



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

[Home Page](#)

[Title Page](#)



Page 1 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



中国科学院研究生院

# 运筹通论II

刘克

中科院数学与系统科学研究院 北京100190

邮箱地址: kliu@amss.ac.cn



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

[Home Page](#)

[Title Page](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

Page 2 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

## 第三部分 HC与摄动MDP

[Home Page](#)

[Title Page](#)



Page 3 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

- 1 哈密尔顿圈问题
- 2 有向图到MDP的嵌入
- 3 平稳策略的分类
- 4 约束折扣MDP与HC
- 5 HCP嵌入MDP的摄动
- 6 更为精细的分析
- 7 开问题和有关猜想



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 4 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

# 1 哈密尔顿圈问题

在离散数学和组合最优化领域中遇到的很多困难问题都是因为这些  
问题所考虑的范围是离散的。我们这里主要讲述了如何将这种问题  
嵌入到一类凸域上考虑，而处理凸域上的问题时，人们有着方便的手  
段。域的凸化技术是通过把原问题的关键元素赋予概率解释而实现  
的，即使原来的问题是一个确定性的问题，也可以这样做。

Home Page

Title Page

◀ ▶

◀ ▶

Page 5 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

尽管人们很早就或多或少的想到这样做，例如，Bertsekas利用了随机优化的分析技术研究了最短路问题。但我们这里介绍的方法是Filar和Krass 在考虑哈密尔顿圈和旅行售货商这两个问题时开创的。具体来说，就是把这两个问题嵌入到单摄动的马尔可夫决策过程中去考虑。每一个确定性平稳策略对应原问题的一个子图（如果存在哈密尔顿圈，必然有一个确定性平稳策略与之对应）。事实上，Filar和Krass所考虑的优化范围是在随机平稳策略类上，它是由一个凸多面体构成，其每一个顶点都对应一个确定性平稳策略。因此，这里蕴涵着丰富的研究内容。

Home Page

Title Page

« »

◀ ▶

Page 6 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

后来，Chen和Filar，Feinberg分别继续了Filar和Krass的研究工作，Liu，Filar和Liu又得到了进一步的结果，随后，Andramonov等人、Filar和Lasserre、Ejov等以及Borkar等都继续这方面的工作。我们主要针对哈密尔顿圈问题展开讨论，原则上讲这里的技术也适用于很多其它离散数学中的问题。

[Home Page](#)

[Title Page](#)

◀◀

▶▶

◀

▶

Page 7 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

威廉·哈密尔顿爵士1859年在探求一个含有十二面体所有顶点的圈的结构时提出了一族以他的名字命名的图。若一个图 $G$ 有一个生成圈 $Z$ ，则称图 $G$ 为一个哈密尔顿图，而 $Z$ 是一个哈密尔顿圈。给定一个图，确定它是否是哈密尔顿图的问题与找到这个图的一个哈密尔顿圈问题具有相同的计算复杂度，都属于NPC难问题，他们同属于21世纪数学领域列举的七大难题之一：NPC难问题是否等价于P问题。

**哈密尔顿圈问题**（Hamiltonian Cycle Problem 简记为HCP）可以简单的用一句话描述：给定一个图，找到一个包含这个图所有顶点的简单的圈（哈密尔顿圈，简记为HC）或者证明HC不存在。HCP是著名的NPC难问题，它的诱人之处除了问题本身以外，还因为它与另一个著名问题，**旅行售货商问题**（Travelling Salesman Problem，简记为TSP）有着密切的联系。后者的有效算法在运筹学、优化以及计算机科学中有着巨大的影响。从数学的角度来看，TSP的主要困难就是HCP。

[Home Page](#)[Title Page](#)[◀](#) [▶](#)[◀](#) [▶](#)[Page 8 of 64](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)





哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

« »

◀ ▶

Page 9 of 64

Go Back

Full Screen

Close

Quit

一个有向图 $G$ 被记为 $G = (V(G), E(G))$ , 其中:  $V(G)$ 表示包含图 $G$ 的所有顶点的集合;  $E(G)$ 表示包含图 $G$ 的所有边的集合。当图 $G$ 有 $N$ 个顶点时, 我们记 $V(G) = \{1, 2, \dots, N\}$ , 而边的集合 $E(G)$ 中的每一条有向边都可以用 $V(G)$ 中的一个有序的对子 $(i_1, i_2)$ 表示, 其中 $i_1$ 表示这条有向边的起点,  $i_2$ 表示这条有向边的终点,  $i_1, i_2 \in V(G)$ 。

如果 $E(G)$ 中一个子集合 $\{(i_0, i_1), (i_1, i_2), \dots, (i_{N-1}, i_N)\}$ 满足条件 $\{i_0, i_1, \dots, i_{N-1}\} = V(G)$ 而且 $i_N = i_0$ , 那么就称 $\{(i_0, i_1), (i_1, i_2), \dots, (i_{N-1}, i_N)\}$ 是一个从 $i_0$ 出发的哈密尔顿圈(HC)。



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

有向图 $G$ 的一个子图 $H$ 是指图 $H$ 的顶点集合与图 $G$ 的顶点集合一致，而图 $H$ 的边集合 $E(H)$ 是图 $G$ 的边集合 $E(G)$ 的一个子集，即 $E(H) \subset E(G)$ 。也可以粗略的说：图 $G$ 的一个子图就是由图 $G$ 去掉一些边构成的一个图。如果图 $G$ 是哈密尔顿图，那么它就存在一个哈密尔顿圈为它的子图；反之亦然。

利用组合的方法研究哈密尔顿图问题或者哈密尔顿圈问题有着悠久的历史，也有着丰富的研究成果，很多学者做出了贡献，例如迪拉克、奥尔、波塞或者哈拉里。

Home Page

Title Page

« »

◀ ▶

Page 10 of 64

Go Back

Full Screen

Close

Quit

## 2 有向图到MDP的嵌入

我们把有向图 $G$ 的每一个顶点 $i \in V(G) = \{1, 2, \dots, N\}$ 看成MDP问题的一个状态，那么我们可以构造一个状态空间 $S = \{1, 2, \dots, N\}$ ，其中每一个状态对应有向图 $G$ 的相应顶点，状态空间的状态个数恰为有向图 $G$ 的顶点个数。有向图 $G$ 的每个顶点 $i$ 向外连接的边就定义为MDP问题中相应状态 $i$ 的行动集合 $A(i)$ 中的一个行动，即： $A(i) = \{(i, j) | (i, j) \in E(G)\}$ ；而在选定一个行动之后的一步状态转移概率是退化的分布，它反映了从状态 $i$ 出发后用行动 $j$ 必然在下一步到达状态 $j$ ；即

$$p(j|i, a) = \delta_{aj} = \begin{cases} 1, & a = (i, j) \text{ 且 } (i, j) \in E(G) \\ 0, & \text{否则.} \end{cases}$$

为了方便且在不引起混淆的情况下，我们有时将状态 $i$ 的行动 $a = (i, j)$ （有向边）记为这条有向边的头 $j$ ，即 $a = j$ 。

通过这样的转换，每一个确定性平稳策略 $f = (i_1, i_2, \dots, i_N)$ （或者 $f^\infty$ ）对应着一个 $G$ 的子图，而且这个子图只有 $N$ 条边。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HCP  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 11 of 64

Go Back

Full Screen

Close

Quit

**例2.1:**  $G$ 是一个有5个顶点的完全图, 其中顶点集合 $V(G) = \{1, 2, 3, 4, 5\}$ , 边集合 $E(G) = \{(i, j), i, j = 1, 2, 3, 4, 5, \text{ 而且 } i \neq j\}$ 。不难看出, 图 $G$ 共有1024个子图。

平稳策略 $f = (2, 1, 1, 2, 4)$ 对应的子图

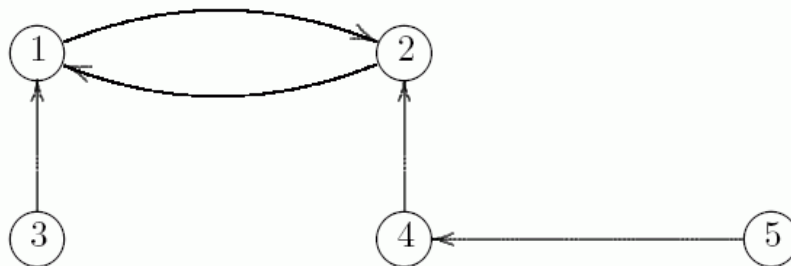


图  $G$  的子图  $f$



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀

▶

◀

▶

Page 12 of 64

Go Back

Full Screen

Close

Quit



它诱导的转移概率矩阵 $P(f)$ 为:

$$P(f) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

[Home Page](#)

[Title Page](#)

[◀◀](#)

[▶▶](#)

[◀](#)

[▶](#)

Page 13 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



它平稳策略  $g = (2, 3, 4, 5, 1)$  对应着一个HC， 它的转移概率矩阵为：

$$P(g) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Page 14 of 64

Go Back

Full Screen

Close

Quit



从上面的例子可以看出，转换后的MDP在一个平稳策略 $f$ 下诱导的马尔可夫链从每一个状态出发的轨迹都是惟一确定的。

对于一个随机平稳策略 $\pi_0$ （或者 $\pi_0^\infty$ ）就是状态空间 $S$ 上的一个 $N$ 维概率分布向量，其中 $\pi_0(\cdot|i) = (\pi_0(1|i), \pi_0(2|i), \dots, \pi_0(N|i))$ ，这里 $\pi_0(k|i)$ 表示在顶点 $i$ 处选择行动 $k$ （即有向边 $(i, k)$ 或者指向顶点 $k$ ）的概率。很明显有： $\sum_{k=1}^N \pi_0(k|i) = 1$ ，而且当 $(i, k) \notin A(i)$ 时， $\pi_0(k|i) = 0$ 。它也诱导出一个转移概率矩阵

$$\mathbf{P}(\pi_0) = [p(j|i, \pi_0)], \quad i, j = 1, 2, \dots, N,$$

其中对所有的 $i, j \in S$

$$p(j|i, \pi_0) = \sum_{k=1}^N p(j|i, k) \pi_0(k|i).$$

很明显，如果边 $(i, k) \notin E(G)$ ，我们总认为有 $p(j|i, k) = 0$ 。



**定义2.1:** . 随机平稳策略 $\pi_0 \in \Pi_s$ 被称为**双随机平稳策略**, 如果由它诱导出的转移概率矩阵 $P(\pi_0)$ 是双随机的。更精确的说, 就是概率矩阵 $P(\pi_0)$ 的行和与列和都是1。双随机平稳策略的全体我们记为 $\Pi_{ds}$ 。

很明显, 在例2.1中, 平稳策略 $f$ 就不是双随机策略, 而对应于HC的平稳策略 $g$ 则是双随机策略。容易得出:  $\Pi_{ds} \subseteq \Pi_s$ ; 而且对任意的 $\pi_0 \in \Pi_{ds}$ , 由 $\pi_0$ 诱导的马尔可夫链要么对应于一个HC, 要么对应于图 $G$ 中几个不交子圈的并。





哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

不失一般性，我们总可以考虑初始状态总是从1出发的情形。此时，我们总将状态1称为家顶点。我们将把MDP中的（确定性）平稳策略 $f$ 称为HC或者图 $G$ 中的哈密尔顿圈，如果由这个平稳策略导出的子图的边集合 $\{(1, f(1)), (2, f(2)), \dots, (N, f(N))\}$ 是图 $G$ 的一个HC。

如果平稳策略 $f$ 构成的子图包含一个圈长小于 $N$ 的圈，记为 $m$ ，我们就称 $f$ 有一个 $m$ -子圈。

这样直接将图 $G$ 对应为MDP所带来的最大困难就是由不同的平稳策略诱导出来的马尔可夫链会具有多链的结构（具有多个遍历类）。

Home Page

Title Page

◀

▶

◀

▶

Page 17 of 64

Go Back

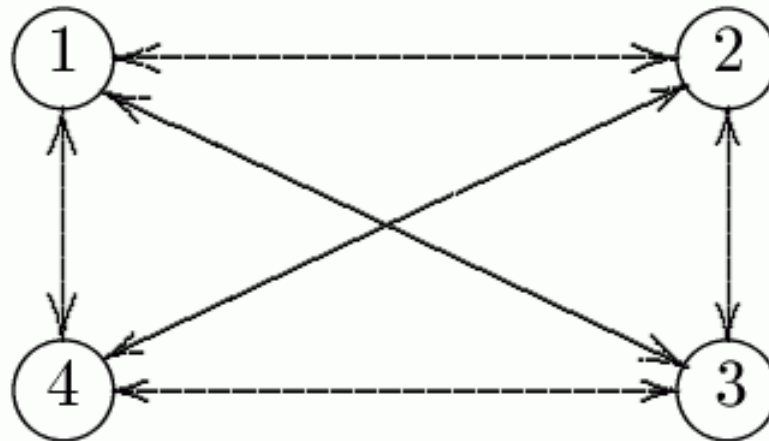
Full Screen

Close

Quit



**例2.2:**  $G_4$ 是有4个顶点的完全图，其中顶点集合 $V(G) = \{1, 2, 3, 4\}$ ，边集合 $E(G) = \{(i, j), i, j = 1, 2, 3, 4, \text{ 而且 } i \neq j\}$ 。



哈密顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 18 of 64

Go Back

Full Screen

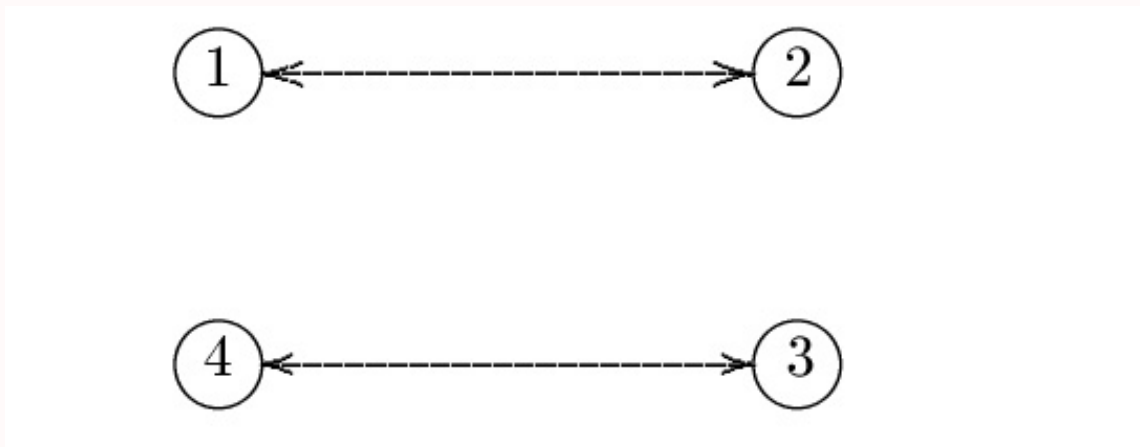
Close

Quit

在上面的图中，平稳策略 $f_1$ 满足 $f_1(1) = 2$ ,  $f_1(2) = 3$ ,  $f_1(3) = 4$ 以及 $f_1(4) = 1$ 诱导出一个 $G$ 的子图 $G_{f_1} = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$ 恰为一个HC，所对应的转移概率矩阵是：

$$\mathbf{P}(f_1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

只有包含4个顶点的一个遍历类。而平稳策略 $f_2$ 满足 $f_2(1) = 2$ ,  $f_2(2) = 1$ ,  $f_2(3) = 4$ 以及 $f_2(4) = 3$ 诱导出一个 $G$ 的子图 $G_{f_2} = \{(1, 2), (2, 1), (3, 4), (4, 3)\}$ 包含了2个圈长为2的子圈，



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 19 of 64

Go Back

Full Screen

Close

Quit



它对应的转移概率矩阵是:

$$\mathbf{P}(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

尽管 $f_1$ 和 $f_2$ 都是双随机的情形, 但是后者有两个遍历类。

当然随机平稳策略可以被看成为平稳策略的凸组合。例如随机平稳策略 $\pi_0$ 由4个概率向量组成:  $\pi_0(1) = (\pi_0(2|1), \pi_0(3|1), \pi_0(4|1)) = (1, 0, 0)$ ,  $\pi_0(2) = (\pi_0(1|2), \pi_0(3|2), \pi_0(4|2)) = (0.7, 0.3, 0)$ ,  $\pi_0(3) = (\pi_0(1|3), \pi_0(2|3), \pi_0(4|3)) = (0, 0, 1)$ 和 $\pi_0(4) = (\pi_0(1|4), \pi_0(2|4), \pi_0(3|4)) = (0.3, 0.7, 0)$ 。显然, 有 $\pi_0 = 0.3f_1 + 0.7f_2$ , 诱导的转移概率矩阵是:

$$\mathbf{P}(\pi_0) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.7 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 1 \\ 0.3 & 0 & 0.7 & 0 \end{bmatrix}.$$

### 3 平稳策略的分类

将一个图  $G = \{V(G), E(G)\}$  嵌入到MDP之后，这个MDP的平稳策略类  $F$  可以根据每个策略（即子图）对应于图  $G$  的圈的结构进行分类。具体来说，一个平稳策略  $f \in F$ ，我们定义子图  $G_f = \{V(G_f), E(G_f)\}$ ，其中顶点集合为  $V(G_f) = V(G)$  而边集合为

$$(i, j) \in E(G_f) \Leftrightarrow f(i) = j.$$

另外，我们记从状态1出发且长为  $m$  的圈为：

$$c_m^1 = \{(i_1 = 1, i_2), (i_2, i_3), \dots, (i_m, i_{m+1} = 1)\}, \quad m = 2, 3, \dots, N$$

显然， $c_N^1$  就是一个HC。如果子图  $G_f$  包含一个圈  $c_m^1$ ，我们记为  $G_f \supset c_m^1$ 。定义：

$$C_m^1 = \{f \in F | G_f \supset c_m^1\},$$

它是从状态1出发含有一个长为  $m$  的圈的平稳策略全体，其中  $m = 2, 3, \dots, N$ 。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

« »

◀ ▶

Page 21 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

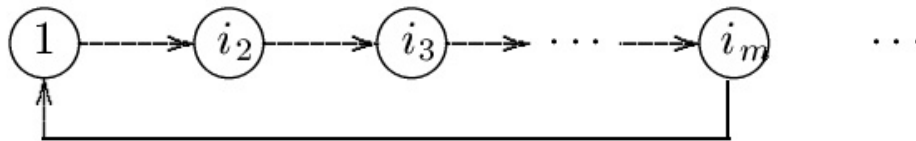
约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

当然， $C_N^1$ 是包含HC的全体平稳策略集合（有可能是空集）。对一般的 $m = 2, 3, \dots, N - 1$ ，根据原图 $G$ 的构造和状态1的选择， $C_m^1$ 也可能是空的。一个典型的平稳策略 $f \in C_m^1$ 所生成 $G$ 中的子图 $G_f$ 具有下面图的形式，其中 $\dots$ 表示其它的那些顶点或者状态。



Home Page

Title Page

◀ ▶

◀ ▶

Page 22 of 64

Go Back

Full Screen

Close

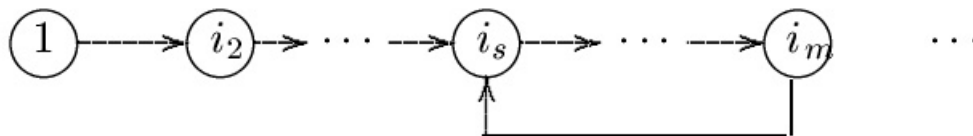
Quit



因此，我们可以把平稳策略类 $F$ 划分为：

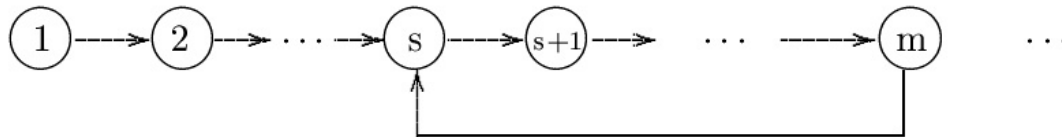
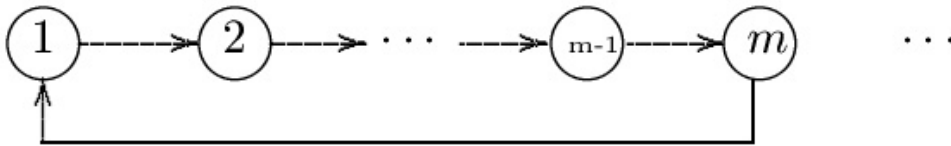
$$F = \left[ \bigcup_{m=2}^N C_m^1 \right] \cup \mathcal{N}_c,$$

其中 $\mathcal{N}_c$ 表示不包含在任何一个 $C_m^1$ 的那些平稳策略。 $\mathcal{N}_c$ 中的典型平稳策略 $f$ 所生成 $G$ 中的子图 $G_f$ 具有下面的形式。





在固定初始状态1之后，对每一个平稳策略 $f$ 我们将考虑它构成的链的结构，而它的链结构在对图 $G$ 中除了1顶点以外的顶点做置换下不变。因此，不失一般性的可以只考虑两种类型的平稳策略 $f_m \in C_m^1$ 和 $f_m^s \in \mathcal{N}_c$ 就可以了，其中 $f_m$ 和 $f_m^s$ 所生成 $G$ 中的子图分别具有下面的形式。







哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

## 4 约束折扣MDP与HC

回想马氏决策关于折扣模型的定义：对任意的 $\beta \in (0, 1)$ ，策略 $\pi \in \Pi$ 和状态 $i \in S$ ，从状态 $i$ 出发的折扣总报酬是

$$V_{\beta}(i, \pi) = \sum_{t=0}^{\infty} \beta^t E_{\pi}^i[r(Y_t, \Delta_t)] \equiv V_{\beta}(i, \pi, r). \quad (1)$$

最后的 $V_{\beta}(i, \pi, r)$ 特别表示总折扣报酬是由报酬函数 $r$ 导致的。

Home Page

Title Page

◀ ▶

◀ ▶

Page 25 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

在2节中我们根据一个给定的有向图 $G$ 已经定义了一个与之相应的MDP，包括状态空间、行动空间和转移概率矩阵等元素。下面我们就HC问题定义报酬函数 $r$ ，它的具体形式就是：

$$r(i, a) = \begin{cases} 1, & \text{如果 } i = 1; \\ 0, & \text{其他。} \end{cases} \quad (2)$$

利用MDP的方法研究HC问题，巧妙的定义报酬是一个很有效的手段。

定义了报酬之后,我们就把在有向图 $G$ 中判定HC存在或者寻找HC问题转换到MDP中寻找满足一定条件的策略问题了。我们首先考虑约束折扣问题，在HC情况下的形式。由于初始的出发状态总是1，我们有下面的结论。

[Home Page](#)

[Title Page](#)

[«](#) [»](#)

[◀](#) [▶](#)

Page 26 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



**定理4.1:** 对任意的 $m = 2, 3, \dots, N$ , 下面的两个条件是等价的。

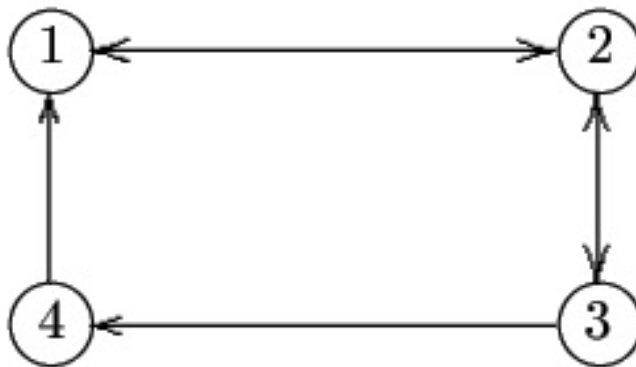
$$1) \quad V_{\beta}(1, \pi, r) = (1 - \beta^m)^{-1}; \quad (3)$$

$$2) \quad P_{\pi}\{Y_t = 1 | Y_0 = 1\} = \begin{cases} 1, & \text{如果 } t \in \{0, m, 2m, \dots\}; \\ 0, & \text{否则.} \end{cases} \quad (4)$$

一个策略 $\pi$ 被称为是哈密尔顿的, 就是对所有的 $t = 0, 1, 2, \dots$ , 用策略 $\pi$ 在时刻 $tN, tN + 1, \dots, (t + 1)N$ 的状态 $i_{tN}, i_{tN+1}, \dots, i_{(t+1)N}$ 是一个HC (关于概率测度 $P_{\pi}\{\dots | Y_0 = 1\}$ 几乎必然成立)。因此, 一个平稳哈密尔顿策略就对应于有向图的一个HC。在给定一个哈密尔顿策略后, 很容易由它可以构造出一个HC。如果策略 $\pi$ 是哈密尔顿的, 那么很明显定理4.1中2) 的(4)式对 $m = N$ 时成立, 因此有(3)式也成立。但反之并不成立。



**例4.1:** 有向图 $G$ 有4个顶点, 其中顶点集合 $V(G) = \{1, 2, 3, 4\}$ , 边集合 $E(G) = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 1)\}$ 。图 $G$ 共有4个平稳策略。



$$f_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 2 \\ 3 \\ 2 \\ 1 \end{bmatrix}, \quad f_3 = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 1 \end{bmatrix}, \quad f_4 = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix}.$$

哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 28 of 64

Go Back

Full Screen

Close

Quit



考虑一个策略  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$  定义为:

$$\pi_{0+4k} = f_2, \pi_{1+4k} = f_2, \pi_{2+4k} = f_2, \pi_{3+4k} = f_4, \quad k = 0, 1, \dots$$

也就是  $\pi = (f_2, f_2, f_2, f_4, f_2, f_2, f_2, f_4, \dots)$  是一个以4为周期的策略, 它在  $m = 4 = N$  时满足条件(4)式和条件(3)式, 但这个策略不是哈密尔顿策略。

**定理4.2:** 如果随机平稳策略  $\pi = \pi_0^\infty \in \Pi_s$ , 而且当  $m = N$  时满足条件(4), 那么随机平稳策略  $\pi$  是确定性的, 即存在  $f \in F$  使得  $\pi_0 = f$ 。如果  $1 < m < N$  时条件(4)成立, 那么随机平稳策略诱导的马氏链从状态1出发只停留在一个以  $m$  为周期的遍历类中, 特别的循环定义

$$\begin{aligned} S_1 &= \{1\}; \\ S_2 &= \{i | \pi_0(i|1) > 0, i \in S, 1 \in S_1\}; \\ S_3 &= \{i | \pi_0(i|j) > 0, i \in S, j \in S_2\}; \\ &\vdots \\ S_m &= \{i | \pi_0(i|j) > 0, i \in S, j \in S_{m-1}\}; \\ S_{m+1} &= S_1, \end{aligned}$$

其中  $S_{k+1}$  是从状态1出发第  $k + 1$  步进入的可能状态集合 ( $l = 0, 1, \dots$ ), 而且满足  $S_{k_1} \cap S_{k_2} = \emptyset, k_1 \neq k_2$ 。



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

**定理4.3:** 下面的结论相互等价:

- 1) 随机平稳策略  $\pi = \pi_0^\infty \in \Pi_s^d$  是哈密尔顿的。
- 2) 随机平稳策略  $\pi = \pi_0^\infty \in \Pi_s$  是哈密尔顿的。
- 3) 随机平稳策略  $\pi = \pi_0^\infty \in \Pi_s^d$ , 且存在  $\beta \in (0, 1)$  满足  $V_\beta(1, \pi, r) = (1 - \beta^N)^{-1}$ 。
- 4) 随机平稳策略  $\pi = \pi_0^\infty \in \Pi_s$ , 且存在  $2N - 1$  个不同的  $\beta_k \in (0, 1)$  满足  $V_{\beta_k}(1, \pi, r) = (1 - \beta_k^N)^{-1}$ ,  $k = 1, 2, \dots, 2N - 1$ 。

Home Page

Title Page

◀ ▶

◀ ▶

Page 30 of 64

Go Back

Full Screen

Close

Quit



由定理4.3知道，HCP等价于下面相应的MDP的两个公开问题：

**问题4.1：** 对某 $\beta \in (0, 1)$ 寻找平稳策略 $f \in F$ 满足：

$$V_{\beta}(1, f, r) = \frac{1}{1 - \beta^N},$$

或者确定其不存在。

**问题4.2：** 对某些不同的 $\{\beta_k \in (0, 1) | k = 1, 2, \dots, 2N - 1\}$ 寻找随机平稳策略 $\pi = \pi_0^{\infty} \in \Pi_s$ 满足：

$$V_{\beta_k}(1, \pi, r) = \frac{1}{1 - \beta_k^N}, \quad k = 1, 2, \dots, 2N - 1,$$

或者确定其不存在。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

## 5 HCP嵌入MDP的摄动

前面我们用折扣的MDP讨论了HC问题。平均准则的MDP也是研究HC问题的有效工具，下面我们利用平均准则来讨论HC问题的。

将有向图 $G$ 嵌入到MDP中之后，任意一个平稳策略 $f$ （子图）对应的转移概率矩阵 $P(f)$ 所诱导的马氏链很可能具有多链的结构，而在平均准则的讨论中，马氏链的状态分类强烈的影响着分析的手段和方法。处理平均准则问题时的很多分析都是基于MDP策略的链结构是单链的（指任意平稳策略诱导的马氏链只有一个遍历类加上一些可能的瞬时状态），因为在这样的条件下，会减少分析上的巨大困难。完全遍历的MDP是一个单链的MDP，而且在完全遍历的MDP中任何一个平稳策略都只有一个遍历类，并且瞬时状态的集合是空集。

[Home Page](#)[Title Page](#)[«](#) [»](#)[◀](#) [▶](#)[Page 32 of 64](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



为了克服一般子图对应的平稳策略 $f$ 诱导的马氏链具有多链的状态结构，可以利用对转移概率摄动的手段解决这个问题。我们取一个参数 $\epsilon \in (0, 1)$ 之后，将需要讨论的MDP的转移概率做 $\epsilon$ 的摄动。因此，原来MDP的四重组

$$\{S, A(i), p(\cdot|i, a), r(i, a)\}$$

被摄动成为

$$\{S, A(i), p_\epsilon(\cdot|i, a), r(i, a)\},$$

其中 $S$ ,  $A(i)$ 和 $r(i, a)$ 都不做改变，只有转移概率 $p(j|i, a)$ 改成为：

$$p_\epsilon(j|i, a) \triangleq \begin{cases} 1 & \text{如果 } i = 1 \text{ 且 } a = j, \\ 0 & \text{如果 } i = 1 \text{ 且 } a \neq j, \\ 1 & \text{如果 } i > 1 \text{ 且 } a = j = 1, \\ \epsilon & \text{如果 } i > 1, a \neq j \text{ 且 } j = 1, \\ 1 - \epsilon & \text{如果 } i > 1, a = j \text{ 且 } j > 1, \\ 0 & \text{如果 } i > 1, a \neq j \text{ 且 } j > 1. \end{cases}$$

其中 $\epsilon \in (0, 1)$ 。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page



Page 33 of 64

Go Back

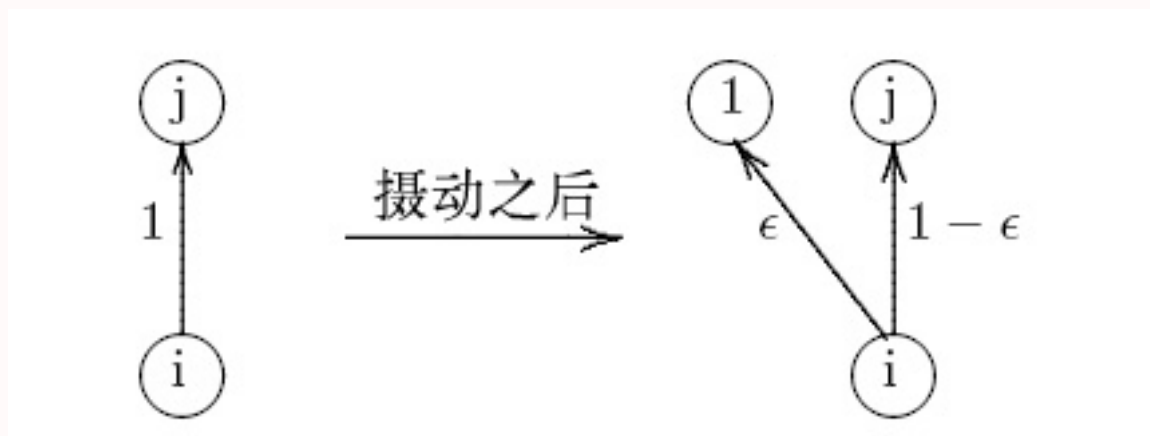
Full Screen

Close

Quit

注意，这样的摄动保证了每一个平稳策略  $f \in F$  所诱导的马氏链都是完全遍历的，它的一步转移概率矩阵我们记做： $P_\epsilon(f)$ 。摄动的目的一个是保证不会具有多链的结构；另一个就是使得家顶点与其他顶点有所区别，并且确定性策略诱导的马氏链的转移概率矩阵保持稀疏性质。

注意顶点1就是家顶点。每一对可行的顶点  $i, j$  (都不等于1) 都对应于一条确定的有向边  $(i, j)$ ，摄动就是将这条确定的有向边  $(i, j)$  用两条随机的有向边  $(i, 1)$  和  $(i, j)$  替代，选择这两条边的概率（也可看成是赋予权重）分别为  $\epsilon$  和  $1 - \epsilon$ 。这个随机摄动还可以解释为：以概率  $1 - \epsilon$  选择有向边  $(i, j)$ ，即从状态  $i$  移动到状态  $j$ ，同时以概率  $\epsilon$  折回家顶点，即状态1。



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 34 of 64

Go Back

Full Screen

Close

Quit

我们以前面的平稳策略 $f_2$ 的转移概率矩阵为例，

$$\mathbf{P}(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

它有两个正常返类：状态 $\{1, 2\}$ ；和状态 $\{3, 4\}$ 。做了非对称线性摄动之后的一步转移概率矩阵就是：

$$\mathbf{P}_\epsilon(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \epsilon & 0 & 0 & 1 - \epsilon \\ \epsilon & 0 & 1 - \epsilon & 0 \end{bmatrix}.$$

很明显，经过摄动之后， $\mathbf{P}_\epsilon(f_2)$ 只有一个正常返类 $\{1, 2\}$ 附加一些瞬时状态 $\{3, 4\}$ 。状态 $\{3, 4\}$ 成为瞬时状态，是由于摄动造成的。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 35 of 64

Go Back

Full Screen

Close

Quit



## 5.1. 摄动下子图的稳态分布

将图 $G$ 嵌入到MDP问题中后，MDP的每一个平稳策略 $f$ 诱导了一个 $G$ 的子图，所对应马氏链的平稳概率分布利用随机过程的知识很容易计算出来。摄动下，任意平稳策略和随机平稳策略诱导的马氏链都是单链的。据马氏链的理论，随机平稳策略 $\pi = \pi_0^\infty$ 对应的稳态分布 $\mathbf{p}_\epsilon^*(\pi)$ 具有形式：

$$\mathbf{p}_\epsilon^*(\pi) = (p_1^*(\pi), p_2^*(\pi), \dots, p_N^*(\pi)), \quad (5)$$

摄动之后转移概率矩阵 $\mathbf{P}_\epsilon(\pi)$ 的极限矩阵 $\mathbf{P}_\epsilon^*(\pi)$ 为：

$$\lim_{n \rightarrow \infty} \mathbf{P}_\epsilon^n(\pi) = \mathbf{P}_\epsilon^*(\pi) = \begin{bmatrix} \mathbf{p}_\epsilon^*(\pi) \\ \mathbf{p}_\epsilon^*(\pi) \\ \vdots \\ \mathbf{p}_\epsilon^*(\pi) \end{bmatrix}. \quad (6)$$

对于平稳策略 $f^\infty$ ，我们将上面两个式子中的 $\pi$ 以 $f$ 替代。



首先定义符号:

$$d_m(\epsilon) \triangleq 1 + \sum_{i=2}^m (1 - \epsilon)^{i-2}, \quad (7)$$

其中  $m = 2, 3, \dots, N$ 。

**定理5.1:** 对于  $\epsilon > 0$ , 如果平稳策略  $f \in C_m^1$ , 那么它的非对称线性摄动下的稳态分布  $\mathbf{p}_\epsilon^*(f)$  为:

$$\mathbf{p}_\epsilon^*(f) = \left( \frac{1}{d_m(\epsilon)}, \frac{1}{d_m(\epsilon)}, \frac{1 - \epsilon}{d_m(\epsilon)}, \dots, \frac{(1 - \epsilon)^{m-2}}{d_m(\epsilon)}, 0, \dots, 0 \right). \quad (8)$$

特别的, 当  $m = N$  时有:

$$\mathbf{p}_\epsilon^*(f) = \left( \frac{1}{d_N(\epsilon)}, \frac{1}{d_N(\epsilon)}, \frac{1 - \epsilon}{d_N(\epsilon)}, \dots, \frac{(1 - \epsilon)^{N-2}}{d_N(\epsilon)} \right). \quad (9)$$



哈密顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Page 38 of 64

Go Back

Full Screen

Close

Quit

下面我们考察  $f \in \mathcal{N}_c$  的情形。当  $f \in \mathcal{N}_c$  时，状态1在未摄动时是瞬时状态，因此调整除了家顶点以外的顶点的顺序，可以认为  $f$  的一般形式是：

$$f = ((1, 2), (2, 3), \dots, (k, k+1), \dots, (k+m, k), \dots),$$

其中  $2 \leq k \leq N$ ,  $0 \leq m \leq N - k + 1$ 。将从状态1出发经过  $k - 1$  步进入一个圈长为  $m$  的子图全体记为  $C_m^k$ ，那么  $\mathcal{N}_c$  可以分为一些相互不交的  $C_m^k$  的并。

首先我们考虑一个特殊问题  $f \in C_{N-1}^2$ 。

**定理5.2:** 假设  $f = (2, 3, \dots, N, 2) \in C_{N-1}^2$ , 对应的摄动转移概率矩阵为:

$$\mathbf{P}_\epsilon(f) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & 1-\epsilon & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon & 0 & 0 & 0 & \cdots & 0 & 1-\epsilon \\ \epsilon & 1-\epsilon & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (10)$$

则对应的平稳概率分布为:

$$\left( \frac{\epsilon}{1+\epsilon}, \frac{1}{k_N}, \frac{1-\epsilon}{k_N}, \dots, \frac{(1-\epsilon)^{N-2}}{k_N} \right), \quad (11)$$

其中

$$k_N = (1+\epsilon) \sum_{i=0}^{N-2} (1-\epsilon)^i. \quad (12)$$



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 39 of 64

Go Back

Full Screen

Close

Quit



一般的问题,  $f \in C_{N-i+1}^i$  的情形, 也就是整个非摄动的子图中只有状态  $1, 2, \dots, i-1$  是瞬时的, 其余的状态构成一个遍历类, 这里  $N \geq i \geq 3$  (定理5.2 实际上是考虑  $i = 2$  的情况)。

**引理5.1:** 对于  $\epsilon > 0$ , 如果平稳策略  $f \in C_{N-i+1}^i$ , 其中  $N \geq i \geq 3$ , 那么, 它在非对称线性摄动下的稳态分布  $\mathbf{p}_\epsilon^* = (p_1, p_2, \dots, p_N)$  为:

$$\mathbf{p}_\epsilon^* = \left( \frac{\epsilon}{1+\epsilon}, \frac{\epsilon}{1+\epsilon}, \frac{\epsilon(1-\epsilon)}{1+\epsilon}, \dots, \frac{\epsilon(1-\epsilon)^{i-3}}{1+\epsilon}, \right. \\ \left. \frac{1}{N-i+1} + O(\epsilon), \dots, \frac{1}{N-i+1} + O(\epsilon) \right). \quad (13)$$





哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 41 of 64

Go Back

Full Screen

Close

Quit

**定理5.3:** 对于 $\epsilon > 0$ 和 $N \geq i \geq 3$ , 如果平稳策略 $f \in C_{N-i+1}^i \subset \mathcal{N}_c$ , 它在非对称线性摄动下的稳态分布 $\mathbf{p}_\epsilon^*(f)$ 为:

$$\mathbf{p}_\epsilon^*(f) = \left( \frac{\epsilon}{1+\epsilon}, \frac{\epsilon}{1+\epsilon}, \frac{\epsilon(1-\epsilon)}{1+\epsilon}, \dots, \frac{\epsilon(1-\epsilon)^{i-3}}{1+\epsilon}, \right. \\ \left. \frac{(1-\epsilon)^{i-2}}{k_N(i)}, \dots, \frac{(1-\epsilon)^{i-3+j}}{k_N(i)}, \dots, \frac{(1-\epsilon)^{N-2}}{k_N(i)} \right), \quad (14)$$

其中 $j = 1, 2, \dots, N - i + 1$ 以及

$$k_N(i) = (1+\epsilon) \sum_{j=0}^{N-i} (1-\epsilon)^j = \frac{1+\epsilon}{\epsilon} (1 - (1-\epsilon)^{N-i+1}). \quad (15)$$



根据定理5.1的结论我们知道，状态1如果包含在一个圈中，那么状态1对应的摄动后的稳态分布取值为 $\frac{1}{d_m(\epsilon)}$ ，其中 $m$ 是圈长。很明显，当摄动因子 $\epsilon$ 充分接近于0的时候，有：

$$\frac{1}{d_N(\epsilon)} < \frac{1}{d_{N-1}(\epsilon)} < \cdots < \frac{1}{d_2(\epsilon)}. \quad (16)$$

也就是说一个HC对应的状态1的稳态分布取值是最小的。如果我们按照公式(2)定义报酬 $r(i, a)$ ，准则值取平均报酬准则，在这样构造的MDP问题中一个HC恰好对应所有含状态1的圈中最小值，即平均报酬在含圈的平稳策略中达到最小。一个自然的想法就是把状态1的稳态分布值作为目标函数，但遗憾的是，根据定理5.3的结论，如果状态1不在圈上，摄动后的稳态分布取值为 $\frac{\epsilon}{1+\epsilon}$ 。因此说明，当摄动因子 $\epsilon$ 充分接近于0的时候，HC仅仅对应次最小值。

## 5.2. 几个例子

在考虑用优化MDP问题求解HC问题时，如果仅把状态1的稳态分布值作为目标函数，并不能保证HC对应的目标函数达到最小值。如果考虑到基本矩阵的因素之后，又会是怎么样呢？在通过数值例子说明之前，我们先给出一个重要的定义： $\mathcal{L}$ 函数。

**定义5.1：** 对任意的平稳策略  $f \in F$ ，根据偏差矩阵(参考第??章)的定义：

$$\mathbf{H}(f) = [\mathbf{I} - \mathbf{P}(f) + \mathbf{P}^*(f)]^{-1} - \mathbf{P}^*(f), \quad (17)$$

则平稳策略  $f$  的偏差就是  $\mathbf{h}(f) = \mathbf{H}(f)\mathbf{r}(f)$  (参见公式(??))。我们特别的将状态1的稳态分布值  $p_1^*(f)$  与偏差  $\mathbf{h}(f)$  的第1个分量  $h(1, f)$  的和记为  $\mathcal{L}(f)$ ，即

$$\mathcal{L}(f) = p_1^*(f) + h(1, f), \quad (18)$$

称为平稳策略  $f$  的  $\mathcal{L}$  函数。事实上，很容易看出  $\mathcal{L}$  函数恰好是矩阵  $[\mathbf{I} - \mathbf{P} + \mathbf{P}^*]^{-1}$  的第(1, 1)个分量。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

« »

◀ ▶

Page 43 of 64

Go Back

Full Screen

Close

Quit



**例5.1:** 考虑两个顶点的图。此时，状态空间 $S = \{1, 2\}$ 。在每个状态上都有两个行动 $\{1, 2\}$ ，其中行动 $i$ 表示选择 $(j, i)$ 这条边。所以这个图有4个子图，分别为：

$$\begin{aligned} f_1 &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}; & f_2 &= \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \\ f_3 &= \begin{bmatrix} 2 \\ 1 \end{bmatrix}; & f_4 &= \begin{bmatrix} 2 \\ 2 \end{bmatrix}. \end{aligned}$$

根据定理5.1、定理5.3和公式(17)，有：

$$\begin{aligned} \mathbf{h}(f_3) &= [\mathbf{I} - \mathbf{P}_\epsilon(f_3) + \mathbf{P}_\epsilon^*(f_3)]^{-1} \mathbf{r}(f_3) - \mathbf{P}_\epsilon^*(f_3) \mathbf{r}(f_3) \\ &= \left[ \left( \mathbf{I} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right)^{-1} - \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right] \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix} \end{aligned}$$



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

以及

$$p_1^*(f_3) = \frac{1}{d_2(\epsilon)} = \frac{1}{2}.$$

所以，平稳策略 $f_3$ 的 $\mathcal{L}$ 函数为：

$$\mathcal{L}(f_3) = p_1^*(f_3) + h(1, f_3) = \frac{1}{2} + \frac{1}{4}$$

同样计算，我们有：

$$\begin{aligned}\mathcal{L}(f_1) &= 1; \\ \mathcal{L}(f_2) &= 1; \\ \mathcal{L}(f_4) &= 1 - \frac{\epsilon}{1 + \epsilon}.\end{aligned}$$

很明显作为HC的 $f_3$ ，其 $\mathcal{L}$ 函数 $\mathcal{L}(f_3)$ 当 $\epsilon \rightarrow 0$ 的时候，在所有平稳策略中达到最小。

Home Page

Title Page

◀

▶

◀

▶

Page 45 of 64

Go Back

Full Screen

Close

Quit



**例5.2:** 考虑三个顶点的图。此时, 状态空间 $S = \{1, 2, 3\}$ , 对每个状态 $i \in S$ ,  $A(i) = \{1, 2, 3\}$ 。这个图共有27个子图。我们先考虑对应HC的平稳策略 $f = (2, 3, 1)$ 其转移概率矩阵为:

$$\mathbf{P}(f) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

根据定理5.1、定理5.3和公式(17), 有:

$$\begin{aligned} \mathbf{h}(f) &= [\mathbf{I} - \mathbf{P}_\epsilon(f) + \mathbf{P}_\epsilon^*(f)]^{-1} \mathbf{r}(f) - \mathbf{P}_\epsilon^*(f) \mathbf{r}(f) \\ &= \begin{bmatrix} \frac{6-3\epsilon}{9-6\epsilon+\epsilon^2} - \frac{1}{3-\epsilon} \\ \star \\ \star \end{bmatrix}, \end{aligned}$$

其中 $\star$ 是我们并不关心的量,



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 47 of 64

Go Back

Full Screen

Close

Quit

以及

$$p_1^*(f) = \frac{1}{d_3(\epsilon)} = \frac{1}{3 - \epsilon}.$$

所以, 当 $\epsilon \rightarrow 0$ 时

$$\mathcal{L}(f) = p_1^*(f) + h(1, f) = \frac{6 - 3\epsilon}{9 - 6\epsilon + \epsilon^2} \rightarrow \frac{2}{3}.$$

类似的, 平稳策略 $f_1 = (3, 1, 2)$ 也对应着一个HC, 经过同样的计算有:

$$\mathcal{L}(f_1) = \frac{6 - 3\epsilon}{9 - 6\epsilon + \epsilon^2}.$$

平稳策略 $f_2 = (2, 1, 1)$ ,  $f_3 = (2, 1, 2)$ ,  $f_4 = (2, 1, 3)$ ,  $f_5 = (3, 1, 1)$ ,  $f_6 = (3, 2, 1)$ 和 $f_7 = (3, 3, 1)$ 是圈长为2的HC, 根据例5.1, 有

$$\mathcal{L}(f_2) = \mathcal{L}(f_3) = \cdots = \mathcal{L}(f_7) = \frac{3}{4}.$$



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 48 of 64

Go Back

Full Screen

Close

Quit

平稳策略  $f_8 = (1, 1, 1)$ ,  $f_9 = (1, 1, 2)$ ,  $f_{10} = (1, 1, 3)$ ,  $f_{11} = (1, 2, 1)$ ,  $f_{12} = (1, 2, 2)$ ,  $f_{13} = (1, 2, 3)$ ,  $f_{14} = (1, 3, 1)$ ,  $f_{15} = (1, 3, 2)$  以及  $f_{16} = (1, 3, 3)$  有

$$\mathcal{L}(f_9) = \cdots = \mathcal{L}(f_{16}) = 1.$$

所有剩下的10个平稳策略  $f_{17}$  到  $f_{26}$  状态1都不在非摄动子图的任何圈上，经过类似的计算，我们有：

$$\mathcal{L}(f_{17}) = \cdots = \mathcal{L}(f_{26}) = \frac{1 + \epsilon + \epsilon^2}{1 + 2\epsilon + \epsilon^2},$$

当  $\epsilon \rightarrow 0$  的时候趋于1。

因为

$$\frac{2}{3} < \frac{3}{4} < 1,$$

所以HC对应的  $\mathcal{L}$  函数在摄动接近于0的时候在平稳策略类中达到最小。





根据例子，我们得到下面的表格。

表 7.1 数例结果的比较

图的顶点数	HC 的 $\mathcal{L}$ 函数	
	精确解	逼近解
2	$\frac{3}{4}$	$\frac{3}{4}$
3	$\frac{6-3\epsilon}{9-6\epsilon+\epsilon^2}$	$\frac{2}{3} + O(\epsilon)$
4	$\frac{10-11\epsilon+4\epsilon^2}{16-24\epsilon+17\epsilon^2-6\epsilon^3+\epsilon^4}$	$\frac{5}{8} + O(\epsilon)$
5	$\frac{15-26\epsilon+19\epsilon^2-5\epsilon^3}{25-60\epsilon+76\epsilon^2-58\epsilon^3+28\epsilon^4-8\epsilon^5+\epsilon^6}$	$\frac{3}{5} + O(\epsilon)$

容易从表中看到HC的 $\mathcal{L}$ 函数与图的顶点数之间有着近似的关系：

$$\frac{N+1}{2N} + O(\epsilon), \quad (19)$$

其中 $N$ 表示图的顶点数目。

### 5.3. 摄动下HC的性质

下面要证明公式(19)具有普遍意义，而且对应HC的平稳策略的 $\mathcal{L}$ 函数在平稳策略类中达到最小。我们证明几个定理。

**定理5.4:** 对于 $N$ 个顶点的图 $G$ 的子图(HC) $f = (2, 3, \dots, N, 1) \in C_N^1$ ，摄动后的一步转移概率矩阵为：

$$\mathbf{P}_\epsilon(f) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 1 - \epsilon & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & 1 - \epsilon & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon & 0 & 0 & 0 & \cdots & 0 & 1 - \epsilon \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (20)$$

那么，有

$$\mathcal{L}(f) = \frac{N+1}{2N} + O(\epsilon). \quad (21)$$



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 50 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

**推论5.1:** 对于子图  $f = (2, 3, \dots, N, 1) \in C_N^1$ , 存在  $\epsilon^* > 0$ , 当  $\epsilon \in [0, \epsilon^*)$  时定理5.4都成立。

**推论5.2:** 任何HC对应的子图  $f \in C_N^1$ , 都有

$$\mathcal{L}(f) = \frac{N+1}{2N} + O(\epsilon).$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 51 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 52 of 64

Go Back

Full Screen

Close

Quit

**定理5.5:** 对于 $N$ 个顶点的图 $G$ 的子图 $f = (2, 3, \dots, m, 1, \dots)$ , 摄动后的一步转移概率矩阵具有形式

$$\mathbf{P}_\epsilon(f) = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & \mathbf{C} \end{bmatrix},$$

其中 $\mathbf{A}$ 为 $m \times m$ 的子矩阵, 形式是:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 1 - \epsilon & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon & 0 & 0 & \cdots & 0 & 1 - \epsilon \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

那么, 有

$$\mathcal{L}(f) = \frac{i+1}{2i} + O(\epsilon). \quad (22)$$

**推论5.3:** 任何由平稳策略  $f \in C_m^1$  对应的子图, 都有

$$\mathcal{L}(f) = \frac{m+1}{2m} + O(\epsilon).$$

**定理5.6:** 对于  $N$  个顶点的图  $G$  的子图  $f = (2, 3, \dots, N, 2)$ , 摄动后的一步转移概率矩阵具有形式

$$\mathbf{P}_\epsilon(f) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & 1-\epsilon & \cdots & 0 & 0 \\ \epsilon & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon & 0 & 0 & 0 & \cdots & 0 & 1-\epsilon \\ \epsilon & 1-\epsilon & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

那么, 有

$$\mathcal{L}(f) = 1 + O(\epsilon). \quad (23)$$



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 53 of 64

Go Back

Full Screen

Close

Quit



**定理5.7:** 对于 $N$ 个顶点的图 $G$ 的任意子图 $f = (2, \dots, i, \dots, N, i)$ , 则有

$$\mathcal{L}(f) = 1 + O(\epsilon).$$

**定理5.8:** 对于 $N$ 个顶点的图 $G$ 的任意子图 $f \in \mathcal{N}_c$ , 有

$$\mathcal{L}(f) = 1 + O(\epsilon).$$

**定理5.9:** 如果平稳策略 $f^* \in F$ 对应于一个HC, 则当摄动因子 $\epsilon$ 充分小的时候, 其 $\mathcal{L}$ 函数在平稳策略类中达到最小, 即

$$\begin{aligned} \mathcal{L}(f^*) &= [\mathbf{I} - \mathbf{P}_\epsilon(f^*) + \mathbf{P}_\epsilon^*(f^*)]^{-1}(1, 1) \\ &= \min_{f \in F} \{\mathcal{L}(f)\}, \end{aligned} \quad (24)$$

其中 $\mathbf{H}^{-1}(1, 1)$ 表示矩阵 $\mathbf{H}$ 逆的第 $(1, 1)$ 个分量。



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

## 6 更为精细的分析

定理5.9已经说明当摄动因子充分小的时候，HC所对应的平稳策略在平稳策略类中的 $\mathcal{L}$ 函数值达到最小。但是并没有更加精细的分析摄动小到什么时候，HC一定是平稳策略类中最小的（最优的）平稳策略。

Home Page

Title Page

◀

▶

◀

▶

Page 55 of 64

Go Back

Full Screen

Close

Quit



我们首先需要知道的是HC对应平稳策略的精确的 $\mathcal{L}$ 函数值。

**定理6.1:** 如果平稳策略 $f \in F$ 对应于一个HC, 当非对称线性摄动因子为 $\epsilon$ 时, 其 $\mathcal{L}$ 函数为:

$$\mathcal{L}(f) = \frac{1}{d_N(\epsilon)} - \frac{1}{d_N^2(\epsilon)} + \frac{1}{\epsilon d_N(\epsilon)} - \frac{1 - \epsilon + (N - 1)(1 - \epsilon)^{N-1}}{\epsilon d_N^2(\epsilon)}. \quad (25)$$





**推论6.1:** 当 $m > 2$ , 如果平稳策略 $f \in C_m^1$ , 则有

$$\mathcal{L}(f) = \frac{1}{d_m(\epsilon)} - \frac{1}{d_m^2(\epsilon)} + \frac{1}{\epsilon d_m(\epsilon)} - \frac{(1 - \epsilon) + (m - 1)(1 - \epsilon)^{m-1}}{\epsilon d_m^2(\epsilon)}.$$

**定理6.2:** 如果摄动因子 $\epsilon \leq \frac{1}{N^2}$ , 对任意的平稳策略 $f_{N-1} \in C_{N-1}^1$ 和对应HC的平稳策略 $f = f_{HC} \in C_N^1$ , 相应的 $\mathcal{L}$ 函数满足:

$$\mathcal{L}(f) = \mathcal{L}(f_{HC}) < \mathcal{L}(f_{N-1}). \quad (26)$$



由此，我们得到主要的定理。

**定理6.3:** 设图为 $G$ ，那么

1) 如果 $G$ 含有子图HC且对应的平稳策略为 $f_{HC}$ ，充分必要条件是 $N \geq 3$ 且当 $\epsilon \leq \frac{1}{N^2}$ 时，

$$\min_{f \in F} \mathcal{L}(f) = \mathcal{L}(f_{HC}) \triangleq \mathcal{L}_{HC},$$

其中

$$\mathcal{L}_{HC} = \frac{1}{d_N(\epsilon)} - \frac{1}{d_N^2(\epsilon)} + \frac{1}{\epsilon d_N(\epsilon)} - \frac{(1 - \epsilon) + (N - 1)(1 - \epsilon)^{N-1}}{\epsilon d_N^2(\epsilon)};$$

2) 如果 $G$ 不含有子图HC，充分必要条件是 $\forall \epsilon \leq \frac{1}{N^2}$ ， $N \geq 3$ ，成立

$$\min_{f \in F} \mathcal{L}(f) > \mathcal{L}_{HC}.$$



## 7 开问题和有关猜想

根据定理5.9或者定理6.3的结论，早在1997年我们给出了下面的猜想，后来文献又精细化了这个猜想。我们这里采用精细化后的叙述，两者的差异仅在于摄动因子有了明确的上界。

**猜想7.1:** 图 $G$ 含有子图HC的充分必要条件是当 $\epsilon \leq \frac{1}{N^2}$ 时:

$$\min_{f \in \Pi_s} \mathcal{L}(f) = \mathcal{L}_{HC},$$

其中 $\mathcal{L}_{HC}$ 引用在定理6.3中定义，而 $\Pi_s$ 则表示随机平稳策略类。



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

最早在1994年Filar等开始利用马氏链的平稳分布思想导出一些算法求解HCP。后来，我们最早给出了 $\mathcal{L}$ 函数的定义，证明了以此为决策目标的马氏决策问题，由HC对应的平稳策略在平稳策略类中达到最小，并提出了猜想7.1的原形。但遗憾的是，到目前为止，还不能够证明这个猜想。如果猜想7.1成立，完全有可能在随机平稳策略类中建立有效的算法求解确定性平稳策略对应的HC，而所有的确定性平稳策略构成了随机平稳策略类形成的凸多面体的所有顶点。如果猜想7.1不成立，就说明存在一个“内点”，它是由某个随机平稳策略 $\pi_{LM} \in \Pi_s$ 使目标值达到最小，即 $\mathcal{L}(\pi_{LM}) < \mathcal{L}_{HC}$ 。类似的，由于全局最优解严格的比每一个基本解都小这一点也阻碍了利用线性规划优势的可能。所以，这个猜想的确认将会非常有助于有效的启发式算法的设计。借鉴一些文献中的方法有可能帮助这个猜想的证明。

[Home Page](#)

[Title Page](#)

[«](#) [»](#)

[◀](#) [▶](#)

Page 60 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



哈密尔顿圈问题

有向图到MDP的嵌入

平稳策略的分类

约束折扣MDP与HC

HCP嵌入MDP的摄动

更为精细的分析

开问题和有关猜想

利用Maple通过数值的方法文献Ejov等人验证了猜想7.1的部分成立。具体的方法是考虑任意两个平稳策略 $f, g \in F$ 的凸组合，即考察随机平稳策略

$$\pi(\lambda, f, g) \triangleq \lambda f + (1 - \lambda)g, \quad \lambda \in [0, 1],$$

其转移概率为

$$\mathbf{P}_\epsilon(\pi(\lambda, f, g)) = \lambda \mathbf{P}_\epsilon(f) + (1 - \lambda) \mathbf{P}_\epsilon(g).$$

Home Page

Title Page

◀

▶

◀

▶

Page 61 of 64

Go Back

Full Screen

Close

Quit



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

Home Page

Title Page

◀ ▶

◀ ▶

Page 62 of 64

Go Back

Full Screen

Close

Quit

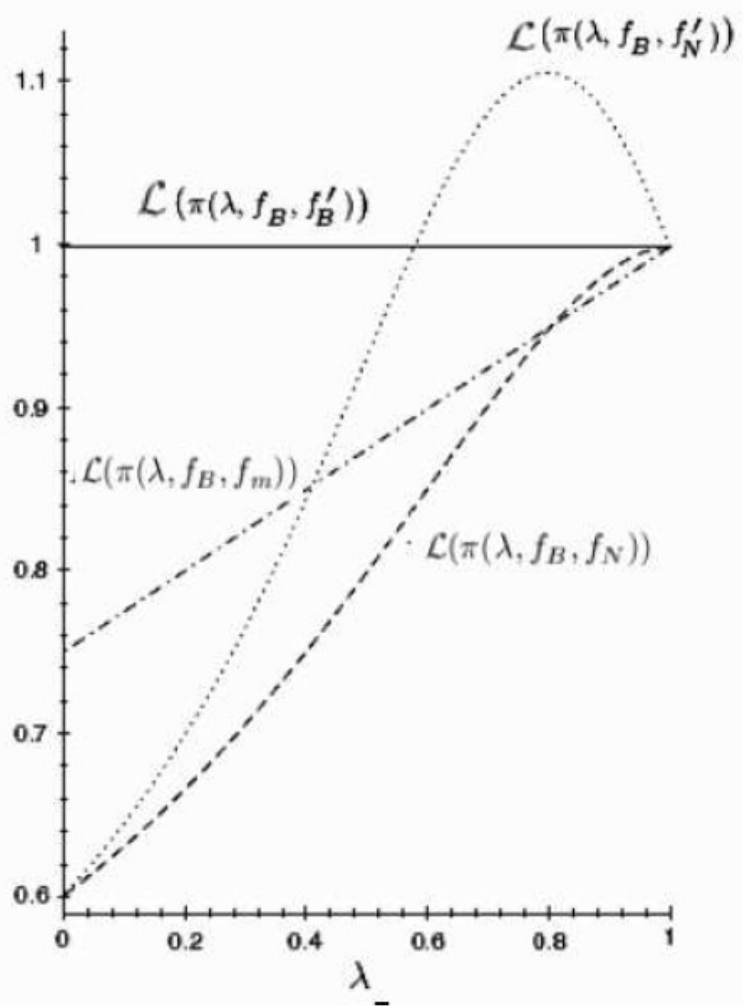
在 $N \leq 20$ 时验证了 $\pi(\lambda, f, g)$ 的目标函数,  $\mathcal{L}$ 函数, 满足:

$$\mathcal{L}(\pi(\lambda, f, g)) \geq \mathcal{L}(f_{HC}),$$

而且等号仅在策略 $\pi(\lambda, f, g) = f_{HC}$ 的时候成立, 即取 $g = f_{HC}$ 并且 $\lambda = 0$ 。下图描述了平稳策略 $f_B, g_B$ 取值于集合 $B$ ,  $f_{HC}$ 以及 $f_m \in C_m^1$ ,  $m < N$ 等不同情况的 $\mathcal{L}$ 函数的比较。



哈密尔顿圈问题  
 有向图到MDP的嵌入  
 平稳策略的分类  
 约束折扣MDP与HC  
 HCP嵌入MDP的摄动  
 更为精细的分析  
 开问题和有关猜想



Home Page

Title Page

◀ ▶

◀ ▶

Page 63 of 64

Go Back

Full Screen

Close

Quit

# 谢谢大家!



哈密尔顿圈问题  
有向图到MDP的嵌入  
平稳策略的分类  
约束折扣MDP与HC  
HCP嵌入MDP的摄动  
更为精细的分析  
开问题和有关猜想

[Home Page](#)

[Title Page](#)



Page 64 of 64

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)