

**HARBIN INSTITUTE OF TECHNOLOGY**

2018年春季学期本科生课程考核

（读书报告、研究报告）

从石头剪刀布出发看博弈论

|  |  |  |  |
| --- | --- | --- | --- |
| 考核科目 | 策略思维----带你走进博弈的奇妙世界 | | |
| 学生所在院（系） | 经济管理学院 | | |
| 学生所在专业 | 经济管理试验班 | | |
| 学生姓名 | 桑浩鑫 | | |
| 学号 | 1171000221 | | |
| 考核结果 |  | 评阅人 |  |

石头剪刀布是在生活中十分常见的一种游戏，本文将以石头剪刀布游戏，及其几个变体游戏为基础，利用传统以及模拟的方法，从完全信息静态博弈到完全信息动态博弈来分析这几个游戏。

# （一）完全信息静态博弈

## （1）零和博弈

最简单的石头剪刀布游戏是由两名参与者在同一时间选择并展示石头、剪刀或者布三者之一，游戏胜负的判定依据：相同为平局，石头胜剪刀，剪刀胜布，布胜石头。若记胜者得1分，负者得-1分，平局双方得0分，则可得如下博弈矩阵。

参与者2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参  与  者  1 |  | 石头 | 剪刀 | 布 |
| 石头 | 0 , 0 | 1 , -1 | -1 , 1 |
| 剪刀 | -1 , 1 | 0 , 0 | 1 , -1 |
| 布 | 1 , -1 | -1 , 1 | 0 , 0 |

这是一个非常典型的零和博弈，可以判断不存在纯策略的纳什均衡。同时这个博弈中也不存在严格劣策略。

假设参与人1的战略是其中且，参与人2的战略是其中且，则对于参与人1战略的最优反应（best reply）如下：

















同理可以得到参与人1对于参与人2战略的最优反应，由纳什均衡定义可知：存在混合策略纳什均衡。

现在我们将这个简单的石头剪刀布游戏稍作修改，将记录得分的规则改为：

记使用石头的胜者得1分，负者得-1分；使用剪刀的胜者的2分，负者得-2分；使用布的胜者得5分，负者得-5分；平局双方得0分。则可得如下博弈矩阵。

参与者2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参  与  者  1 |  | 石头 | 剪刀 | 布 |
| 石头 | 0 , 0 | 1 , -1 | -5 , 5 |
| 剪刀 | -2 , 2 | 0 , 0 | 2 , -2 |
| 布 | 5 , -5 | -2 , 2 | 0 , 0 |

同样，这个博弈也没有纯策略纳什均衡，也没有严格劣策略。

这一次我们通过计算求纳什均衡时作如下分析：假设纳什均衡中，参与人1的战略是其中，那么石头、剪刀和布三个策略的收益应当相等。用反证法证明，若三个策略的收益不等，不妨设石头策略收益更大，因那么参与人1会增加混合策略中石头策略的概率，以此来获得更大的收益，这与参与人1处于纳什均衡的假设矛盾，因此纳什均衡中，参与人1使用石头、剪刀和布三个策略的收益相等。同理，纳什均衡中参与人2三个策略的收益也相同，并且等于参与人1的收益，否则参与人1和参与人2总会选择受益较大的那个策略。又知这是一个零和博弈，两个参与人的收益不可能同时为正或同时为负，那么混合策略纳什均衡中两个参与人三个策略的收益都为0。

由此知：  解得：

对于更加一般的情况，我们将记录得分的规则改为：记使用石头的胜者得a分，负者得-a分；使用剪刀的胜者的b分，负者得-b分；使用布的胜者得c分，负者得-c分；平局双方得0分。利用上述办法同样可以可以求得混合策略纳什均衡是 。

## （2）非零和博弈

如果我们将石头剪刀布游戏的得分规则更加一般化，也就是同一策略获胜的得分与失败的减分不相等，换句话说，整个游戏变成了更加一般的两个参与人，三种可选策略的博弈，下面举出一例。首先游戏的变体应当满足不同参与人同策略具有相同收益的原则，也就是收益是具有对称性的。于是我们假设参与人1与参与人2的博弈矩阵如下。

参与者2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参  与  者  1 |  | 石头 | 剪刀 | 布 |
| 石头 | 1 , 1 | -1 , -2 | -2 , -1 |
| 剪刀 | -2 , -1 | 2, 2 | 5,4 |
| 布 | -1 , -2 | 4 , 5 | -2 , -2 |

我们可以得到上述博弈中存在纯策略纳什均衡（石头，石头），（剪刀，布）和（布，剪刀）。

看过上面的例子后我们来分析一下最为一般的情况。由于我们只关心对手某一特定策略下我们的最优反应，所以我们用下面的收益流向图来表示这一博弈。

石头

剪刀 布

图中有向边指向的的策略是对于有向边出发的策略的最优反应，例如布指向剪刀，即意味着剪刀是面对布收益最高的策略，有向边的权值，就是该最优反应的收益。另外需要指出的是上述利益流向图一定满足以下条件：

1. 每个点的出度至少为1。因为三个策略中，作为对于该策略的反应，至少有一个是收益最大的。
2. 由同一点出发的有向边，权值一定相同。由同一点出发的有向边都是该策略的最优反应，则获得的收益相同，否则只有其中收益最大的才是最优反应。

例如上一个博弈矩阵用收益流向图表示如下。

1

石头

5

剪刀 布

4

由收益流向图我们可以得到下列结论：

1. 收益流向图的类型有种。
2. 当收益流向图中出现了自环或是由两个点和两条边构成的环，则该博弈存在纯策略的纳什均衡。
3. 当存在三个点和三条边构成的环时，该博弈可以利用形如式的方法来求混合策略纳什均衡。



上述三个结论的简单证明如下：

1. 每个点的出度至少为1，则从一个点的出发的情况有种，故收益流向图的类型有种。
2. 如果出现自环，则双方都选择这个策略就是符合纳什均衡；如果出现由两个点和两条边构成的环，则说明两个策略互为最优反应，故符合纳什均衡。
3. 如果出现三个点和三条边构成的环，当不考虑自环和其他反向的边时，是符合式的使用情景的，因而可据此求出混合策略纳什均衡。



由结论（3）得到的混合策略与纯策略可以构成新的混合策略，纯策略间也可以构成混合策略，因而对于该类博弈可求其所有纳什均衡。

## （3）模拟寻找纳什均衡

纳什均衡就是一个不断改变策略以寻求最优反应的过程，因而我们可以考虑用模拟的方法寻找最优反应，并求得纳什均衡。

在一个游戏中，有2个参与者参与人1、参与人2，两人分别有s和t个策略，两人的收益矩阵参与人2知，利用模拟分析两人博弈情况的方法如下：

（1）设参与人1的s个决策分别有个有效单位，我们把参与人1的p个决策的所有有效单位称为参与人1的决策池，记为A。同理，参与人2的t个决策分别有个有效单位，形成参与人2的决策池B。

（2）参与人1、参与人2每次博弈均随机的从各自的决策池中选择一个有效单位，该有效单位对应的策略就作为自己本次博弈的决策。

（3）参与人1、参与人2两人每次博弈结束后，会依据本次博弈的收益调整自己的决策池，若收益为+a，则在决策池中添加a个本次博弈所使用决策对应的有效单位，若收益为-a，则在决策池中剔除a个本次博弈所使用决策对应的有效单位。

（4）多次重复上述步骤，观察参与人1与参与人2的策略池内各策略对应有效单位占比的变化。

例如：参与人1、参与人2进行石头剪刀布，那么参与人1有三个决策：石头、剪刀和布，设三个决策依次有1,2,3个有效单位，那么参与人1的决策池中就有6个有效单位，其中1个代表石头，2个代表剪刀，3个代表布。参与人1会从自己的决策池中随机挑选一个有效单位，其中石头剪刀布依次被选中的概率是1/6，2/6，3/6。若规定胜者记+1分，负者记-1分，如果某次博弈中参与人1使用决策“石头”，参与人2使用决策“剪刀”，则参与人1的决策池中增加一个石头对应的有效单位，参与人2的决策池中减少一个剪刀对应的有效单位。

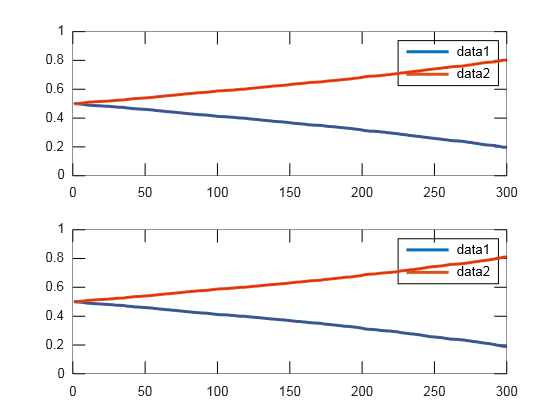
接下来举几个利用模拟来寻找纳什均衡的例子，数据图中上图是参与人1，下图是参与人2，其中每200次记录一次参与人1与参与人2决策池中各决策对应有效单位的占比。

1. 囚徒困境

设参与人1与参与人2决策池的初始状态，都是cooperate与defect对应的有效单位是10000，进行60000次博弈。

参与者2

|  |  |  |  |
| --- | --- | --- | --- |
| 参  与  者  1 |  | Data1（cooperate） | Data2（defect） |
| Data1（cooperate） | 4 , 4 | -5 , 5 |
| Data2（defect） | 5 , -5 | -3 , -3 |

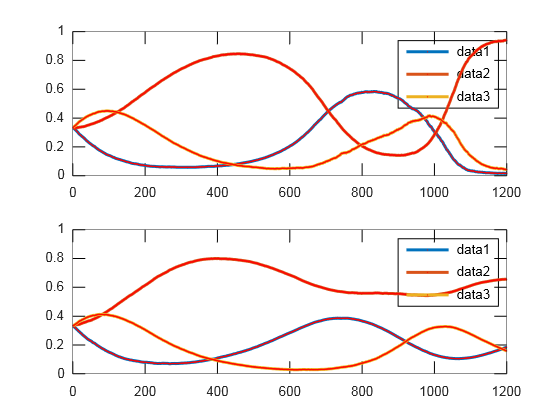


（2）剪刀石头布（零和博弈变体）

设参与人1与参与人2决策池的初始状态，都是石头、剪刀与布对应的有效单位是10000，进行240000次博弈。

参与者2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参  与  者  1 |  | Data1（石头） | Data2（剪刀） | Data3（布） |
| Data1（石头） | 0 , 0 | 1 , -1 | -5 , 5 |
| Data2（剪刀） | -2 , 2 | 0 , 0 | 2 , -2 |
| Data3（布） | 5 , -5 | -2 , 2 | 0 , 0 |

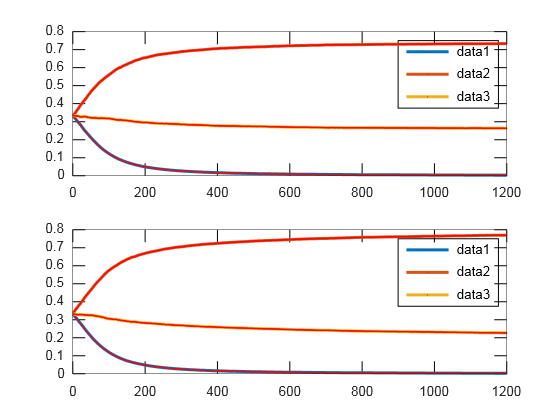


1. 剪刀石头布（非零和博弈变体）

设参与人1与参与人2决策池的初始状态，都是石头、剪刀与布对应的有效单位是10000，进行240000次博弈。

参与者2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参  与  者  1 |  | Data1（石头） | Data2（剪刀） | Data3（布） |
| Data1（石头） | 1 , 1 | -1 , -2 | -2 , -1 |
| Data2（剪刀） | -2 , -1 | 2, 2 | 5,4 |
| Data3（布） | -1 , -2 | 4 , 5 | -2 , -2 |



综合上述例子可以看出，当博弈存在纯策略纳什均衡时，模拟结果收敛，当不存在纯策略纳什均衡时，模拟结果呈现周期性波动。

下面来考察模拟结果呈现周期性波动的情况。对于石头剪刀布零和博弈变体，我们更改得分规则如下表，获胜得分依次表示使用石头、剪刀和布获胜的得分，以及分别使用上述三个策略失败的减分。表中数据是整个模拟过程中，所有记录点三个策略比例的平均值。可以发现这个平均值与混合策略的纳什均衡十分接近，因而波动情况实则反应了博弈的混合策略纳什平衡。

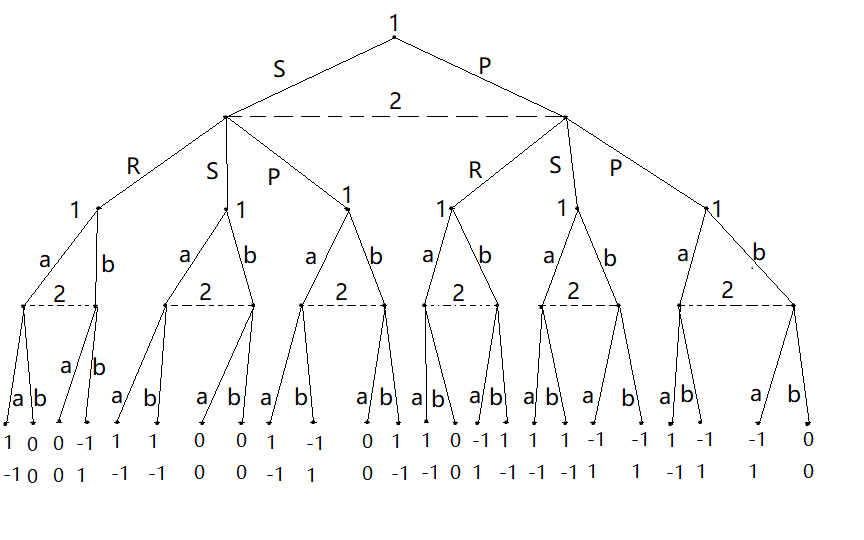
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 获胜得分  种群名称 | 1 , 2 , 3 | | 1 , 3 , 4 | | 1 , 1 , 2 | |
| 石头 | 0.33355 | 0.33416 | 0.36951 | 0.37256 | 0.25555 | 0.25327 |
| 剪刀 | 0.49905 | 0.49956 | 0.50129 | 0.49720 | 0.48642 | 0.48717 |
| 布 | 0.16739 | 0.16628 | 0.12920 | 0.13025 | 0.25803 | 0.25957 |

事实上，通过模拟的方法求得的纳什平衡事实上只是博弈中纳什平衡的一个，其更为接近生物学中进化稳定战略（Evolutionary Stable Strategy,ESS），话句话说通过模拟的方法只能求得收益高（在生物学上具有更快繁殖速率），对称性好（在生物学上稳定）的纳什均衡。简单来说，模拟求得的ESS是纳什均衡，但纳什均衡并不一定都是ESS。

## （二）完全信息动态博弈

在生活中有一个很常见的剪刀石头布的变体游戏，符合不完全信息动态博弈的特征，具体规则如下：参与人1与参与人2首先同时选择并展示剪刀、石头和布三个策略之一，保持这一状态的情况下，再一次同时选择并展示三个策略之一，随后参与人1与参与人2可以选择收回刚刚自己的两个策略之一，以两人剩下的两个策略来判断两人的胜负。这一规则可以简单地概括为“左一拳，右一拳，情况不妙收一拳”，也就是展示两个选择之后，选择一个作为自己的最终策略。

由上可知，参与人1与参与人2的战略集均有3\*3\*2个元素，则共有18\*18种战略组合。对于初次策略选择，事实上可以归结为两种情况，两人的策略相同（概率为2/3）或是两人策略不同（概率为1/3）。我们取两人策略不同情况的一部分来制作其博弈树。不妨设参与人1第一次选择了石头，参与人2第二次选择剪刀，开始节点是参与人1选择，同时我们忽略较为简单的参与人1在第二次选择相同策略——石头，博弈树如下所示，其中收益向量上方是参与人1的收益，下方是参与人2的收益。（R表示石头，S表示剪刀，P表示布，a表示选择第一次的策略作为最终的策略，b表示选择第二次的策略作为最终的策略，获胜得+1分，失败则得-1分）



根据子博弈完美均衡(subgame perfect equilibrium)满足每一个子博弈都是纳什平衡，我们以上图所示博弈树的最左侧的一个子博弈，即参与人1已经展示了石头和剪刀，参与人2已经展示了剪刀和石头，两人即将做出最终决策的子博弈。这个博弈的博弈矩阵如下。

参与者2

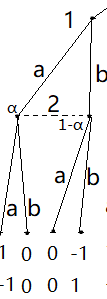
|  |  |  |  |
| --- | --- | --- | --- |
| 参  与  者  1 |  | a | b |
| a | 1 , -1 | 0 , 0 |
| b | 0 , 0 | -1 , 1 |

我们可以得到这个子博弈的纳什均衡是(a,b)，若纳什均衡不存在，则不存在子博弈完美均衡。通过相同的方法我们可以求得其他子博弈的纳什均衡(RPa,SSa)和(RPb,SSb)，对应的收益都是(1,-1)，进而简化博弈树，最终可以得到上图博弈树的子博弈完美均衡是(RSa,SRb)。

如果参与人1第二次的策略与第一次相同，同样选择石头，那么此时的完美子博弈均衡也可以有更小的子博弈的纳什均衡考虑，包括：(RRa,SRb),(RRb,SRb)收益为(0,0)；(RRa,SPb),(RRb,SPb)收益为(-1,1)。所以此时的完美子博弈平衡是(RRa,SRb)和(RRb,SRb)。再与上述博弈树结合考虑，对于参与人1和参与人2首先选择策略石头和剪刀的情况，整体的子博弈完美均衡为(RSa,SRb),(RRa,SRb)和(RRb,SRb)。

对于初次策略选择，若参与人1与参与人2策略相同，不妨设两人都选择石头，可以利用类似的方法得到此时整个博弈的子博弈完美均衡，进而依据整个博弈的对称性可求所有子博弈完美均衡。

接下来我们选取一个子博弈来考察其完美贝叶斯均衡(perfect Bayesian equilibrium)。由于贝叶斯均衡都是完美子博弈均衡，所以我们只需要从上述求得的完美子博弈平衡中考察即可。我们取上图所示博弈树的最左侧的一个子博弈，即参与人1已经展示了石头和剪刀，参与人2已经展示了剪刀和石头，两人即将做出最终决策的子博弈。其中我们假设参与人1选择a的信念(belief)为α，0≤α≤1。



若参与人2选择a的收益小于等于选择b的收益，可列式：解得：，也就是说，参与人2选择b的收益始终大于选择a的收益。

现在假设参与人2选择b，那么参与人1必定选择a也即，满足，所以说参与人2最好确信参与人1会选择a，也就是存在一个完美贝叶斯均衡(RSa,SRb)其中α=1。通过类似的方法可以得到上述整个博弈树的完美子博弈均衡，同时也是确定的完美贝叶斯均衡。