仿真

要研究多个体系统，首先要建模。那么什么是模型？

- 模型是对一个系统的简化的数学描述
- 在真实系统中，很多因素可能都是重要的但不是所有的因素都要包含在模型里面
- 根据我们研究的目的，只把相关的重要因素包含进来

多自主体系统建模有两种基本的方式：平均场建模 (mean field modelling) 和基于 agent 的建模方式 (agent-based modelling) ，这两种方法的特点是：

- 平均场方法（粗）：把环境对物体的作用进行集体处理，以平均作用效果替代单个作用效果加和的方法。不考虑空间结构对群体行为的影响，一般采用微分方程或差分方程对整个系统中个体的数量或者比例进行建模，所得到的解是随时间演化的空间上的平均量
- 基于 Agent 的建模（细）：考虑空间结构对 agent 行为的影响，agent 之间的相互作用在不同空间位置可能是不同的

对于有的现象，两种方式都有人用，比如捕食 - 被捕食系统；而对有的现象，则用 MAS 建模更合适，比如鸟群中的同步、合作起源、语言起源、免疫系统等等。

平均场建模方法举例：种群模型。

在很多系统中，如人口、鱼群等，关注的是种群的密度（或者数量）的变化规律。事实上，在生态系统中，影响种群密度的因素有很多，比如相互作用的模式，有竞争（利用共同资源）、互惠、宿主 - 寄生，并随着季节变化、年龄结构、空间分布等因素的影响。

先考虑单一种群模型。设 R 是一个离散的种群的增长率，即

$$x(t+1) = Rx(t),$$

其中 $x(t)$ 为第 t 代的密度。显然

- 当 $0 < R < 1$ 时, $x(t) \rightarrow 0$;
- 当 $R = 1$ 时, $x(t) = x(0)$;
- 当 $R \geq 1$ 时, $x(t) \rightarrow \infty$;

对于连续种群, 设 $x(t)$ 为 t 时刻种群的数量, 则

$$\dot{x}(t) = \lim_{\delta t \rightarrow 0} \frac{x(t + \delta t) - x(t)}{\delta t}$$

表示种群的变化律, $\frac{\dot{x}(t)}{x(t)}$ 表示种群的相对增长率, 假设其为 r , 则

$$\dot{x}(t) = rx(t),$$

当 $r > 0$ 时, $x(t) \rightarrow \infty$ 。显然这是不合理的, 因为资源是有限的, 同一种群内部存在着竞争, 较大的种群隐含着较小的增长率。当种群内部有竞争关系时, 假设增长率为 x 的单调递减函数, 这样就给出了 logistic 增长模型:

$$\dot{x}(t) = rx(t)(1 - \frac{x(t)}{K}).$$

利用变量分离法, 我们可以得到

$$x(t) = \frac{Kx(0)e^{rt}}{K + x(0)(e^{rt} - 1)},$$

可以看出

- 当 $x(0)$ 非常小时, 几乎指数增长;
- 当 $x(0)$ 逐渐增大, 增长率慢慢变小, $x(t)$ 最终趋于常数 K (种群的饱和密度)。

对于离散的竞争模型, 会出现混沌与分岔 (见 MAS-2.ppt)。

下面介绍捕食 - 被捕食模型, 该模型由两位数学生态学的先驱者 A.J. Lotka 与意大利数学家 V. Volterra 各自建立, 因而称为 Lotka-Volterra 方程, 主要研究研究捕食者 - 被捕食者之间的数量关系, 如野兔与猓狗的种群震荡等。

Volterra 假设, 在没有捕食者时, 被捕食者的增长为常数 a , 但为捕食者密度的线性减函数。没有被捕食者时, 捕食者将走向死亡, 即有负增长率 c ; 而增长率又有捕食者密度的贡献。据此可得到 Lotka-Volterra 方程如下:

$$\begin{cases} \dot{x}(t) = ax(t) - bx(t)y(t) \\ \dot{y}(t) = -cy(t) + dx(t)y(t) \end{cases},$$

其中 $x(t), y(t)$ 分别代表被捕食者和捕食者的数量, a 代表被捕食者的出生率, c 代表捕食者的死亡率, b, d 代表两个物种的相互作用。该方程的解为:

$$(1) x(t) = y(t) = 0;$$

$$(2) \text{ 若 } y(t) = 0, \text{ 则 } x(t) = x(0)e^{at}. \text{ 捕食者不存在, 被捕食者因无天敌, 呈指数增长;}$$

$$(3) \text{ 若 } x(t) = 0, \text{ 则 } y(t) = y(0)e^{-ct}. \text{ 因捕食者 } y \text{ 仅以 } x \text{ 为食, 则 } y(t) \text{ 呈指数下降;}$$

(4) 若 $x(t) \neq y(t) \neq 0$; xy 项表示 x 与 y 的相互作用。它表示物种 x 与 y 相遇的几率, 而系数的正负反映 y 捕食 x 的后果。我们用微分方程的定性理论进行分析。方程的平衡点可以通过求解下面的方程得到:

$$\begin{cases} 0 = ax(t) - bx(t)y(t) \\ 0 = -cy(t) + dx(t)y(t) \end{cases},$$

除去 $x(t) = 0, y(t) = 0$ 的情形, 我们有 $x(t) = \frac{c}{d}, y(t) = \frac{a}{b}$ 。利用相平面图 (图 3) 分析可得:

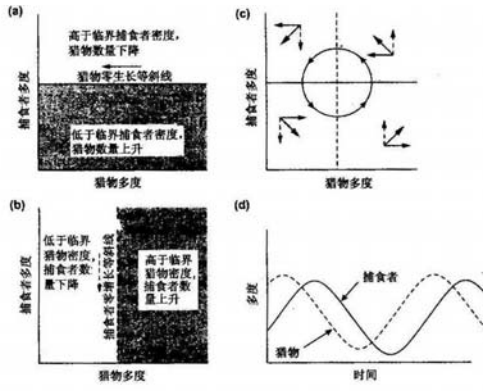


图 3

- 当捕食者少于其平衡点 a/b ，被捕食者会增加；反之当捕食者多于其平衡点 a/b ，被捕食者会减少；
- 当捕食者由少而多的通过平衡点 a/b ，被捕食者会达到其极大值 (此值 $>$ 被捕食者之平衡点 c/d)；
- 当捕食者由多而少地通过平衡点 a/b ，被捕食者数量达到极小值 ($< c/d$)。

事实上，当 $x(t) \neq y(t) \neq 0$ 时，系统存在周期解。根据 Lotka-Volterra 方程，我们有

$$\frac{\dot{x}(t)}{\dot{y}(t)} = \frac{ax(t) - bx(t)y(t)}{-cy(t) + dx(t)y(t)},$$

利用变量分离法，我们可以得到

$$\dot{x}(t) \frac{-c + dx(t)}{x(t)} = \dot{y}(t) \frac{a - by(t)}{y(t)},$$

两边同时积分，

$$-c \log x(t) + dx(t) + C = a \log y(t) - by(t),$$

其中 C 为由初始条件决定的参数。由上面的相平面图我们可以知道系统做周期运动。

在捕食 - 被捕食者模型中，还可以考虑内部竞争的影响，详见 MAS-2.ppt。

通过种群模型的研究，我们看到平均场的建模方法比较简洁，并且现有的一些理论，比如微分方程等可以拿来应用做定性（量）的理论分析。但是，平均场方法只能得到在空间平均意义下一些量的变化规律，无法也不可能得到系统的结构组织变化行为。

我们求助于基于 agent 的建模方法。仍以捕食者 - 被捕食者为例。

初始时刻：Begins with a randomly distributed population of fish, sharks, and empty cells in a 1000x2000 cell grid (2 million cells) Initially, 50% of the cells are occupied by fish (Fish = 1 (red pixel)), 25% are occupied by sharks (Sharks = -1 (yellow pixel)), 25% are empty (Empty cells = 0 (black pixel)).

演化规则：

- if the current cell is empty, If there are ≥ 4 neighbors of one species, and ≥ 3 of them are of breeding age (Fish breeding age ≥ 2 , Shark breeding age ≥ 3), and there are < 4 of the other species: then create a species of that type +1 = baby fish (age = 1 at birth) -1 = baby shark (age = $|-1|$ at birth);
- If the current cell contains a fish, Fish live for 10 generations. If ≥ 5 neighbors are sharks, fish dies (shark food); If all 8 neighbors are fish, fish dies (overpopulation); If a fish does not die, then increase its age.

- If the current cell contains a shark, Sharks live for 20 generations. If ≥ 6 neighbors are sharks and fish neighbors = 0, the shark dies (starvation); A shark has a $1/32$ (.031) chance of dying due to random causes; If a shark does not die, then increase its age.

注意, what happens in each cell is determined only by rules that apply locally, yet which often yield long-term large-scale patterns.

演化结果:

- Borders tended to ‘harden’ along vertical, horizontal and diagonal lines;
- Borders of empty cells form between like species;
- Clumps of fish tend to coalesce and form convex shapes or ‘communities’.

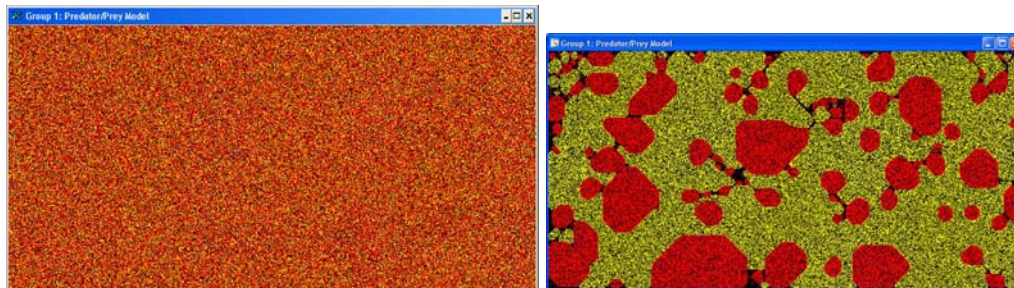


图 4 左图为初始时刻鱼与鲨鱼的分布, 右图为 10500 代的结果。

上述模型实际上是一元胞自动机模型。元胞自动机实质上是定义在由离散的有限状态的元胞组成的元胞空间上, 按照一定的局部规则, 在离散的时间维度上演化的动力学系统。它的特点是:

- 空间被一定形式的网格分割成很多单元, 每一个单元称为元胞, 而状态只能在有限的、离散状态集中选取;
- 所有元胞遵守相同的作用规则, 依照确定的局部规则进行更新;
- 通过局部相互作用进行演化。

元胞自动机不是由数学函数或者物理方程确定, 而是由一系列的演化规则构成。元胞 i 在 t 时刻的演化规则可以写为

$$F(S_i^{t+1}) = f(S_j^{t+1}, j \in N_i^t),$$

其中 N_i^t 为个体 i 在 t 时刻的邻居。

元胞自动机由以下四个要素组成:

- 元胞: An automaton consists of a grid/lattice of cells, each can be in a number of states
- 元胞空间: 元胞所分布的空间上的网格点的集合就是元胞空间。理论上可以是任意欧几里德空间的规则划分。对于二维元胞自动机, 常见的有下面三种网格排列: 三角网格, 四方网格, 六方形网格。
- 邻居: Neighbourhoods determine the extent of the interaction between cells in the grid. 对于二维的元胞自动机, 常用的两种邻居关系是: von Neumann 邻域和 Moore 邻域。
- 演化规则: 刻画每个元胞如何随着邻居状态的变化而演化的

当然，边界也很重要，常用的边界条件有：周期型边界、固定边界、绝热边界、映射边界。

例：Conway 的生命游戏。

- 每个元胞有两个状态：0 代表死，1 代表生
- 元胞分布在 2 维空间格子上
- 邻域采用 Moore neighborhood
- 演化规则：
 1. 生存 (survival)：对于一个活的元胞，如果它的邻居中有两个或者三个元胞是活的，那么该元胞继续生存下去
 2. 死亡 (die)：对于一个活元胞，如果它的邻居中有四个或者四个以上的元胞是活的，或它的邻居中只有一个或没有活的元胞，那么该元胞也将死去。
 3. 出生 (born)：对一个空 (死) 的元胞，如果它的邻居中有 3 个 (不能多也不能少)，那么该元胞将成为一个活的元胞。

系统经过演化会出现一些有趣的 pattern，比如 Blinker, Glider, Lightweight spaceship 等。

在 Lotka-Volterra 方程描述的模型里面，只有关于种群数量的变化规律（这是因为平均场方法只能得到关于空间的平均解）；而在元胞自动机模型里面，我们既可以看到种群数量的变化规律，还能看到空间的 pattern 变化规律。但 Lotka-Volterra 方程描述的模型更为简洁，所以更容易得到数学上的分析结论。

仿真语言：netlogo。

4 复杂网络

网络和图的定义没有很严格的不同，都是由一组节点以及定义在节点间的连边组成。节点可以是个体，飞机，鸟，人等，边反映了节点间的作用。复杂网络是对多个个体系统的一种高度的抽象和简化。在复杂网络中，我们关心整个网络上拓扑结构的涌现现象。

传统的图论中节点的个数相对较少，并且以往的研究中大多是通过组合和代数的工具去研究图上的性质。但是随着技术的发展，尤其是超级计算机运算能力的提高，我们能收集到越来越多的复杂数据，因而网络的规模越来越大。另一方面，不同学科间的不同的系统有可能用相同的网络进行建模，这促使我们去研究大规模的复杂网络。

我们先定义衡量网络拓扑结构的几个概念，这些概念反映了整个网络统计意义下的拓扑性质，针对单个节点意义不大。

1. 平均路径长度 (Average path length)。

对于网络 $G = (V, E)$ ，其中 V, E 分别为节点和边的集合，连接节点 i 与 j 的路径是指一个节点集合 $i = i_0, i_1, \dots, i_l = j$ ，使得任意相邻的两个节点之间有边相连， l 称为该路径的长度。两个节点之间的最短路径 (也称为距离)，是指连接两节点之间连边数最少的路径。平均路径长度 $\langle l \rangle$ 定义为任意两个节点之间的距离的平均值，即

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{j \neq i} d_{ij}.$$

对于无向网络，平均路径长度可以通过下式计算：

$$\langle l \rangle = \frac{2}{N(N-1)} \sum_{j > i} d_{ij}.$$

平均路径长度反映了网络中任何两者之间平均通过多少人连接起来。严格来讲，网络平均路径长度只对连通图才是有限的。这里的连通图是指网络中任意两个节点之间都存在着路径。但大型网络往往是不连通的，这就导致了整个网络的平均路径长度为无穷大。为了避免这种发散的情况，可以把平均路径长度定义为存在有连通路径的节点对之间的距离的平均值。这对于网络中相当部分的节点都属于连通巨片的网络比较合适。我们说网络具有小世界性质是指平均路径长度满足 $l \sim \log N$ 。

2. 聚类系数 (Clustering coefficient)

聚类系数反映了你的朋友也是朋友的可能性。对于任意的节点 i ，其聚类系数定义为：

$$C_i = \frac{E_i}{T_i} = \frac{2E_i}{k_i(k_i - 1)},$$

其中 k_i 为节点 i 的度，也就是点 i 的邻居数， E_i 为节点 i 及其邻居形成的实际的三角形的个数，也就是以节点 i 为中心顶点的三角形的个数，而 T_i 是这些邻居可能形成的三角形的个数，即 $T_i = \frac{k_i(k_i - 1)}{2}$ 。整个网络的聚类系数定义为 $C = \langle C(i) \rangle$ 。

3. 度分布 (degree distribution)

对于任意的节点 i ，它的度定义为与之相连的节点个数，也就是点 i 的邻居数。度分布 $P(k)$ 定义为对任意的节点它的度为 k 的概率。

度分布刻画了整个网络上节点度的传播。

4. 介数 (betweenness) 中心性 (也称为介数)：利用经过该节点的最短路径的数目来刻画节点的重要性。节点 i 的介数通过下面的公式进行计算：

$$BC_i = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}},$$

其中 g_{st} 表示节点 s 到 t 的最短路径的条数， n_{st}^i 表示节点 s 到 t 的最短路径中经过节点 i 的条数。

下面介绍一些典型网络的拓扑结构性质。

1. 环形图 (Ring graph)：N 个节点均匀分布在圆环上，每个节点和最近的 K 个邻居相连，左右各为 $K/2$ 个邻居。环形图的聚类系数：

$$C = \langle C_i \rangle = \frac{2C_{K/2}^2 + [K/2 - 1 + \dots + K/2 - (K/2 - 1)]}{K(K - 1)/2} = \frac{3(K - 2)}{4(K - 1)}.$$

平均路径长度：

$$\langle l \rangle \approx \frac{N}{2K}.$$

度分布：

$$P(k) = \begin{cases} 1, & \text{if } k=K; \\ 0, & \text{otherwiae.} \end{cases}$$

2. 随机图 (random graphs).

随机图模型是由 Paul Erdos 与 Alfred Renyi 在上世纪 60 年代提出的，该模型简洁直观，我们从中可以得到一些直观的感觉。另外，该模型还展现出相变现象。

随机图 $G(n, p)$ 是一个由 n 个节点组成的图的集合，这些图中任意节点对之间以相同的概率 p 连接，不同节点对之间的连接概率是独立的。（节点对之间有边的概率是 p ，无边的概率是 $(1 - p)$ ）。

ER 随机图 $G(n, p)$ 的构造算法：

(1) 初始化：给定 n 个节点以及连边概率 $p \in (0, 1]$ 。

(2) 随机连边：选择一对没有边相连的不同的节点，生成一个随机数 $r \in (0, 1)$ ，如果 $r < p$ ，那么在这对节点之间添加一条边，否则不添加边。重复以上过程直至所有的节点对都被选择过一次。

给定网络节点数 n 和连边生成概率 p ，网络中任一给定的节点恰好与其他 k 个节点有边相连的概率为 $p^k(1-p)^{n-1-k}$ 。共有 $\binom{n-1}{k}$ 种选取 k 个其他节点的方式，因此网络中任一给定节点度为 k 的概率为：

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

平均度为 $\langle k \rangle = p(n-1)$ 。当 n 很大 p 很小时，度分布近似为 Poisson 分布（参照 MAS-3.ppt）。

聚类系数：

$$C = p = \frac{\langle k \rangle}{n-1}.$$

平均路径长度：

$$\langle l \rangle = \frac{\ln n}{\ln \langle k \rangle} (1 + o(1)).$$

临界现象：

- 当连边概率 p 从 0 开始增大时，网络中初始阶段的 N 个孤立节点也开始形成一些小的连通片。也就是说，当 p 很小时，网络是由大量的碎片组成的。此时处于亚临界状态。
- 当 p 超过某个临界值 $p_c \sim 1/n$ 时，网络中会突然涌现出一个包含相当部分节点的连通的巨片，此时处于超临界状态。
- 当 $p > \ln(n)/n$ 时，随机图几乎总是连通的。

3. 小世界模型 (small world networks)

WS 模型是由 D.J. Watts 与 S.H. Strogatz 于 1998 年提出的，它是环形图与随机图的结合。环形图具有较高的聚类系数，而随机图具有较小的平均路径长度，而 WS 的小世界模型同时具备了这些好的性质。

WS 模型的生成过程：

(1) 从规则图开始：给定一个含有 N 个点的环状最近邻耦合网络，每个节点都与它左右相邻的各 $K/2$ 个节点相连， K 为偶数

(2) 随机化重连：以概率 p 随机地重新连接网络中原有的每条边，即把每条边的一个端点保持不变，另一个端点改取为网络中随机选择的一个节点，其中规定不得有重边和自环。

从生成过程来看，WS 模型在 p 很小的时候，具有规则图的特性；当 p 慢慢增大时，逐步向随机图的性质转变。

WS 模型的平均路径长度为：

$$l(N, p) \sim \frac{K}{N} f(pKN),$$

其中 $f(u)$ 为一标度函数满足

$$f(u) \begin{cases} \text{constant}, & \text{if } u \ll 1; \\ \frac{\ln u}{u}, & \text{if } u \gg 1. \end{cases},$$

Newman 与合作者给出了如下的近似形式:

$$f(u) = \frac{4}{\sqrt{u^2 + 4u}} \tanh^{-1} \frac{u}{\sqrt{u^2 + 4u}},$$

其中 $\tanh^{-1} = \frac{1}{2} \ln \frac{1+u}{1-u}$. 我们可以看出, 当 $NKp \ll 1$ 时, $l(N, p) \sim \frac{N}{K}$; 当 $NKp \gg 1$ 时, $l(N, p) \sim \frac{\ln pKN}{K^2 p}$.

WS 模型的聚类系数为

$$C = \frac{3(K-2)}{4(K-1)}(1-p)^3 + O(1/N).$$

当 $k \geq K/2$ 时, 度分布为:

$$P(k) = \sum_{n=0}^{\min(k-K/2, K/2)} \binom{K/2}{n} (1-p)^n p^{K/2-n} \frac{(pK/2)^{k-K/2-n}}{(k-K/2-n)!} e^{-pK/2},$$

当 $k < K/2$ 时, $P(k) = 0$.

1999 年, M.E.J. Newman 与 D.J. Watts 提出了 NW 模型, 该模型的生成过程为:

- 从规则图开始: 给定一个含有 N 个点的环状最近邻耦合网络, 每个节点都与它左右相邻的各 $K/2$ 个节点相连, K 为偶数满足 $1 \ll \ln(N) \ll K \ll N$.
- 随机加边: 在以概率 p 选择出的节点对之间加边, 其中规定不得有重边和自环.

NW 模型的平均路径长度、聚类系数及度分布见 MAS-3.ppt.

4. 无标度网络 (Scale-free networks)

无标度网络是指度分布不存在单一的特征标度, 那么什么样的分布函数才具有无标度性质呢? 答案是幂律分布.

定理 4.1 考虑一个概率分布函数 $f(x)$, 假设 $f(1)f'(1) \neq 0$. 如果对任意给定常数 a , 存在常数 b 使得函数 $f(x)$ 满足如下的“无标度条件”:

$$f(ax) = bf(x), \quad (4.1)$$

那么必有

$$f(x) = f(1)x^{-\gamma}, \gamma = -\frac{f'(1)}{f(1)}.$$

证明: 在 (4.1) 中取 $x = 1$, 有 $f(a) = bf(1)$, 从而 $b = f(a)/f(1)$, 因而

$$f(ax) = \frac{f(a)f(x)}{f(1)},$$

上式对任意的 a 都成立, 两边对 a 求导可得

$$x \frac{df(ax)}{d(ax)} = \frac{f(x)}{f(1)} \frac{df(a)}{da}.$$

取 $a = 1$, 可得

$$x \frac{df(x)}{d(x)} = \frac{f'(1)}{f(1)} f(x),$$

两边积分可得:

$$\ln f(x) = \frac{f'(1)}{f(1)} \ln x + \ln f(1),$$

两边取指数即可得定理。 ■

无标度网络的性质：很少的节点具有很多的连接，而大部分节点具有很少的连接。1999 年 Barabási 与 Albert 提出了 BA 模型，可以证明 BA 模型具有无标度性质。

BA 模型生成过程：

(1) 增长：从一个具有 m_0 个节点的连通网络开始，每次引入一个新的节点并且连到 m 个已存在的节点上，这里 $m \leq m_0$ 。

(2) 优先连接：一个新节点与一个已经存在的节点 i 相连接的概率 Π_i 与节点 i 的度 k_i 之间满足如下关系：

$$\Pi_i = \frac{k_i}{\sum_j k_j}.$$

BA 模型的平均路径长度：

$$\langle l \rangle \sim \frac{\ln N}{\ln \ln N}.$$

但 BA 模型没有表现出很强的聚类性质，其聚类系数 $C \sim N^{0.75}$ 。

下面我们用平均场的方法证明 BA 模型的度分布服从幂律分布。平均场方法只是一种近似的分析方法，这里我们需要做如下近似：1. 时间 t 不再是离散的，而是连续的，并且是可微的；2. 节点的度值也不再是整数，而是可以为任意实数。

证明：记 $k_i(t)$ 为节点 i 在 t 时刻的度，那么当一个新节点加入到系统中时，节点 i 的度 $k_i(t)$ 改变的速率正比于 Π_i ，

$$\frac{\partial k_i(t)}{\partial t} = m \Pi_i = \frac{m k_i}{\sum_{j=1}^{m_0+t} k_j},$$

注意到 $\sum_{j=1}^{m_0+t} k_j \approx 2mt$ 。因此有

$$\frac{\partial k_i(t)}{\partial t} = \frac{m k_i}{2mt} = \frac{k_i}{2t}.$$

利用初始条件 $k_i(t_i) = m_0$ 可得

$$k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}.$$

下面计算概率分布函数。

$$P(k_i(t) < k) = P\left(t_i > \frac{m^2 t}{k^2}\right).$$

我们假设节点以相等的时间间隔进入网络，那么 t_i 的概率密度为：

$$P(t_i) = \frac{1}{m_0 + t}.$$

因而有

$$P\left(t_i > \frac{m^2 t}{k^2}\right) = 1 - P\left(t_i \leq \frac{m^2 t}{k^2}\right) = 1 - \frac{m^2 t}{k^2(m_0 + t)}.$$

进一步，我们求得概率密度函数为

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = 2m^2 \frac{t}{m_0 + t} \frac{1}{k^3} = 2m^2 k^{-3}.$$

证毕。 ■

除了平均场方法，也可以用主方程的方法进行证明，参照 MAS4.ppt。

对于节点或者边失效，BA 模型具有一定的拓扑鲁棒性。Barabasi 小组比较了随机图和 BA 无标度网络的连通性对节点去除的鲁棒性。考虑两种节点去除策略：一是随机故障策略，即完全随机地去除网络中的一部分节点；二是蓄意攻击策略，即从去除网络中度最高的节点开始，有意识地去掉网络中一部分度高的节点。通过仿真发现：无标度网络对随机节点故障具有极高的鲁棒性；而对蓄意攻击具有高度的脆弱性。这些都归因为节点度分布的极度不均匀性。

Albert and Barabási 于 2000 年提出了扩展的 BA 模型。模型生成过程如下：

1) Start with m_0 isolated nodes

2) At each time step, we perform the following operations:

- With probability p , we add $m(\leq m_0)$ new edges. One end of a new edge is selected randomly, the other with probability

$$\Pi(k_i) = \frac{k_i + 1}{\sum_j (k_j + 1)}.$$

- With probability q , we rewire m edges. For this we randomly select a node and remove an edge l_{ij} connected to it, replacing it with a new edge l'_{ij} that connects i with node j' chosen with probability $\Pi(k'_j)$.
- With probability $1 - p - q$, we add a new node. The new node has m new edges that with probability $\Pi(k_i)$ are connected to nodes i already present in the network.

除了 BA 模型与扩展的 BA 模型，还有其它方式可以生成无标度网络，比如 Price 模型，参照 MAS4.ppt。

以上研究主要集中在网络的拓扑结构的涌现，下面介绍网络上的动力学。我们以传染病模型为例。

网络上的传播（蔓延、扩散、流行）在许多实际网络中都广泛存在，它们基于不同的网络结构，表现形式也不完全相同。比如：社会网络中的疾病传播，通信网络中的病毒传播，社会网络中的信息传播，电力网络中的相继故障，经济网络中的危机扩散等。近年来人们开始关注网络结构对传播行为的影响，我们重点介绍传染病模型。一方面它的研究较多，另一方面在一定程度上可以推广到其它网络的传播行为上。在典型的传染病模型中，种群内的 N 个个体的状态可以分为以下三类：

- 易染状态 S(Susceptible)：一个个体在感染之前处于易染状态，即该个体有可能被邻感染。
- 感染状态 I(Infected)：感染上某种病毒的个体处于感染状态，该个体会以一定的概率感染邻居。
- 移除状态 R(Recovered)：个体经历过一个完整的感染周期后，该个体就不再被感染。

经典的传染病模型假设所有个体是完全混合的，可以用平均场的方法进行建模。我们介绍三个基本的传染病模型。

1. SI 模型。SI 模型是最简单的情形：个体一旦被感染，永远处于感染状态。记 $S(t)$ 和 $I(t)$ 分别为时刻 t 的易染人数和感染人数， $S(t) + I(t) = N$ 。假设一个易染个体在单位时间里与感染个体接触并被感染的概率为 β 。由于易染个体的比例为 $S(t)/N$ ，时刻 t 网络中总共有 $I(t)$ 个感染个体，所以易染个体的数目按照如下变化率减少：

$$\frac{dS}{dt} = -\beta \frac{SI}{N},$$

相应地，感染个体的数目按照如下变化率增加：

$$\frac{dI}{dt} = \beta \frac{SI}{N}.$$

记 $s(t) = S(t)/N, i(t) = I(t)/N$, 显然 $s(t) + i(t) = 1$. 很容易有

$$\frac{di(t)}{dt} = \beta s(t)i(t) = \beta(1 - i(t))i(t),$$

利用变量分离法,

$$i(t) = \frac{i_0 \exp\{\beta t\}}{1 - i_0 + i_0 \exp\{\beta t\}}, \quad i(0) = i_0.$$

初始阶段, 绝大多数为易染个体, 任何一个感染个体很容易接触到易染个体并把病毒传染给后者, 因此感染个体数量随着时间指数增长; 但随着易染个体个体数量的减少, 感染个体数量的增加也呈现饱和状态。

2. SIR 模型。与 SI 模型的假设类似, 不同之处在于感染个体以某个定常速率 γ 被移除。记 $s(t), i(t), r(t)$ 为易染人群、感染人群和移除人群的比例, 则 SIR 模型如下:

$$\begin{cases} \frac{ds(t)}{dt} = -\beta s(t)i(t) \\ \frac{di(t)}{dt} = \beta s(t)i(t) - \gamma i(t) \\ \frac{dr(t)}{dt} = \gamma i(t) \\ s(t) + i(t) + r(t) = 1 \end{cases}.$$

根据上式中的第一式和第三式可得

$$\frac{1}{s(t)} \frac{ds(t)}{dt} = -\frac{\beta}{\gamma} \frac{dr(t)}{dt},$$

两边积分

$$s(t) = s_0 e^{-\beta r(t)/\gamma}, \quad s(0) = s_0.$$

据此及 SIR 模型中的第三式可得

$$\frac{dr(t)}{dt} = \gamma(1 - s_0 e^{-\beta r(t)/\gamma} - r(t)).$$

令上式中的右端等于 0 , 可得移除状态的稳态值如下:

$$r(t) = 1 - s_0 e^{-\beta r(t)/\gamma}.$$

对于大规模网络, 初始时刻只有少数感染个体, 没有移除个体, 从而有 $s_0 \approx 1, i_0 \approx 0, r_0 = 0$, 记 $\lambda = \frac{\beta}{\gamma}$, 于是有

$$r(t) = 1 - e^{\lambda r(t)},$$

$\lambda = 1$ 是 SIR 模型的传染临界值: 如果 $\lambda < 1$, 那么 $r = 0$, 意味着病毒无法传播; $\lambda > 1$, 那么 $r > 0$, 并且随着 λ 值的增大, r 值也增大, 意味着病毒在网络中扩散的范围也增大。

3. SIS 模型。SIR 模型与 SIS 模型的区别在于感染个体恢复之后的状态。在 SIR 模型中, 一个感染个体恢复之后处于移除状态; 而在 SIS 模型中, 每一个感染个体以定常速率 (记为 γ) 再变为易染个体。记 $s(t)$ 和 $i(t)$ 为分别为时刻 t 的易染和感染人群占的比例, 则 SIS 模型描述如下:

$$\begin{cases} \frac{ds(t)}{dt} = \gamma i(t) - \beta s(t)i(t) \\ \frac{di(t)}{dt} = \beta s(t)i(t) - \gamma i(t) \\ s(t) + i(t) = 1 \end{cases}.$$

根据上式中的第二式, 很容易得到

$$\frac{di(t)}{dt} = \beta(1 - i(t))i(t) - \gamma i(t).$$

当 $\lambda = \frac{\beta}{\gamma} > 1$ 时, 此方程的稳态值为

$$i(t) = 1 - \frac{\gamma}{\beta}.$$

SIS 模型的解为:

$$i(t) = \frac{i_0(\beta - \gamma) \exp\{(\beta - \gamma)t\}}{(\beta - \gamma - \beta i_0) + \beta i_0 \exp\{(\beta - \gamma)t\}}.$$

$\lambda = 1$ 是 SIS 模型的传染临界值: 当 $\lambda > 1$ 时, 方程收敛到稳态值当 $\lambda < 1$ 时, 方程收敛到零, 不扩散。

经典的传染病模型中假设个体间以等概率接触, 但现实世界中一个个体通常只能和网络中很少的节点直接接触, 也就只能传播给那些直接接触的节点。我们以 SIS 为例研究一下网络结构对传染病模型的影响。

均匀网络上的 SIS 模型。这里的均匀网络是指: 每个节点的度都差不多, 集中在均值 $\langle k \rangle$ 附近。如 ER 随机图和 WS 小世界网络, 它们的度分布在 $\langle k \rangle$ 附近有个尖峰, 当 $k \gg \langle k \rangle$ 或 $k \ll \langle k \rangle$ 时, 度呈指数下降。我们假设如果一个易染节点的邻居中至少有一个感染节点, 该节点被感染的概率假设为常数 β , 而一个感染节点恢复到易染节点的概率假设为常数 γ 。定义传播概率为:

$$\lambda = \frac{\beta}{\gamma}.$$

记网络在时刻 t 感染个体的密度为 $\rho(t)$ 。不考虑度之间的相互性, 并另 $\gamma = 1$, 均匀网络上的 SIS 模型为:

$$\frac{d\rho(t)}{dt} = -\rho(t) + \lambda \langle k \rangle (1 - \rho(t))\rho(t),$$

其中右边的第一部分为被感染个体以单位速率恢复为易染个体, 第二部分为单个感染个体产生的新感染个体的平均密度, 它与传播率、结点的度以及与健康易染个体相连的概率成比例。系统的感染个体的稳态密度为:

$$\rho = \begin{cases} 0, & \text{if } \lambda < \frac{1}{\langle k \rangle} \\ 1 - \frac{1}{\lambda \langle k \rangle} & \text{if } \lambda > \frac{1}{\langle k \rangle} \end{cases}$$

传播的临界值为 $\lambda_c = \frac{1}{\langle k \rangle}$: 如果传播率 $\lambda \leq \lambda_c$, 那么感染个体指数衰减至 0, 无法扩散; 如果传播率 $\lambda > \lambda_c$, 感染个体能够将病毒传播扩散并使得整个网络感染个体总数最终稳定于某一平衡态。

对于非均匀网络, 我们要考虑节点的度的不同带来的影响。将均匀网络上的 SIS 模型拓展到非均匀网络上, 记 $\rho_k(t)$ 为度为 k 的节点感染的概率:

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k(1 - \rho(t))\Theta_k,$$

右边第一项表示被感染个体以单位速率恢复为易染个体, 第二项表示一个健康节点被一个与之相连的感染节点传染的概率与传播率 λ 、结点的度 k 以及度为 k 的节点的感染邻居密度 Θ_k 成比例。对于度不相关的网络, 传播临界值为:

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}.$$

那么我们怎样对节点采取免疫措施使得这些节点不再具有传播能力, 最终导致病毒不能够在整个网络上传播。免疫节点意味着这些节点以及与之相连的边可以从网络上去除。常用的免疫策略有三种:

- 随机免疫: 随机免疫一部分个体。假设免疫节点密度为 g , 从平均场的观点来看, 随机免疫意味着传播率由 λ 变为 $(1 - g)\lambda$ 。对于均匀网络, 随机免疫对应的免疫临界密度 g_c 为 $g_c = 1 - \frac{1}{\lambda \langle k \rangle}$ 。对于无标度网络, 随机免疫对应的免疫临界密度 g_c 为

$$g_c = 1 - \frac{\langle k \rangle}{\lambda \langle k^2 \rangle} \rightarrow 1.$$

对于大规模无标度网络，如果采取随机免疫策略，需要对网络中几乎所有的点都要进行免疫才能保证消灭病毒传染。

- 目标免疫是希望通过有选择地对少量节点进行免疫以获得尽可能好的免疫效果。对 BA 模型，可以选择节点度大的节点进行免疫。目标免疫对应的免疫密度临界值为 $g_c \sim e^{-2/m\lambda}$ 。即使传播率 λ 在很大范围内取不同值，都可以得到很小的免疫密度临界值。目标免疫的主要问题是需要了解网络的全局信息以找到控制病毒传播的 hub 节点。
- 熟人免疫：从 N 个节点中随机选出一定比例的节点，再从每一个被选出的节点中随机选择一个邻居节点进行免疫。这意味着象 BA 模型这样的网络，度大的节点是很容易被选做免疫节点，这样很容易达到整个网络的免疫。

5 集体行为行为的分析。

多个体系统系统在微观层面上由多个相互作用的个体组成，系统在整体层面上展现出个体单独存在时所不具备的行为特征，称为集体行为。比如，蚁群的群体智能，鱼群鸟群的同步，Internet 网等展现出的小世界、无标度性质，人群恐慌，语言进化等。

目前，集体行为的研究主要有三类：

- 分析 (Analysis)：给定个体的规则，系统在整体上会展现出什么样的行为？怎样分析？
- 分布式控制 (Distributed control)：给定期望的行为（比如队形、覆盖等），如何利用每个个体能接收到局部信息设计个体的控制律以使系统整体上展现出期望的行为？
- 干预 / 软控制 (Intervention/Soft Control)：给定期望的行为，如何在不破坏已有个体规则的基础上进行干预以使系统达到期望的行为？

以上三类研究不是彼此孤立的，而是相互关联的。

我们以同步为例进行研究。我们说多个体系统达到同步，如果系统中所有个体的状态达到一致，即存在共同的 \bar{x} 使得

$$\lim_{t \rightarrow \infty} x_i(t) = \bar{x}, \quad \forall i,$$

其中 $x_i(t)$ 表示个体 i 在时刻 t 的状态。

同步存在于很多系统中，如物理多体系统、分布式计算、神经科学、通信系统、生物群体、群体决策、社会经济系统等。

一、我们首先考虑不带噪声的一维的一阶连续时间：

$$\dot{x}_i(t) = u_i(t), \quad i = 1, 2, \dots, n,$$

其中 $x_i(t)$ 表示个体 i 在时刻 t 的状态， $u_i(t)$ 表示个体 i 的控制律。每个个体可以利用邻居的状态做分布式控制律，我们按照如下方式设计个体的控制律：

$$u_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)),$$

其中 $N_i(t)$ 表示个体 i 在时刻 t 的邻居。将上述控制律带入系统，可以得到如下的闭环系统：

$$\dot{x}_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)), \quad i = 1, 2, \dots, n.$$

我们用图 $G_t = \{V, E_t\}$ 表示个体之间的连接，其中 V 是由所有个体组成的集合， E_t 是边的集合：

$$E_t = \{(i, j) : j \in N_i(t)\}.$$

当控制律设计好后，剩下的问题就是分析。我们考虑网络的拓扑结构如何影响系统的同步行为。

1. 当图 $G_t = \{V, E_t\}$ 为无向静态图时，即图 G_t 不随时间而变化，记做 $G = \{V, E\}$ 。 G 是无向图，意味着 $(i, j) \in E$ 等价于 $(j, i) \in E$ 。

记 $x(t) = [x_1(t), \dots, x_n(t)]'$ 。则系统可以写成如下的矩阵形式：

$$\dot{x}(t) = -L(G)x(t),$$

其中 $L(G)$ 是图 G 的 Laplacian 矩阵。系统的解为：

$$x(t) = e^{-L(G)t}x(0).$$

图 G 的邻接矩阵 $A = [a_{ij}]$ 的元素 a_{ij} 定义为：如果 $(i, j) \in E$ ，则 $a_{ij} = 1$ ，否则 $a_{ij} = 0$ 。节点 i 的度定义为与节点 i 相连的邻居的个数， $d_i = \sum_{j=1}^n a_{ij}$ 。图 G 的最大度和最小度定义为 $d_{\max} = \max_i d_i$ ， $d_{\min} = \min_i d_i$ 。度矩阵定义为 $T = \text{diag}(d_1, d_2, \dots, d_n)$ 。矩阵的 Laplacian 矩阵定义为 $L = T - A$ 。矩阵 L 的性质如下：

引理 5.1 无向图 G 的 Laplacian 矩阵 L 具有如下性质：

1. 矩阵 L 为非负定矩阵。
2. 若图 G 为连通图，则零特征值只有一个，其特征值按照如下顺序排列

$$0 = \lambda_1(L(G)) < \lambda_2(L(G)) \leq \dots \leq \lambda_n(L(G)),$$

与零特征值对应的特征向量为 $\mathbf{1} = [1, 1, \dots, 1]^\tau$ 。

证明：当 G 为无向图时，矩阵 L 是对称的，所有特征值必为实数，且 0 为 L 的特征值，与之对应的特征向量为 $\mathbf{1}$ 。根据圆盘定理，我们知道 L 的所有特征值非负。从而 L 为非负定矩阵。设 ξ 为矩阵 L 的 0 特征根对应的特征向量，则有

$$\xi' L \xi = \sum_{(i,j) \in E} (\xi_i - \xi_j)^2 = 0,$$

若图 G 是连通的，则由上式可以推出 $\xi_1 = \xi_2 = \dots = \xi_n$ ，这意味着矩阵 L 的零特征根对应的特征向量空间是一维的，所以 0 根是一维的，从而

$$0 = \lambda_1(L(G)) < \lambda_2(L(G)) \leq \dots \leq \lambda_n(L(G)),$$

且与零特征值对应的特征向量为 $\mathbf{1}$ 。 ■

记与特征值 $\lambda_i (i = 1, 2, \dots, n)$ 对应的单位正交特征向量为 u_1, \dots, u_n ，且记

$$U = [u_1, u_2, \dots, u_n],$$

则

$$U^\tau L U = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \triangleq \Lambda.$$

将上式带入系统的解可得

$$\begin{aligned} x(t) &= e^{-Lt}x(0) = e^{-U\Lambda U^T t}x(0) = Ue^{-\Lambda t}U^T x(0) \\ &= U \begin{pmatrix} e^{-\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{-\lambda_2 t} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & e^{-\lambda_n t} \end{pmatrix} U^T x(0), \end{aligned}$$

从而

$$x(t) = e^{-\lambda_1 t}(u_1^T x(0))u_1 + e^{-\lambda_2 t}(u_2^T x(0))u_2 + \cdots + e^{-\lambda_n t}(u_n^T x(0))u_n.$$

当图 G 连通时, 我们有

$$\begin{aligned} \|x(t) - (u_1^T x(0))u_1\| &= \left\| x(t) - \left(\frac{\sum_{j=1}^n x_j(0)}{n} \right) \mathbf{1} \right\| \\ &\leq ce^{-\lambda_2 t}. \end{aligned}$$

据此, 我们得到下面的定理:

定理 5.1 如果图 G 是连通的, 则对任意的初始条件系统都能达到同步, 代数连通度 λ_2 决定了系统的收敛速率。

注 5.1 事实上, 图 G 的连通性也是系统同步必要条件。即系统同步的充必条件是邻居关系图 G 是连通的。

2. 当图 G 为有向静态图时, 我们借鉴无向图的情形来设计个体的规则:

$$\dot{x}_i(t) = \sum_{j \in N_i(t)} w_{ij}(x_j(t) - x_i(t)), \quad i = 1, 2, \dots, n,$$

其中 w_{ij} 表示节点 j 对节点 i 影响的权值。 $w_{ij} > 0$ 意味着图 G 中存在节点 j 到节点 i 的边, 记做 $(i, j) \in E$ 。我们可以定义邻接矩阵 $A = [a_{ij}]$ 为: 如果 $(j, i) \in E$, 则 $a_{ij} = w_{ij}$; 否则 $a_{ij} = 0$ 。节点 i 的入度 $d_i^{in} = \sum_{j=1}^n a_{ij}$ 。度矩阵定义为 $T = \text{diag}(d_1^{in}, d_2^{in}, \dots, d_n^{in})$ 。图 G 的 Laplacian 矩阵定义为 $L = T - A$ 。

在有向图中, 我们引入生成树的概念。

定义 5.1 一个有向图 G 是一个有向树 (*directed tree* 或 *rooted out-branching*), 如果

- (1) 图 G 中不包含有向的圈;
- (2) 存在一个节点 v_r (根), 使得对任意的其它节点 v , 都存在由节点 v_r 到 v 的有向路径。

定义 5.2 有向图的一个生成树是一个连接所有节点的有向树。一个有向图包含一个生成树如果图中边的子集形成一个生成树。

引理 5.2 (矩阵 - 树定理) 设 v 为一个加权有向图 \mathcal{D} 的节点, 则

$$\det L_v(\mathcal{D}) = \sum_{T \in \mathcal{T}_v} \prod_{e \in T} w(e),$$

其中 \mathcal{T}_v 是图 \mathcal{D} 中以节点 v 为根的生成树的集合, $\prod_{e \in T} w(e)$ 是生成树 T 中的边对应的权重的乘积, $L_v(\mathcal{D})$ 是矩阵 $L(\mathcal{D})$ 删掉第 v 行, 第 v 列后剩下的矩阵。

引理 5.3 一个有向图包含有向生成树的充要条件是

$$\text{rank}(L) = n - 1.$$

证明：矩阵 L 的特征多项式为

$$P(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1\lambda + \alpha_0.$$

根据 Laplacian 矩阵的定义，可知 0 是一特征根，所以 $\lambda_0 = 0$ 。因而

$$\text{rank}(L) = n - 1 \Leftrightarrow \alpha_1 \neq 0.$$

而 $\alpha_1 = \sum_i \det(L_i)$ ，其中 L_i 去掉的 i 行和第 i 列之后的矩阵。根据矩阵 - 树定理 (matrix-tree) 当且仅当图 G 中包含一个 rooted out-branching 子图。定理得证。 ■

注 5.2 由上面的引理我们知道，0 是矩阵 L 的一个特征值，且当图 G 包含生成树时， L 有唯一的 0 特征根。另外，根据圆盘定理可知 L 的所有特征值都具有正实部。

引理 5.4 设 $L = PJ(\Lambda)P^{-1}$ 为矩阵 L 的若当分解，当图 G 包含一个有向生成树时，我们能找到非奇异矩阵 P ，使得

$$J(\Lambda) = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & J_{n_l}(\lambda_l) \end{pmatrix}, 1 + n_2 + \cdots + n_l = n,$$

其中 $\lambda_i (i = 2, \dots, l)$ 有正实部， $J_{n_i}(\lambda_i)$ 是与特征值 λ_i 对应的若当块。并且有

$$\lim_{t \rightarrow \infty} e^{-J(\Lambda)t} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

从而 $\lim_{t \rightarrow \infty} e^{-Lt} = p_1 q_1^T$ ，其中 p_1 和 q_1^T 分别为 P 的第一列和 P^{-1} 的第一行，并满足 $p_1^T q_1 = 1$ 。

定理 5.2 对有向图 G ，如果所有个体按照前面的方式更新状态，则对于所有的初始状态，系统达到同步的充分必要条件是图 G 包含一个生成树，并且

$$\lim_{t \rightarrow \infty} x(t) = p_1 q_1^T x(0) = \frac{q_1^T x(0) \mathbf{1}}{\sqrt{n}},$$

其中 p_1 和 q_1 分别为矩阵 L 的关于零的右特征向量和左特征向量，且 $p_1^T q_1 = 1$ 。

据前面的分析可以直接得证。

二、考虑离散时间多个体系统，我们假设个体 $i (i = 1, \dots, n)$ 的状态 $x_i(t)$ 按照如下方式进行更新：

$$x_i(t+1) = \frac{\sum_{j \in N_i(t)} x_j(t)}{n_i(t)}, \quad i = 1, 2, \dots, n, \quad t = 0, 1, 2, \dots, \quad (5.1)$$

其中 $N_i(t)$ 表示个体 i 的邻居集合， $n_i(t)$ 表示集合 $N_i(t)$ 的元素个数。记 $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ ，则上系统可以写成如下的矩阵形式：

$$x(t+1) = P(t)x(t),$$

其中 $P(t) = [p_{ij}(t)]$ 定义如下:

$$p_{ij}(t) = \begin{cases} \frac{1}{n_i(t)}, & \text{if } j \in N_i(t) \\ 0, & \text{otherwise} \end{cases}.$$

同样地, 我们可以用一个图序列 $G_t = \{V, E_t\}$ 表示个体之间的相互作用, 其中 V 为所有个体组成的集合, 边 $(i, j) \in E_t$ 当且仅当 $j \in N_i(t)$ 。下面介绍非负矩阵的一些性质。

方阵 $A = (a_{ij})_{n \times n}$ 称为非负 (正) 的, 如果它的所有元素满足 $a_{ij} \geq (>) 0$, 记作 $A \geq (>) 0$; 进一步, 非负矩阵 A 称为随机矩阵如果其行和满足 $\sum_{j=1}^n a_{ij} = 1, i = 1, \dots, n$ 。对任意的随机矩阵 A , 都有 $A\mathbf{1} = \mathbf{1}$, 所以 1 一定是随机矩阵的一个特征根, 对应的一个特征向量为 $\mathbf{1} = [1, 1, \dots, 1]^T$ 。非负矩阵的乘积仍为非负矩阵。非负矩阵 $A = [a_{ij}]$ 称为本原矩阵 (primitive matrix), 如果存在正整数 k , 使得矩阵 M^k 的所有元素都是正的。

对任意的两个矩阵 $A = (a_{ij})$ 与 $B = (b_{ij})$, $A \geq (>) B$ 意味着 $a_{ij} \geq (>) b_{ij}$ 。

引理 5.5 如果 A 为 $n \times n$ 矩阵且 $A \geq 0$, 则 $\rho(A)$ 一定是 A 的特征值, 并且存在一个非负向量 $x \geq 0, x \neq 0$, 使得 $Ax = \rho(A)x$ 。

引理 5.6 本原矩阵的所有特征值中一定有且仅有一个具有最大模, 其它特征值的模一定小于该最大值。

引理 5.7 如果矩阵 $A = [a_{ij}]$ 为非负本原矩阵, 则有

$$\lim_{k \rightarrow \infty} [\rho^{-1}(A)A]^k = L > 0,$$

其中 $L = xy^T, Ax = \rho(A)x, A^T y = \rho(A)y, x > 0, y > 0$ 且 $x^T y = 1$ 。进一步, 如果 λ_{n-1} 为 A 的特征值且满足对于任意的 $\lambda \neq \rho(A)$, 都有 $|\lambda| < |\lambda_{n-1}|$ 且

$$|\lambda_{n-1}|/\rho(A) < r < 1,$$

则存在常数 $C = C(r, A)$ 使得

$$\|(\rho^{-1}(A)A)^k - L\|_\infty \leq Cr^k, k = 1, 2, \dots.$$

1. 先考虑静态无向图情形, 也就是图 G_t 不随时间变化, 记做 G 。根据矩阵 P 的定义, 我们知道 P 为非负随机矩阵。进一步当图 G 连通并且在每个节点上包含自圈 (这一点非常重要) 时, 此时矩阵 P 的对角线元素都大于 0, 很容易可以验证 P 为一随机本原矩阵。进一步根据前面的引理, 有

$$\lim_{k \rightarrow \infty} [P]^k = \mathbf{1}c,$$

其中 c 为矩阵 P 对应于特征值 1 的左特征向量。

据此我们很容易得到:

定理 5.3 如果图 G 为每个节点都包含自圈的连通图时, 多个体系统 (5.1) 可以达到同步。

下面的问题是: 当系统达到同步时, 系统同步的速率是多少?

矩阵 $P^k = [p_{ij}^{(k)}]$ 的元素为:

$$p_{ij}^{(k)} = \sum_{j_1, j_2, \dots, j_{k-1}} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{k-1} j}.$$

根据矩阵 P 的定义, 我们知道其非零元素 $p_{ij} \geq \frac{1}{n}$, 其中 n 为系统中个体的总个数。进一步, 当图 G 连通时, 任意两个节点 i 与 j 之间都存在着路径 $i = j'_0, j'_1, j'_2, \dots, j'_l = j$, 且路径长度 $l \leq n-1$, 这意味着 $P^{n-1} = [p_{ij}^{(n-1)}]$ 的元素

$$p_{ij}^{(n-1)} = \sum_{j_1, j_2, \dots, j_{n-1}} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j} \geq p_{ij'_1} p_{j'_1 j'_2} \cdots p_{j'_{l-1} j} p_{jj} \cdots p_{jj} \geq \left(\frac{1}{n}\right)^{n-1} > 0.$$

记 $x_{\max}(t) = \max_{1 \leq i \leq n} x_i(t)$, $x_{\min}(t) = \min_{1 \leq i \leq n} x_i(t)$, 显然存在两行 i 和 i' , 使得:

$$x_{\max}(n-1) = \sum_{j=1}^n p_{ij}^{(n-1)} x_j(0), \quad x_{\min}(n-1) = \sum_{j=1}^n p_{i'j}^{(n-1)} x_j(0),$$

因而

$$\begin{aligned} x_{\max}(n-1) - x_{\min}(n-1) &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right) x_j(0) \\ &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^+ x_j(0) - \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^- x_j(0), \end{aligned}$$

其中 $(a)^+ = \max(a, 0)$, $(a)^- = \max(-a, 0)$ 。因为对任意的正整数 k , P^k 仍为随机矩阵, 所以我们有

$$\begin{aligned} \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^+ &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^- \\ &\leq 1 - n \left(\frac{1}{n} \right)^{n-1} = 1 - \left(\frac{1}{n} \right)^{n-2} \triangleq \lambda. \end{aligned}$$

据此, 我们得到

$$x_{\max}(n-1) - x_{\min}(n-1) \leq \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^+ (x_{\max}(0) - x_{\min}(0)) \leq \lambda(x_{\max}(0) - x_{\min}(0)).$$

以此类推, 我们得到

$$x_{\max}(k(n-1)) - x_{\min}(k(n-1)) \leq \lambda^k (x_{\max}(0) - x_{\min}(0)).$$

注意到由状态更新 (5.1), 可得

$$x_{\max}(t+1) = \sum_{j=1}^n p_{ij} x_j(t) \leq x_{\max}(t), \quad x_{\min}(t+1) \geq x_{\min}(t),$$

上式意味着 $x_{\max}(t)(x_{\min}(t))$ 是单调递减(增)序列, 因而 $x_{\max}(t)$ 与 $x_{\min}(t)$ 必存在极限, 同时

$$\begin{aligned} x_{\max}(t) - x_{\min}(t) &\leq x_{\max} \left(\left\lfloor \frac{t}{n-1} \right\rfloor (n-1) \right) - x_{\min} \left(\left\lfloor \frac{t}{n-1} \right\rfloor (n-1) \right) \\ &\leq \lambda^{\lfloor \frac{t}{n-1} \rfloor} (x_{\max}(0) - x_{\min}(0)) \rightarrow 0, \quad t \rightarrow \infty, \end{aligned}$$

其中 $\lfloor \cdot \rfloor$ 表示从下取整。因而 $\lim_{t \rightarrow \infty} x_{\max}(t) = \lim_{t \rightarrow \infty} x_{\min}(t)$, 这意味着系统可达到同步。

2. 当图 G 为静态有向图。我们可以得到下面的结论:

定理 5.4 如果图 G 中每个节点上都有自圈, 且图 G 包含有向生成树, 则多个体系统 (5.1) 可以达到同步。

当图 G 中每个节点上都有自圈, 且图 G 包含有向生成树时, 矩阵 P 为随机矩阵, 且对角线元素大于 0。为简单起见, 我们假设节点 1 是生成树的根, 那么对于任意的节点 i , 总存在由节点 1 到节点 i 的有向路径, 记做 $1 = j_0, j_1, j_2, \dots, j_l = i$, 此时 $p_{j_k j_{k-1}} \geq \frac{1}{n} > 0$ 。矩阵 $P^{n-1} = [p_{ij}^{(n-1)}]$ 的第一列元素满足

$$p_{i1}^{(n-1)} = \sum_{j_1, j_2, \dots, j_{n-1}} p_{ij_1} \cdots p_{j_{n-2} j_{n-1}} \geq p_{ii} \cdots p_{ii} p_{ij_{l-1}} \cdots p_{j_1 1} \geq \left(\frac{1}{n} \right)^{n-1} > 0.$$

类似于无向图的证明，我们可以得到

$$\begin{aligned} x_{\max}(n-1) - x_{\min}(n-1) &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right) x_j(0) \\ &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^+ x_j(0) - \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^- x_j(0), \end{aligned}$$

其中

$$\begin{aligned} \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^+ &= \sum_{j=1}^n \left(p_{ij}^{(n-1)} - p_{i'j}^{(n-1)} \right)^- \\ &\leq 1 - \left(\frac{1}{n} \right)^{n-1} \triangleq \lambda. \end{aligned}$$

据此，我们有

$$\begin{aligned} x_{\max}(t) - x_{\min}(t) &\leq x_{\max} \left(\left\lfloor \frac{t}{n-1} \right\rfloor (n-1) \right) - x_{\min} \left(\left\lfloor \frac{t}{n-1} \right\rfloor (n-1) \right) \\ &\leq \lambda \left\lfloor \frac{t}{n-1} \right\rfloor (x_{\max}(0) - x_{\min}(0)) \rightarrow 0, \quad t \rightarrow \infty, \end{aligned}$$

注意到 $x_{\max}(t)(x_{\min}(t))$ 是单调递减（增）序列，因而有 $\lim_{t \rightarrow \infty} x_{\max}(t) = \lim_{t \rightarrow \infty} x_{\min}(t)$ ，这意味着系统达到同步。

3. 图 G_t 为时变无向图。此时矩阵 $P(t)$ 也是时变的，系统的解为：

$$x(t+1) = p(t) \cdots P(0)x(0),$$

系统的同步问题转化为随机矩阵的乘积 $\Pi_{k=0}^t P(k)$ 是否收敛于矩阵 $\mathbf{1}\mathbf{c}^\tau$ ？其中 $\mathbf{1} = [1, 1, \dots, 1]$ ， c 为依赖于矩阵 $P(t)$ 和初始条件的列向量。下面介绍矩阵乘积的一些性质。

定义 5.3 如果随机矩阵 A 满足

$$\lim_{t \rightarrow \infty} A^t = \mathbf{1}\mathbf{c}^\tau,$$

其中 c 为列向量，则矩阵 A 称为 *SIA* 阵（*Stochastic, Indecomposable, aperiodic*），也称为遍历矩阵。

定理 5.5 定理 (Wolfowitz)：假设 $M_1, M_2 \cdots M_m$ 为一有限个 *SIA* 矩阵组成的集合，且对每个长度为有限的序列 $M_{i_1} M_{i_2} \cdots M_{i_j}$ ，矩阵乘积也是 *SIA* 矩阵。则对于任意的无穷长度矩阵 $M_{i_1} M_{i_2} \cdots$ ，总存在列向量 c ，使得

$$\lim_{t \rightarrow \infty} M_{i_1} M_{i_2} \cdots = \mathbf{1}\mathbf{c}^\tau.$$

注 5.3 矩阵集合 $M_1, M_2 \cdots M_m$ 中每个元素都是遍历矩阵不能得到定理的结论（参见 *J. Wolfowitz 1963* 年的论文）。

当邻居关系图 $G_t = (V, E(t))$ 为无向连通图，且在每个时刻 t ，在每个节点处都有一个自圈时，可以验证图 G_t 对应的矩阵 $P(t)$ 一定是 *SIA* 阵，并且任意的 $P(t)$ 的乘积也一定是 *SIA* 矩阵。据此，我们得到：

定理 5.6 假设邻居关系图 G_t 在每个时刻都是连通的，且在每个节点处都存在自圈，则对于任意的初始状态，系统可以达到同步。

事实上，如果图 G_t 不满足在每个时刻都连通，但每隔一段有界的时间是连通的，系统也能达到同步。

定理 5.7 假设存在一个非空有界且相邻接的无穷时间区间序列 $[t_i, t_{i+1}]$ ，使得在每个时间区间内邻居关系图是联合连通的，其中 $t_0 = 0$ ，则系统能达到同步。

定义 5.4 图 $G_1 = \{V, E_1\}$ 与 $G_2 = \{V, E_2\}$ 的并图定义为：

$$G_1 \bigcup G_2 = \{V, E_1 \bigcup E_2\}$$

我们称图集合 $\{G_1, G_2, \dots, G_m\}$ 是联合联通的，如果这些图的并是连通的。

那么当矩阵联合连通时，我们首先化矩阵乘积为求和，见下面的引理：

引理 5.8 假设 m 为一正整数且 $m \geq 2$ ， A_1, A_2, \dots, A_m 为 $n \times n$ 阶的非负矩阵，矩阵 A_i 的主对角线元素都是正的，其中最大值和最小值分别记为 ρ 和 μ 。则

$$A_1 A_2 \cdots A_m \geq \left(\frac{\mu^2}{2\rho} \right)^{m-1} (A_1 + A_2 + \cdots + A_m).$$

证明：记 $\delta = \frac{\mu^2}{2\rho}$ 。我们要用数学归纳法证明：

$$A_1 A_2 \cdots A_i \geq \left(\frac{\mu^2}{2\rho} \right)^{m-1} (A_1 + A_2 + \cdots + A_i) \quad (5.2)$$

对所有的 $i \in \{2, \dots, m\}$ 成立。将矩阵 A_i 重新写作 $A_i = \mu I + B_i$ ，其中 B_i 为非负矩阵，则对于任意的 $j, k \in \{1, 2, \dots, m\}$ ，

$$A_j A_k = (\mu I + B_j)(\mu I + B_k) = \mu^2 I + \mu(B_j + B_k) + B_j B_k.$$

因而

$$\begin{aligned} A_j A_k &\geq \mu^2 I + \mu(B_j + B_k) \geq \mu^2 I + \frac{\mu^2}{2\rho}(B_j + B_k) \\ &\geq \delta((\rho I + B_j) + (\rho I + B_k)) \geq \delta(A_j + A_k). \end{aligned}$$

式 (5.2) 对于 $i = 2$ 显然成立。现在假设 $m \geq 2$ ，且对于任意的 $l \geq 2$ ，有

$$A_1 A_2 \cdots A_l \geq \delta^{l-1} (A_1 + A_2 + \cdots + A_l).$$

我们要证明 (5.2) 对于 $l+1$ 成立。注意到

$$A_1 A_2 \cdots A_{l+1} \geq \delta^{l-1} (A_1 + A_2 + \cdots + A_l)(A_{l+1}),$$

且

$$\begin{aligned} (A_1 + A_2 + \cdots + A_l)(A_{l+1}) &\geq \delta\{(A_1 + A_{l+1}) + \cdots + (A_l + A_{l+1})\} \\ &\geq \delta\{A_1 + \cdots + A_l + A_{l+1}\}. \end{aligned}$$

因而

$$A_1 A_2 \cdots A_{l+1} \geq \delta^l \{A_1 + \cdots + A_l + A_{l+1}\}.$$

引理得证。 ■

引理 5.9 设在区间 $[t, \tau]$ 上图序列 $\{G_t, G_{t+1}, \dots, G_\tau\}$ 是联合连通的，那么矩阵乘积 $P(t)P(t+1) \cdots P(\tau)$ 是 SIA 矩阵。

证明: 记图 $\{G_t, G_{t+1}, \dots, G_\tau\}$ 的并对应的图的度矩阵和邻接矩阵分别为 D 和 A , 很显然矩阵 $D^{-1}A$ 为遍历矩阵, 并且

$$D \geq D_i, \quad A_t + \dots + A_\tau \geq A.$$

利用这些性质和引理 5.8, 我们有

$$\begin{aligned} P(t)P(t+1)\cdots P(\tau) &\geq \gamma(P(t) + P(t+1) + \dots + P(\tau)) \\ &= \gamma(D^{-1}(t)A(t) + D^{-1}(t+1)A(t+1) + \dots + D^{-1}(\tau)A(\tau)) \\ &\geq \gamma D^{-1}(A(t) + A(t+1) + \dots + A(\tau)) \geq \gamma D^{-1}A, \end{aligned}$$

矩阵 $D^{-1}A$ 为遍历矩阵, 那么矩阵乘积 $P(t)P(t+1)\cdots P(\tau)$ 也为遍历矩阵。■

定理 5.7 的证明: 记 T^* 为区间 $[t_{j-1}, t_j]$ 长度的最大值, 且记

$$\Phi(t_j, t_{j-1}) = P(t_j) \cdots P(t_{j-1})$$

则

$$x(t_j + 1) = \Phi(t_j, t_{j-1})\Phi(t_{j-1}, t_{j-2})\cdots\Phi(t_1, 0)x(0)$$

根据图 G_t 的生成方式, 我们知道 $\Phi(t_j, t_{j-1})$ 为有限个, 且每一个为 SIA 阵, 有限个 $\Phi(t_j, t_{j-1})$ 的乘积也为 SIA 阵。根据 Wolfowitz 定理

$$\lim_{j \rightarrow \infty} \Phi(t_j, t_{j-1})\Phi(t_{j-1}, t_{j-2})\cdots\Phi(t_1, 0) = \mathbf{1}c^\tau,$$

其中 c 为一列向量。另外, 注意到 $x_{\min}(t)$ 与 $x_{\max}(t)$ 为单调有界序列, 因而必有极限。据此定理得证。■

三、考虑具有切换拓扑结构的连续系统。

我们先介绍 Laselle 不变原理。对于下面的连续系统

$$\dot{x}(t) = f(x(t)), \tag{5.3}$$

其中 $f: R^n \rightarrow R^n$ 是局部 Lipschitz 连续的, 且不失一般性我们假设 $f(0) = 0$. 记 C^1 为连续可微的函数类。我们说函数 $V: R^n \rightarrow R$ 是正定的, 如果 $V(0) = 0$, 且对所有的 $x \neq 0$ 有 $V(x) > 0$; 函数 V 是半正定的, 如果对所有的 x 有 $V(x) \geq 0$; 如果 V 满足当 $\|x\| \rightarrow \infty$ 时有 $V(x) \rightarrow \infty$, 则称 V 是径向无界的 (radially unbounded)。

定义 5.5 系统 (5.3) 的关于原点的一个 *Lyapunov* 函数是一个实值的、正定的 C^1 函数 $V: R^n \rightarrow R$, 使得对所有沿着 (5.3) 的轨线 $x \neq 0$ 有 $\dot{V}(t) < 0$; 系统 (5.3) 的关于原点的一个弱 *Lyapunov* 函数是一个实值的、正定的 C^1 函数 $V: R^n \rightarrow R$, 使得对所有沿着 (5.3) 的轨线 $x \neq 0$ 有 $\dot{V}(t) \leq 0$ 。

定理 5.8 (Laselle 不变原理) 设 $V: R^n \rightarrow R$ 为 (5.3) 的一个弱 *Lyapunov* 函数, M 为包含在 (关于微分包含) 集合 $\{x \in R^n | \dot{V}(x) = 0\}$ 的最大不变集, 则 (5.3) 的保持有界的解 $x(t)$ 使得

$$\inf_{y \in M} \|x(t) - y\| \rightarrow 0, \text{ 当 } t \rightarrow \infty.$$

考虑如下的连续系统

$$\dot{x}_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)), i = 1, 2, \dots, n, t \in [0, \infty).$$

取 Lyapunov 函数

$$V(x(t)) = \frac{1}{2}x^\tau(t)x(t),$$

很显然

$$\dot{V}(x(t)) = -x^\tau(t)L(G)x(t).$$

当图 G_t 为无向连通图时, 最大不变集包含在

$$\{x \in R^n | \dot{V}(x(t)) = 0\} = \text{span}\{\mathbf{1}\}.$$

根据 Lasalle 不变原理我们可以知道系统收敛到集合 $\text{span}\{\mathbf{1}\}$ 。

当邻居关系图不是在每个时刻都连通时, 我们很难定义一段时间区间上的联合连通性。另外, 在每个时刻邻居关系图都发生变化会产生一些令人困惑的问题。下面考虑如下连续的动态系统:

$$\dot{x}_i(t) = u_i(t), i = 1, 2, \dots, n, t \in [0, \infty),$$

我们对连续的时间轴进行采样 $0 = t_0, t_1, \dots, t_k, \dots$ 满足 $|t_{i+1} - t_i| = \delta > 0$ 。假设在时间区间 $[t_i, t_{i+1}]$ 上邻居图保持不变。对每个个体 i , 我们设计如下的控制律:

$$u_i(t) = \sum_{j \in N_i(t_k)} (x_j(t) - x_i(t)), \quad t \in [t_k, t_{k+1}].$$

将控制律带入系统可得如下的混杂系统:

$$\dot{x}_i(t) = \sum_{j \in N_i(t_k)} (x_j(t) - x_i(t)), t \in [t_k, t_{k+1}], i = 1, 2, \dots, n.$$

记 $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^\tau$, 我们可以将上面的系统写成矩阵的形式。很容易可以得到系统的解为:

$$z(k+1) = x((k+1)\delta) = \exp\{-\delta L(t_k)\}x(k\delta).$$

引理 5.10 矩阵 $\exp\{-\delta L(t_k)\}$ 的性质如下:

- (1) 此矩阵是非负随机矩阵;
- (2) 矩阵的元素 $[\exp\{-\delta L(t_k)\}]_{ij} > 0$ 当且仅当 $i = j$ 或者 i 与 j 之间存在路径。

证明: 将 $\exp\{-\delta L(t_k)\}$ 做如下变形:

$$\exp\{-\delta L(t_k)\} = \exp\{-\delta n\} \exp\{\delta(nI - L(t_k))\},$$

其中 $(nI - L(t_k))$ 为非负矩阵, 且每个正元素必大于或者等于 1。利用如下展开

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!},$$

可知矩阵 $\exp\{-\delta L(t_k)\}$ 为非负矩阵。注意到 $L(t_k)\mathbf{1} = \mathbf{0}$, 因而

$$\exp\{-\delta L(t_k)\}\mathbf{1} = \sum_{i=0}^{\infty} \frac{(-\delta L(t_k))^i}{i!} \mathbf{1} = \mathbf{1}.$$

矩阵 $\exp\{-\delta L(t_k)\}$ 为随机矩阵。进一步,

$$\exp\{\delta(nI - L(t_k))\} = \sum_{k=0}^{\infty} \frac{\delta^k (nI - L(t_k))^k}{k!},$$

很显然, 上矩阵的主对角线元素一定大于 0。当 i 与 j 之间存在长度为 l 的路径时, 根据离散时间系统的分析, 我们可知

$$[(\delta(nI - L(t_k)))^l]_{ij} \geq \delta^{n-1} > 0.$$

而当 i 与 j 之间不存在路径时,

$$[(\delta(nI - L(t_k)))^n]_{ij} = \sum_{j_1, j_2, \dots, j_n} l_{ij_1} l_{j_1 j_2} \cdots l_{j_{n-1} j_n} = 0.$$

引理得证。■

据上面的引理, 我们可以推出:

$$z(k+1) \leq \left(1 - n \left(\frac{\delta^{n-1}}{n!}\right)\right) z(k) = \left(1 - \frac{\delta^{n-1}}{(n-1)!}\right) z(k).$$

系统以指数的速度同步。

6 集体行为行为的控制与干预。

多个体系统的控制不同于单个个体的控制, 我们需要利用每个个体接收到的信息设计 (全部或部分) 个体的规则, 使得系统完成某项任务或者展现某种行为。在分布式控制问题研究中, 一种常用的方法是利用势函数设计个体的规则, 课上我们讲了一个例子, 见 MAS9.ppt。下面我们讲一下牵制控制。

一、牵制控制 (Pinning Control)。

牵制控制本质上也是分布式控制, 不同之处在于牵制控制是控制部分而非全部节点, 它是通过有选择地对网络中的少部分节点施加控制而使得整个网络达到所期望的行为。牵制控制中我们要考虑: 选择什么样的节点进行控制? 选择多少节点进行控制? 施加什么样的控制?

常用的牵制控制策略有两种:

- 特定牵制控制策略 (Specific pinning control scheme): 有选择地选取牵制控制节点, 常见的包括按照某种节点重要性指标选取控制节点, 例如选择网络中度最大或者介数最大的节点;
- 随机控制策略 (Random pinning control scheme): 随机选择若干个节点施加控制。

根本上我们要分析闭环系统的行为, 它受系统本身的动力学、网络结构、耦合强度以及控制策略、控制节点等因素影响。

我们以一个例子说明牵制控制的可行性。

假设系统由四个节点组成, 节点的状态 $x_i(t), i = 1, 2, 3, 4$ 按照如下方式演化:

$$\begin{cases} \dot{x}_1(t) = Ax_1 - 2x_1 + x_2 + x_3 \\ \dot{x}_2(t) = Ax_2 - x_2 + x_1 \\ \dot{x}_3(t) = Ax_3 - x_3 + x_1 \\ \dot{x}_4(t) = Ax_4 \end{cases},$$

其中

$$x_i \in R^2, \quad A = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

我们希望系统能同步, 并且同步到状态 $[0, 0]^T$ 上。

如果不施加控制, 对于由节点 1,2,3 组成的子系统,

$$\begin{pmatrix} \dot{x}_1 \\ x_2 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}$$

其中

$$A = \begin{pmatrix} -1 & 0.5 & 1 & 0 & 1 & 0 \\ 0.5 & -1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 1 & 0 & 0 & 0.5 & 0 \end{pmatrix}.$$

矩阵 A 的特征值为 -2.5000, -1.5000, -0.5000, 0.5000, 0.5000, 1.5000, 系统既不稳定 (稳定需要所有特征值在左半平面) 也不同步。该子系统在不施加控制的情况下不会同步。而节点 4 是孤立节点, 其状态方程为:

$$\begin{cases} \dot{x}_{41} = x_{41} + 0.5x_{42} \\ \dot{x}_{42} = 0.5x_{41} + x_{42} \end{cases}$$

很容易看出 $(0,0)^T$ 是系统的不稳定平衡点。如果不施加控制, 系统会发散。

我们可以在每个节点上都施加反馈控制, 这样系统变为

$$\begin{cases} \dot{x}_1(t) = (A - 2I)x_1 + x_2 + x_3 - \kappa x_1 \\ \dot{x}_2(t) = (A - I)x_2 + x_1 - \kappa x_2 \\ \dot{x}_3(t) = (A - I)x_3 + x_1 - \kappa x_3 \\ \dot{x}_4(t) = Ax_4 - \kappa x_4 \end{cases},$$

即

$$\begin{pmatrix} \dot{x}_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \overline{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \kappa \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

其中

$$\overline{A} = \begin{pmatrix} -1 & 0.5 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0.5 & -1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 1 \end{pmatrix}.$$

取 $\kappa = 2$, 此时矩阵 $\overline{A} - \kappa I$ 的特征值为 -4.5000, -3.5000, -2.5000, -1.5000, -1.5000, -1.5000, -0.5000, -0.5000, 此时系统会同步到 $[0,0]^T$ 点。这意味着对每个节点都进行分布式控制可以保证系统同步到期望的状态。

如果只控制我们对部分节点施加控制, 我们对节点 2,3,4 施加控制, 此时系统为:

$$\begin{pmatrix} \dot{x}_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \overline{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \kappa \begin{pmatrix} 0 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

若取 $\kappa = 11$ ，此时系统的特征值 -11.6962， -11.5000, -10.6962, -10.5000, -10.5000, -9.5000, -1.3038, -0.3038，系统能同步到期望的状态。这意味着牵制控制是可行的。

如果我们只控制节点 1 和 4，此时系统为：

$$\begin{pmatrix} \dot{x}_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \bar{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \kappa \begin{pmatrix} x_1 \\ 0 \\ 0 \\ x_4 \end{pmatrix},$$

取 $\kappa = 2$ ，此时特征值为 -4.0616, -3.0616, -1.5000, -0.5000, -0.5000, 0.0616, 0.5000, 1.0616. 并且可以看出单纯增大反馈增益，不能把系统平衡到 $(0, 0)^T$ 上。

目前该方面理论研究结果还比较少，只能对具体的问题去尝试。

二、干预或者软控制 (intervention or soft control)。

为什么要干预多个个体系统？

- 在很多实际情形中，个体的局部规则是不能设计和人为改变的。
- 系统的动力学行为是由所有个体的初始状态与参数决定的，系统自发产生的行为未必是我们所期望的。
- 怎样在不破坏系统中已有个体规则的基础上引导整个系统达到所希望的行为？

我们讲了两种干预控制方案：软控制与加入“带信息”的个体。并通过案例分析证明了可行性，见 MAS10.ppt。

三、网络的能控性。

能控性是控制论中的一个核心概念。如果系统能在有限时间内，从任意的初态可以控制到任意的终态，那么我们就称此系统是能控的。

对于线形时不变系统

$$\dot{x}(t) = Ax(t) + Bu(t),$$

其中 $x(t) \in R^{N \times 1}$ 为状态向量， $u(t) \in R^{M \times 1}$ 为控制向量， $A \in R^{N \times N}$ 为状态矩阵， $B \in R^{N \times M}$ 为输入矩阵， Kalman 给出了此系统能控的判别式，也就是

$$\text{rank} C = N,$$

其中

$$C = [B, AB, \dots, A^{N-1}B].$$

我们换一个角度去理解线性系统，将 $x(t)$ 看成由 N 个个体的状态排列形成的，每一维对应一个个体，那么状态矩阵 A 描述了个体（节点）之间的连接，每个控制分量也看成个体，输入矩阵 B 反映了控制个体与状态个体之间的关联。

对于图 5 所示的例子，其中 $N = 4, M = 2$ 。很容易计算

$$C = \begin{pmatrix} b_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_2 & a_{21}b_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{31}b_1 & 0 & a_{34}a_{41}b_1 & 0 & 0 & 0 \\ 0 & 0 & a_{41}b_1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

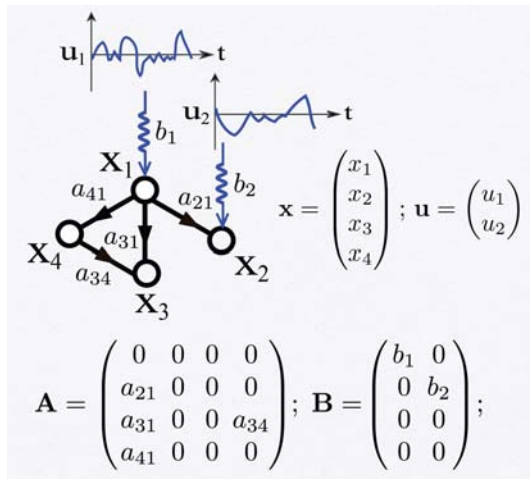


图 5

此时 $\text{rank}(C) = N$ 。系统中有两个驱动节点 x_1, x_2 。

在复杂网络的能控性研究中，我们关心的问题有：

1. 要使得系统能控，最少需要多少驱动节点 (N_D)?
2. 如何有效地辨识它们?
3. 什么样的网络性质来决定 N_D ?
4. 网络能控性对节点 / 边失效或者移出有没有鲁棒性?

要解决这些问题，遇到的主要问题是：

1. 能控性依赖于系统参数，但在很多网络系统特别是生物系统中，系统的参数可能是未知的；
2. 能控矩阵 C 是 $N \times NM$ 维的矩阵，当 M, N 很大时，验证矩阵的秩是不容易的；
3. 如果用蛮力搜索，需要验证 $2^N - 1$ 个组合的秩。

以上这些问题为我们研究大型复杂网络的能控性带来了困难。为了解决这些问题，我们引入结构能控性的概念。结构能控性是由 Ching-Tai Lin 在 1974 年提出的。所谓结构能控性，就是我们不关心矩阵 A 和 B 中元素数值的大小，而只研究网络结构对系统能控性的影响。也就是说，如果对一种网络结构我们能找到一组参数使得系统能控（这意味着具有这种结构的系统不能控的测度为 0），那么该结构就称为结构能控的。图 6 列举了几个例子，通过计算矩阵的秩，我们可以发现 (a)(c)(d) 结构能控，而 (b) 结构不能控。

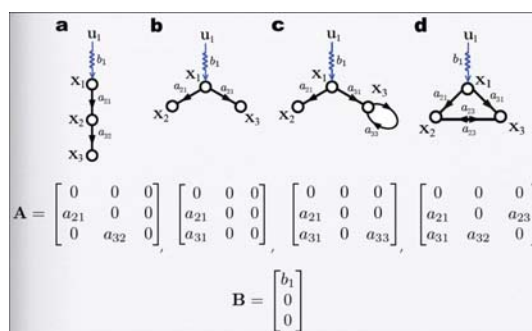


图 6

那么什么样的结构能使得系统结构能控呢？下面介绍几个关于图的概念。参见图 7 的示例。

定义 6.1 1. 有向图 $G(A, B)$ 中的一个状态节点称为不可达的 (inaccessibility)，如果不存在从输入节点到达该状态节点的有向路径。

2. 有向图 $G(A, B)$ 包含一个扩张 (dilation) 当且仅当存在一个子集 $S \subset V_A$, 使得指向集合 S 的节点数目小于集合 S 中的结点数, 即 $|T(S)| < |S|$ 。这里集合 $T(S)$ 定义为直接有边指向集合 S 的所有结点的集合。

3. 干 (stem) 是指源自于输入节点的一条简单路径 (即经过的节点各不相同的路径)。干的起点称为根, 终点称为顶点。

4. 芽 (bud) 是指一个有向的简单圈 C 以及指向 C 中某一节点的一条边 e 。边 e 也称为芽的显著边。

5. 掌是一个递归定义的子图。一个干是一个掌; 给定一个干 S_0 和一些芽 B_1, \dots, B_l , 如果对每个 $i (1 \leq i \leq l)$, B_i 的显著边的始点不是干 S_0 的顶点, 而是唯一一个同时属于 B_i 和 $S_0 \cup B_1 \cup \dots \cup B_{i-1}$ 的节点, 那么 $S_0 \cup B_1 \cup \dots \cup B_l$ 是一个掌。节点不相交的掌的集合称为掌群 (Cacti)。

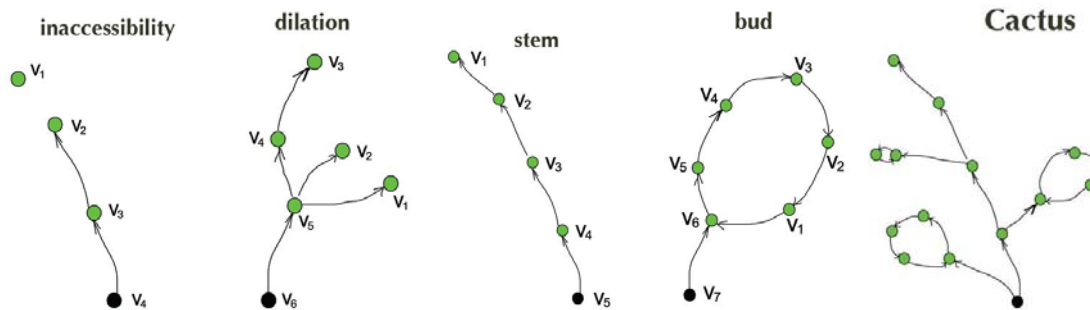


图 7

可以证明掌 (群) 是既不包含不可达点又没有扩张的最小结构。关于结构能控性, 有如下定理:

定理 6.1 下面的三个论断是等价的:

1. 线性系统 (A, B) 是结构能控的;
2. 有向图 $G(A, B)$ 既不包含不可达点又没有扩张;
3. $G(A, B)$ 是由掌群张成的图。

那么给定一状态矩阵 A , 如何选择驱动节点?

定义 6.2 有向网络 $G(A)$ 的边的子集 M 称为一个匹配 (Matching), 如果 M 中的任意两条边都既没有公共的始点也没有公共的终点。如果一个节点是 M 的一条边的终点, 那么该节点就称为匹配节点 (matched node), 否则称为未匹配节点 (unmatched node)。匹配节点数最多的匹配称为最大匹配 (maximum matching)。一个匹配称为完全匹配 (perfect matching), 如果网络中所有节点都是匹配节点。示例见图 8。

下面的节点给出了驱动节点的数量以及怎样选择驱动节点。

定理 6.2 (最少输入定理) 网络 $G(A)$ 结构能控所需要的驱动节点的数为

$$N_D = \max\{1, N - |M|\},$$

其中 $|M|$ 为网络 $G(A)$ 的最大匹配所对应的匹配节点数。

7 博弈论简介。

博弈论研究包含相互依存情况中理性行为的研究。在博弈中, 任何一个局中人受其他局中人行行为的影响, 反过来, 他的行为也影响到其他局中人; 博弈中一个理性的决策必定建立在预测其他局中人的反映之

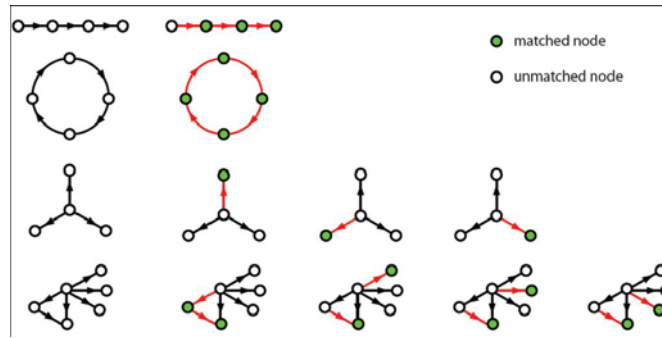


图 8

上，这里的理性指尽量极大化自己的收益，而不是打败对方；一个局中人将自己置身于其他局中人的位置并为他着想从而预测其他局中人将选择的行动，在这个基础上局中人决定自己最理想的行动。博弈本身也可以看成是一个多个体系统。

一个博弈由以下要素组成：

1. 参与人 (player)：也称为局中人，指博弈的参加者，他们有自己的切身利益，有可供选择的行动集合，理性地选择自己的策略（最大化自己的收益），可以是个人，也可以是团体；

2、行动 (action, move)：行动是指参与人在博弈进程中轮到自己选择时所作的某个具体决策，记做 $S_i = \{s_{i1}, s_{i2}, \dots, s_{il_i}\}$ ；

3、收益 (payoff)：收益是参与人在博弈结束后从博弈中获得的效用，一般是所有参与人的策略或行动的函数，这是每个参与人最关心的东西。

$$u_i : \times_{i=1}^n S_i \rightarrow R.$$

4. 策略 (strategy)：策略是指参与人选择行动的规则，即在博弈进程中，什么情况下选择什么行动的预先安排；

5. 信息 (information)：信息信息对博弈来说很重要，信息是指参与人在博弈中所知道的关于自己以及其他参与人的行动、策略及其收益函数等的知识。

根据不同的方式，博弈有不同的类型

- 根据参与人的多少，可将博弈分为两人博弈与多人博弈；
- 根据参与人是否合作，可将博弈分为合作博弈与非合作博弈；在合作博弈中，局中人之间形成一个具有约束力的协议；
- 根据收益的不同，又可分为零和博弈与非零和博弈；
- 从行动的先后次序来分，博弈可以分为静态博弈与动态博弈：静态博弈指在博弈中，参与人同时选择行动，或虽非同时但后行动者并不知道前行动者采取了什么具体行动；动态博弈指的是参与人的行动有先后顺序，且后行动者能够观察到先行动者所选择的行动的博弈；
- 从参与人对其他参与人的各种特征信息的获得差异来分，博弈可分为完全信息博弈和不完全信息博弈。完全信息指的是每一个参与人都知道其他参与人的策略集合及收益函数；否则就是不完全信息。完美信息不同于完全信息，完美信息是指在动态博弈中，如果所有局中人都知道前一步其他局中人的行动，比如围棋等。

常见的一种方式是将后两者结合起来，可以得到以下四种不同类型的博弈：完全信息静态博弈，完全信息动态博弈，不完全信息静态博弈不完全信息动态博弈。

一个博弈有两种表述方式：一种是策略式表述，一种是扩展式博弈。策略式博弈比较适合静态博弈。我们这里只讲这一种表述方式。在策略式表述中，要给出：

1. 博弈参与人的集合，记做 $N = \{1, 2, \dots, n\}$;
2. 每个参与人的策略空间： $S_i = \{s_{i1}, s_{i2}, \dots, s_{ik_i}\}$;
3. 每个参与人的支付函数： $u_i(s_1, s_2, \dots, s_i, \dots, s_n), i = 1, 2, \dots, n$.

我们用 $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$ 代表策略式表述博弈。

下面介绍混合策略。

定义 7.1 在 n 个参与人博弈的策略式表述 $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$ 中，设参与人 i 有 k_i 个纯策略： $S_i = \{s_{i1}, \dots, s_{ik_i}\}$ ，那么概率分布 $\sigma_i = \{\sigma_{i1}, \dots, \sigma_{ik_i}\}$ 称为 i 的一个混合策略，这里 $\sigma_{il} (l = 1, \dots, k_i)$ 是参与人 i 选择 s_{il} 的概率，对于所有的 $l = 1, \dots, k_i, 0 \leq \sigma_{il} \leq 1, \sum_{l=1}^{k_i} \sigma_{il} = 1$ 。

我们用 Σ_i 代表 i 的混合策略空间 ($\sigma_i \in \Sigma_i$)， $\sigma = \{\sigma_1, \dots, \sigma_i, \dots, \sigma_n\}$ 代表混合策略组合，其中 σ_i 为 i 的一个混合策略， $\Sigma = \times_i \Sigma_i$ 代表混合策略组合空间。因为混合策略是纯策略空间上的一个概率分布，所以支付具有不确定性。此时，参与人关心的是期望效用：

$$u_i(\sigma_1, \sigma_2, \dots, \sigma_n) = \sum_{j_1=1}^{k_1} \dots \sum_{j_n=1}^{k_n} \left(\prod_{i=1}^n \sigma_{ij_i} \right) u_i(s_{1j_1}, s_{2j_2}, \dots, s_{nj_n}), \quad i = 1, 2, \dots, n.$$

对于 2 人博弈情形，期望的收益为：

$$u_i(\sigma_1, \sigma_2) = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \sigma_{1i} \sigma_{2j} u_i(s_{1i}, s_{2j}), \quad i = 1, 2, .$$

从上面的描述我们可以看出：每个个体的收益都依赖于所有参与人。那么，博弈的解怎样去描述？

1. 占优策略。记 S_{-i} 为除去参与人 i 的其他参与人所形成的策略集合。

定义 7.2 策略 s_i^* 称为参与人 i 的（严格）占优策略，如果

$$u_i(s_i^*, s_{-i}) > u_i(s'_i, s_{-i}), \forall s_{-i} \in S_{-i}, \forall s'_i \neq s_i^*.$$

注 7.1 占优策略不依赖于其他参与人的策略，也就是说，不管对手选什么，他的最优策略是唯一的，这样的策略称为“占优策略”

例 7.1 囚徒困境模型。博弈有两个参与人，每个参与人可以选择两个行动： C ， D 。其收益矩阵见图 9，其中矩阵中的第一个元素为行参与者的收益，第二个元素为列参与者的收益。根据定义，我们可以知道 (D, D) 是占优策略。

		Prisoner B	
		C	D
Prisoner A	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

图 9

2. 重复剔除劣势策略。先给出劣势策略的定义：

定义 7.3 对于局中人 i 的策略空间中的某个纯策略 s_i ，如果存在（混合）策略 $\sigma_i^* \in \Sigma_i$ 使得

$$u_i(\sigma_i^*, s_{-i}) \geq u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i},$$

成立，且在 S_{-i} 中至少存在一个纯策略组合 $s_{-i}^* \in S_{-i}$ ，使得上式的不等号严格成立，即

$$u_i(\sigma_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*),$$

则称纯策略 s_i 为局中人 i 的弱劣纯策略。进一步，如果

$$u_i(\sigma_i^*, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i},$$

则称纯策略 s_i 为局中人 i 的严劣纯策略。

我们可以通过重复剔除劣势策略给出博弈的解。

例 7.2 考虑如下有两个参与人的博弈，其收益矩阵见图 10 中的最左图。经过重复剔除后我们可以得到占优策略为 (U, L) 。

	L	M	R
U	4,3	5,1	6,2
M	2,1	8,4	3,6
D	3,0	9,6	2,8

	L	R
U	4,3	6,2
M	2,1	3,6
D	3,0	2,8

	L	R
U	4,3	6,2

	L
U	4,3

图 10

3. 纳什均衡。我们先给出定义：

定义 7.4 定义（纳什均衡）：在 n 个参与人博弈的策略式表述 $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$ 中，混合策略组合 $\sigma^* = \{\sigma_1^*, \dots, \sigma_i^*, \dots, \sigma_n^*\}$ 是一个纳什均衡，如果对所有的 $i = 1, 2, \dots, n$ ，下式成立：

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*), \quad \forall \sigma_i \in \Sigma_i.$$

我们先看一下纯策略的纳什均衡。下面给一个例子。

例 7.3 考虑卖家 A 、 B 形成的博弈，他们的行动集有两个：高价，低价。其收益矩阵如图 11。此博弈有两个纯策略纳什均衡点： (H, H) ， (L, L) 。

		B	
		High Price	Low price
A	High Price	(100, 100)	(0, 10)
	Low price	(10, 0)	(5, 5)

图 11

我们可以这样分析:

1. 从 (H,H) 出发。固定 A 的 H 策略, 当 B 的策略由 H 变为 L 时, 收益由 100 变为 10, 所以 B 没有理由变化自己的策略。同样, 当 B 的策略固定为 H 时, A 也没有理由改变自己的策略。所以 (H,H) 是一个纳什均衡。

2. 从 (L,L) 出发。固定 A 的 L 策略, 当 B 的策略由 L 变为 H 时, 收益由 5 变为 0, 所以 B 没有理由变化自己的策略。同样, 当 B 的策略固定为 L 时, A 也没有理由改变自己的策略。所以 (L,L) 是一个纳什均衡。

3. 从 (H,L) 出发。固定 A 的 H 策略, 当 B 的策略由 L 变为 H 时, 收益由 10 变为 100, 所以 B 将自己的策略由 L 变为 H, 博弈的状态变为 (H,H)。(H,L) 不是一个纳什平衡。同理, (L,H) 也不是一个纳什均衡。

有些博弈不存在纯策略纳什均衡。比如硬币匹配博弈。

例 7.4 考虑 A、B 两个参与人形成的博弈, 他们的行动集有两个: head, tail。其收益矩阵如图 12。

		B	
		head	Tail
A	head	(1,-1)	(-1,1)
	Tail	(-1,1)	(1,-1)

图 12

1. 从 (H,H) 出发。固定 A 的 H 策略, 当 B 的策略由 H 变为 T 时, 收益由 -1 变为 1, 所以 B 会改变自己的策略, 博弈的状态变为 (H,T)。所以 (H,H) 不是一个纳什平衡。

2. 从 (H,T) 出发。固定 A 的 H 策略, B 不会改变自己的策略; 但当 B 固定自己的 T 策略, 当 A 的策略由 H 变为 T 时, A 的收益由 -1 变为 1, 所以 A 会改变自己的策略, 博弈的状态变为 (T,T)。所以 (H,T) 不是一个纳什平衡。

3. 从 (T,T) 出发。固定 A 的 T 策略, 当 B 的策略由 T 变为 H 时, B 的收益由 -1 变为 1, 所以 B 会改变自己的策略, 博弈的状态变为 (T,H)。所以 (T,T) 不是一个纳什平衡。

4. 从 (T,H) 出发。固定 B 的 H 策略, 当 A 的策略由 T 变为 H 时, A 的收益由 -1 变为 1, 所以 A 会改变自己的策略, 博弈的状态变为 (H,H)。所以 (T,H) 不是一个纳什平衡。

综上 (H,H)(H,T)(T,T)(T,H) 都不是纳什平衡。此系统不存在纯策略纳什平衡点, 但存在混合策略的纳什平衡点。我们怎么计算?

设 $\sigma_1 = (x, 1-x), \sigma_2 = (y, 1-y)$, 则在 (σ_1, σ_2) 策略下系统的收益为:

$$u_1(\sigma_1, \sigma_2) = xy - x(1-y) - (1-x)y + (1-x)(1-y) = 4xy - 2x - 2y + 1;$$

$$u_2(\sigma_1, \sigma_2) = -xy + x(1-y) + (1-x)y - (1-x)(1-y) = -4xy + 2x + 2y - 1$$

上面的第一式对 x 求导, 第二式对 y 求导, 并另两个导函数等于 0,

$$\begin{cases} 4y - 2 = 0 \\ -4x + 2 = 0 \end{cases}$$

我们可以得到 $x = 1/2, y = 1/2$ 。 (σ, σ) 为该博弈的纳什平衡点，其中 $\sigma = \{1/2, 1/2\}$ 。

用此方法可求得例 7.3 的混合策略纳什平衡点为 (σ, σ) ，其中 $\sigma = \{1/19, 18/19\}$ 。值得注意的是，该方法只适用于求混合策略的纳什均衡点。

注 7.2 1. 占优策略一定是纳什均衡；

2. 对博弈双方来说，纳什平衡未必是最好的，如囚徒困境。纳什平衡只是多个（自私的）参与人追求自己利益最大化的结果；

3. 纳什平衡只是告诉我们，系统一旦达到这个状态，任何参与人没有改变自己策略的动机。但纳什平衡没有告诉我们博弈怎么达到此状态。

看一下 MAS11.ppt 中的其它求纳什均衡的例子。

定理 7.1 对于有限（参与人，纯策略）的博弈，一定存在纳什均衡。

下面我们考虑有多个个体组成的博弈。每两个参与人是按照图 13 所示的收益矩阵进行博弈，假设参数满足 $T + S \leq 2R$ 。

	C	D
C	R	S
D	T	P

图 13

- 当 $T > R > P > S$ 时，对应于囚徒困境
- 当 $T > R > S > P$ 时，对应于雪堆博弈，鹰鸽博弈，胆小鬼博弈
- 当 $R > T > P > S$ 时，对应于猎鹿博弈

考虑均匀混合的种群，每个个体可以与中群众的其他所有个体博弈。假设采用合作策略的个体比例为 x ，选择背叛的比例为 y ，则种群中合作 / 背叛者的收益为：

$$\begin{cases} P_C = Rx + Sy \\ P_D = Tx + Py \end{cases}$$

利用复制动力学（也称模仿者动态）描述策略的动态变化：策略比例的变化速度与收益成正比，

$$\begin{cases} \dot{x} = x(P_C - \phi) \\ \dot{y} = y(P_D - \phi) \end{cases}$$

其中 $\phi_i = xP_C + yP_D$ 是种群的平均收益。这样种群的总数保持不变。利用 $x + y = 1$ 可得：

$$\dot{x} = x(1-x)((R-S-T+P)x + S-P).$$

此方程的相平面图见图 14：

- 当 $T > R > P > S$ 时，平衡点有两个： $0, 1$ ，其中 0 是稳定的平衡点，意味着背叛占主导；
- 当 $T > R > S > P$ ，平衡点有三个： $0, 1, x^* = \frac{P-S}{R-S-T+P}$ ，其中 x^* 是稳定的平衡点，意味着合作与背叛共生。

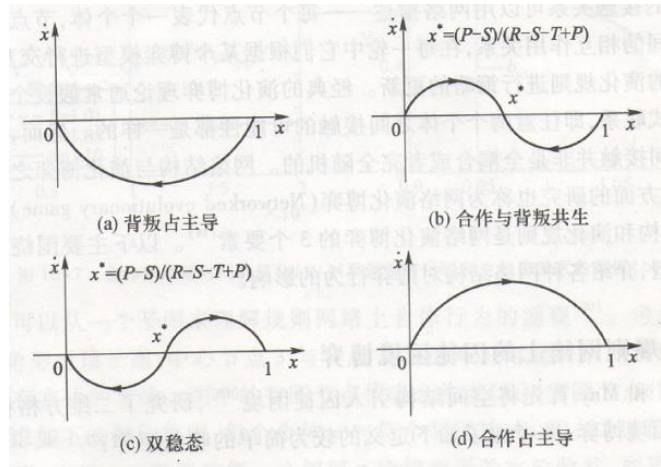


图 14

c) 当 $R > T > P > S$, 平衡点有三个: $0, 1, x^* = \frac{P-S}{R-S-T+P}$, 其中 0 和 1 是稳定的平衡点, 而 x^* 不稳定, 此时是双稳定状态。

d) 当 $R > T$ 且 $P < S$ 时, 平衡点有两个: $0, 1$, 其中 1 是稳定的平衡点, 意味着合作占主导;。

对此我们或许可以从进化稳定策略中得到一些解释。

进化稳定策略由 John Maynard Smith 与 George R. Price 引入。进化稳定策略 (evolutionarily stable strategy) 是这样一种策略: 如果群体中所有个体都采取这种策略, 那么在自然选择的力量下没有变异策略可以入侵该群体。

定义 7.5 策略 x 称为进化稳定策略, 如果对任意的 $y, y \neq x$, 使得下式对任意小的正数 ε 成立,

$$U(x, (1 - \varepsilon)x + \varepsilon y) > U(y, (1 - \varepsilon)x + \varepsilon y).$$

根据收益的线性性质, 上式等价于下面的两个条件同时成立:

$$(1) u(x, x) \geq u(x, y);$$

$$(2) \text{当策略 } y \text{ 满足 } u(x, x) = u(x, y) \text{ 时, } u(x, y) \geq u(y, y).$$

注 7.3 • 演化稳定策略是一个纳什均衡的精练;

- 演化稳定策略一定是纳什均衡; 严格的纳什均衡一定是一个演化稳定策略;
- 纯策略的纳什均衡未必是演化稳定策略, 演化稳定策略也可以是混合策略;
- ESS 不一定存在。

ESS 主要处理单一大群体中的对称配对博弈, 它不能处理互动作用同时发生于多于两个个体的博弈

NE 依赖于参与者的认知能力。假设参与人知道博弈的结构, 并通过预测对方行动实现自己利益的最大化, 并假设这些信息是共同信息

ESS 假设参与者的策略是可遗传的, 参与者并没有对策略进行控制, 并且不必知道博弈的信息。他们只受进化压力的影响: 适应度大的个体生存下来并有多数的后代, 变异发生的概率非常小。它在进化的意义下稳定: 一旦群体中所有个体都选择此策略, 自然选择可以阻止变异的入侵

例 7.5 考虑两人鹰鸽博弈, 每个参与者可以选择两个行动: 鹰策略, 鸽策略。收益矩阵如图 15 所示, 其中 $(b < c)$ 。

		B	
		D	H
A	D	$b/2$	0
	H	b	$(b-c)/2$

图 15

鹰鸽博弈存在三个纳什平衡点： (D, H) , (H, D) 和 (σ^*, σ^*) ，其中 $\sigma^* = (1 - b/c, b/c)$ 。我们考察一下哪个纳什平衡点可能进化稳定策略。

1. 因为 $u(D, D) < u(H, D)$ ，鸽策略可以被鹰策略入侵，鸽策略不是 ESS；
2. 因为 $u(H, H) < u(D, H)$ ，鹰策略可以被鸽策略入侵，鹰策略也不是 ESS；
3. 混合策略 σ^* 是 ESS，因为它不能被其它策略入侵，我们按照定义进行验证。假设入侵策略为 $\sigma = (y, 1 - y)$ ，则

$$u(\sigma^*, \sigma^*) = \frac{b(c-b)}{2c}, u(\sigma, \sigma^*) = \frac{b(c-b)}{2c}.$$

因而对任意的策略 σ ，都有 $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$ 。下面考察是否有 $u(\sigma^*, \sigma) > u(\sigma, \sigma)$ 。

$$u(\sigma^*, \sigma) = \frac{b(b-c)}{2c} + by, \quad u(\sigma, \sigma) = \frac{b-c}{2} + cy - \frac{cy^2}{2},$$

$$u(\sigma^*, \sigma) - u(\sigma, \sigma) = \frac{c}{2}(y - \frac{c-b}{c}) > 0, \text{ 当 } y \neq 1 - \frac{b}{c} \text{ 时.}$$

根据定义可知策略 σ^* 是进化稳定策略。