第四讲：混合策略、纳什定理

Mixed Strategy and Nash Theorem

基于单纯策略的纳什均衡并非总是存在

	石头	剪子	布
石头	0, 0	1, -1	-1, 1
剪子	-1, 1	0, 0	1, -1
布	1, -1	-1, 1	0, 0

- 显然，这个博弈没有基于单纯策略(pure strategy)的纳什均衡

混合策略 (Mixed Strategy)

- 混合策略：博弈者以一定的概率随机选择某些行动
  - 比如1/2的概率出石头，1/3的概率出剪子，1/6的概率出布
- 博弈者i的行动集为 $A_i = \{a_{i1}, a_{i2}, \ldots, a_{im}\}$，他的一个混合策略记为 $\alpha_i = (p_{i1}, p_{i2}, \ldots, p_{im})$，其中$p_{ik}$是行动$a_{ik}$的概率，$\sum_{k=1}^{m} p_{ik} = 1$
  - 例如，石头剪子布中i的一个可能的混合策略是：(1/2，1/3，1/6)
- 也就是说，一个混合策略是在自己行动集上的一个概率分布
  - 一个单纯策略，也可以看作是一个（“退化”了的）混合策略：某一行动的概率是1，其他行动的概率是0
  - 一个策略组合(strategy profile) 记为$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$，为方便有时也记为$(\alpha_i, \alpha_{-i})$

混合策略的（期望）收益

- 例：如何计算右图中博弈者1的混合策略的收益

			博弈者2		
			1/3	1/6	1/2
			石头	剪子	布
博弈者1	1/2	石头	0, 0	1, -1	-1, 1
	1/3	剪子	-1, 1	0, 0	1, -1
	1/6	布	1, -1	-1, 1	0, 0

- 第一步：先考虑1的每一个**单纯策略**的收益
  - 给定2策略：(1/3, 1/6, 1/2)
  - 博弈者1用“石头”的期望收益：0*1/3 + 1*1/6 + (-1)*1/2 = -1/3
  - 博弈者1用“剪子”的期望收益：(-1) *1/3 + 0*1/6 + 1*1/2 = 1/6
  - 博弈者1用“布”的期望收益： 1*1/3 + (-1) *1/6 + 0*1/2 = 1/6
- 第二步：将单纯策略的收益按照相应概率加权平均

(-1/3) *1/2 + (1/6)* 1/3 + (1/6)* 1/6 = -1/12

混合策略纳什均衡

定义：

在一个策略组合$\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_n^*)$中，如果对于任意一名博弈者i来说， 给定其他博弈者的策略α_{-i}^*，他现在的策略α_i^*是他所有可选策略中最优的，也就是说

$u_i(\alpha_i^*, \alpha_{-i}^*) \geq u_i(\alpha_i, \alpha_{-i}^*)$, $\forall \alpha_i \in \boldsymbol{A}_i$，$\boldsymbol{A}_i$是i所有混合策略的集合

那么该策略组合α^*是一个纳什均衡

找出混合策略纳什均衡

- 定义法
  - 可以先猜测均衡在哪里，然后根据定义证明

- 最常用：最优反应法
  - 关键：当一个混合策略是最优反应时，构成该策略的每个单纯策略所带来的收益必须正好相等
    - 如果两个单纯策略所带来的收益一个高一个低，那么我总是采用其中高收益的那个策略所获得的收益，肯定是超过我混合采用这两个策略所获得的收益

例：“对硬币”(Matching Pennies)

- 双方“对上”则博弈者1获胜，“对不上”则2获胜
- 现实应用：“华容道”、战争的攻防
- 容易看出，不存在单纯策略的纳什均衡

		博弈者2	
		正面	反面
博弈者1	正面	1, -1	-1, 1
	反面	-1, 1	1, -1

寻找混合策略最优反应

1的混合策略α_1：(p, 1-p)
2的混合策略α_2：(q, 1-q)

			博弈者2	
			q	1-q
			H	T
博弈者1	p	H	1, -1	-1, 1
	1-p	T	-1, 1	1, -1

- 试着分析博弈者1的最优反应
- 给定对方策略，1的每个**单纯策略**的收益为

$$u_1(H, \alpha_2) = 1q + (-1)(1-q) = 2q-1$$

$$u_1(T, \alpha_2) = (-1)q + 1(1-q) = 1-2q$$

- 1的混合策略的收益*

$$u_1(\alpha_1, \alpha_2) = p\cdot(2q-1) + (1-p)\cdot(1-2q) = 4pq-2p-2q+1$$

*但是我们一般并不用这个总收益来确定最优反应，而是考察1的各个单纯策略的收益

寻找混合策略最优反应

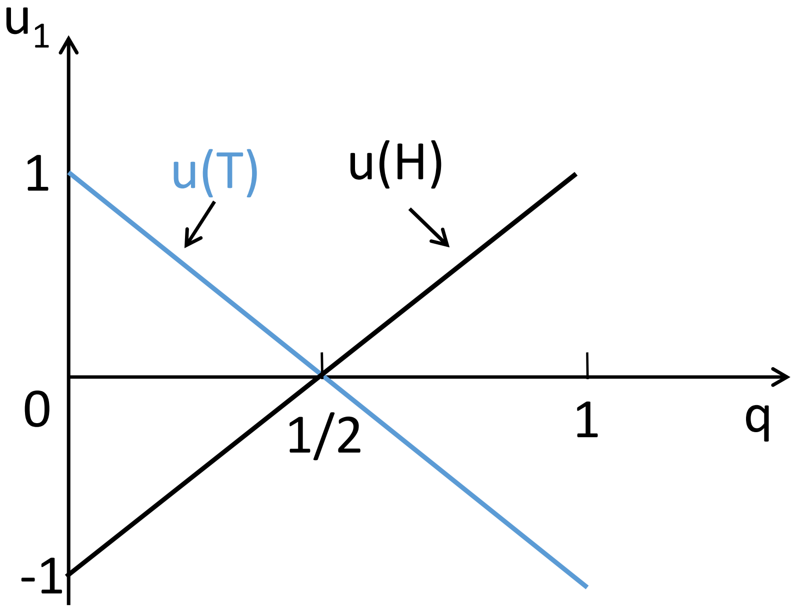
- 1的每个单纯策略收益为

  $u_1(H) = 2q - 1$

  $u_1(T) = 1 - 2q$

- 确定最优：比较两个单纯策略的收益
  - 当q<1/2时，1选T最优
  - 当q>1/2时，1选H最优
  - 当q=1/2时，T和H的收益**相等**
    - **只有这个时候，1随机选择T和H才可**能是最优反应
    - 以什么概率p混合？
    - 从1的收益角度看，**任意p均可**

			博弈者2	
			q	1-q
			H	T
博弈者1	p	H	1, -1	-1, 1
	1-p	T	-1, 1	1, -1



博弈者1的单纯策略收益（作为q的函数）

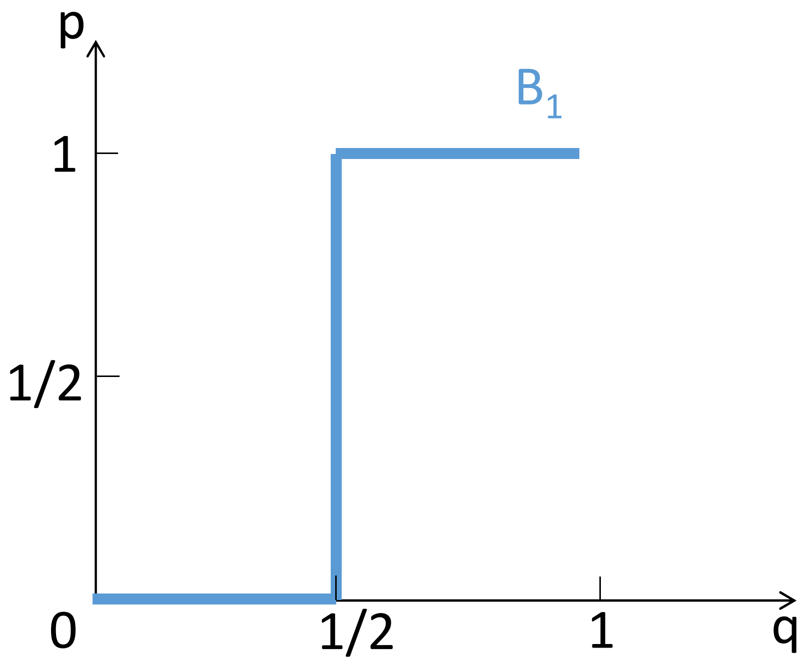
混合策略作为最优反应

- 核心思想：当博弈者的最优反应是**随机混合**某些单纯策略时，这些单纯策略中的每一个带来的收益一定是**相等**的（也就等于该混合策略的收益）
- 所以我们在寻找混合策略的最优反应时，总是在找**某些单纯策略收益相等**的点
- 当一个混合策略是最优反应时，由于它其中的每一个单纯策略的收益相等，那么**无论用什么概率**混合这些策略，带来的期望收益总是一样的

博弈者1的最优反应映射

			博弈者2	
			q	1-q
			H	T
博弈者1	p	H	1，-1	-1，1
	1-p	T	-1，1	1，-1

- 1的最优反应
  - 当q<1/2时，1总是选T最优，即p=0
  - 当q>1/2时，1总是选H最优，即p=1
  - 当q=1/2时，1选T和选H的收益**相等**
    - **只有这个时候**，1采用随机的混合策略才可能是最优反应；且**任意p均可**
- 可以画出最优反应映射，如右图



博弈者1的最优反应映射

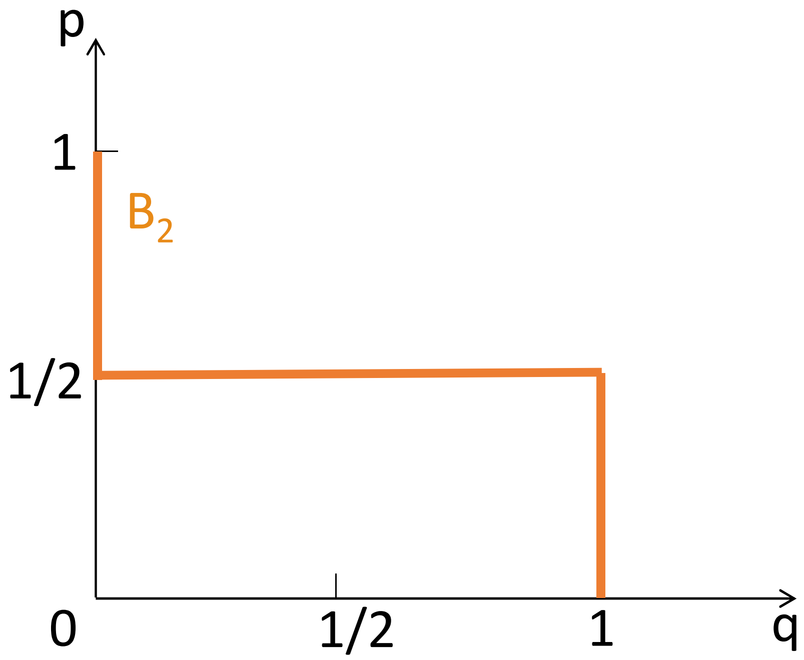
博弈者2的最优反应

- 2的每个单纯策略收益为

  $u_2(H) = -1p + 1\ (1-p) = 1 – 2p$

  $u_2(T) = 1p + -1(1-p) = 2p – 1$
- 何时单纯策略的收益相等？p=1/2
- 2的最优反应
  - 当p < 1/2时，2总是选H最优，即q=1
  - 当p > 1/2时，2总是选T最优，即q=0
  - 当p = 1/2时，2选T和H的收益**相等**
    - **只有这个时候**，2采用随机的混合策略才可能是最优反应；2选择**任意q均可**
- 如右图

			博弈者2	
			q	1-q
			H	T
博弈者1	p	H	1, -1	-1, 1
	1-p	T	-1, 1	1, -1

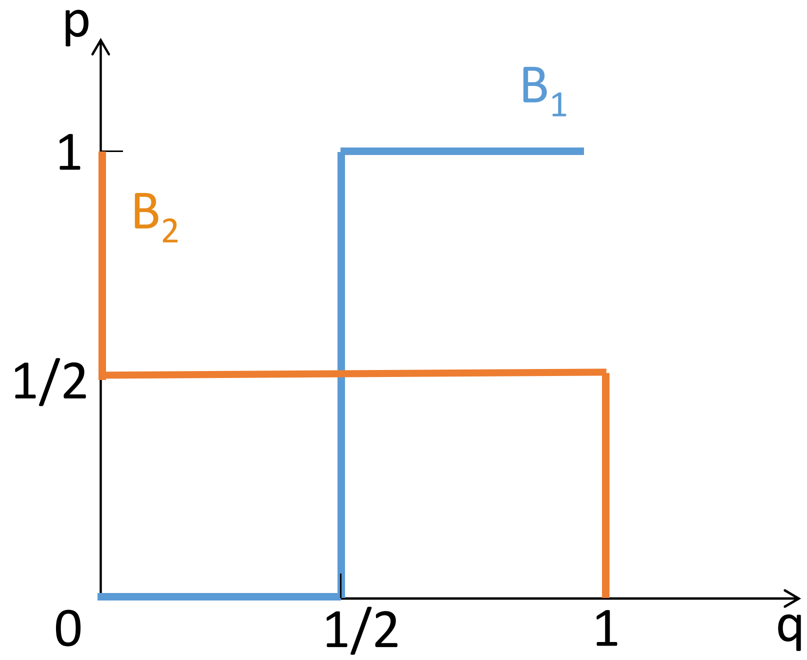


博弈者1和2的最优反应映射

纳什均衡

- (p=1/2, q=1/2) 是纳什均衡
  - –严格应写作 ((1/2, 1/2), (1/2, 1/2))

			博弈者2	
			q	1-q
			H	T
博弈者1	p	H	1, -1	-1, 1
	1-p	T	-1, 1	1, -1



博弈者1和2的最优反应映射

混合策略纳什均衡

定理：一个混合策略组合是纳什均衡，**当且仅当**

- 对任一博弈者而言，他的混合策略中使用概率大于零的每一个单纯策略所带来的收益**相等**
- 任一**未在**该混合策略中使用的单纯策略，其收益**小于等于**在混合策略中使用的单纯策略的收益

- 如何寻找混合策略均衡
  - 寻找某些单纯策略的收益相等的点，而且这些单纯策略的收益需要大于等于其他的单纯策略
  - 另外：严格劣势策略一定不会出现在混合策略均衡中

点球

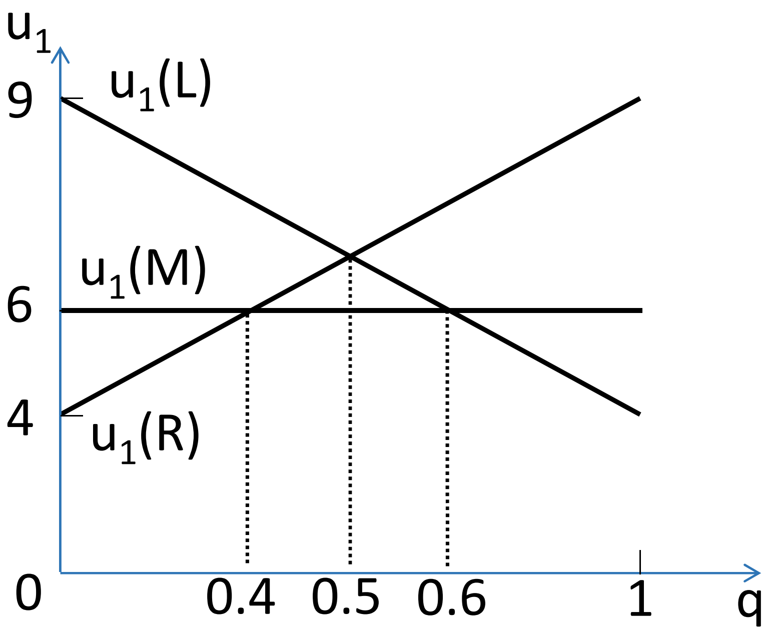
- 没有单纯策略的纳什均衡
- 考虑混合策略
- 射手的每一个单纯策略的收益：

  $u_1(L) = 4q + 9(1-q) = 9-5q$

  $u_1(M) = 6q + 6(1-q) = 6$

  $u_1(R) = 9q + 4(1-q) = 4+5q$
- 如右图
- 可能的混合点：三个
  - 当q=0.4，L最优，不会采用混合
  - 当q=0.6，R最优，不会采用混合
  - 只有当q=0.5, 可能混合L, R，且不使用M

		守门员	
		L	R
射手	L	4, -4	9, -9
	M	6, -6	6, -6
	R	9, -9	4, -4



- 所以对射手而言
  - 只有当对方使用 q=0.5, 自己可能混合L, R，且不使用M，即$p_M=0$

			守门员	
			q	1-q
			L	R
射手	p_L	L	4，-4	9，-9
	p_M	M	6，-6	6，-6
	$1-p_L-p_M$	R	9，-9	4，-4

- 对守门员而言，每个单纯策略的收益（考虑到$p_M=0$）

$u_2(L) = -4p_L - 6p_M - 9(1-p_L-p_M) = 5p_L - 9$

$u_2(R) = -9p_L - 6p_M - 4(1-p_L-p_M) = -5p_L - 4$

- 这两个策略何时收益相等？
  - $p_L=1/2$
- 纳什均衡：$\left(\left(\frac{1}{2},0,\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)$

对于混合策略均衡的解读

- 个体层面的随机行动
  - 每个行动者个体按某种概率在随机的采取行动
- 或者，群体层面的随机行动
  - 一个群体中的不同个体使用不同的单纯策略，那么混合策略中的概率是指使用相应单纯策略的个体在群体中所占的比率

小偷与警察

- 小偷可以选择作案(C) 或者不作案(NC)
- 警察可以选择睡觉(S)或者不睡觉(NS)

警察

		S	NS
小偷	C	V, -D	-P, 0
	NC	0, R	0, 0

- V -- 小偷得手, D -- 警察被罚, P -- 小偷被抓, R -- 警察偷懒，均大于 0
- 有没有单纯策略的均衡?

小偷与警察

- 小偷可以选择作案(C) 或者不作案(NC)
- 警察可以选择睡觉(S)或者不睡觉(NS)

			警察	
			q	1-q
			S	NS
小偷	p	C	V, -D	-P, 0
	1-p	NC	0, R	0, 0

- V -- 小偷得手, D -- 警察被罚, P -- 小偷被抓, R -- 警察偷懒，均大于 0
- 有没有单纯策略的均衡?
- 寻找混合策略均衡：令小偷的混合策略是(p，1-p)，警察是(q，1-q)

小偷的两个单纯策略的收益

$u_1(C) = Vq – P(1-q) = (V+P)q - P$

$u_1(NC) = 0 + 0 = 0$

两者相等意味着：$q = \frac{P}{V+P}$

警察的两个策略的收益

$u_2(S) = -Dp + R(1-p) = R – (D+R)p$

$u_2(NS) = 0 + 0 = 0$

两者相等意味着：$p = \frac{R}{D+R}$

小偷与警察

			警察	
			q	1-q
			S	NS
小偷	p	C	V, -D	-P, 0
	1-p	NC	0, R	0, 0

- 混合策略均衡中

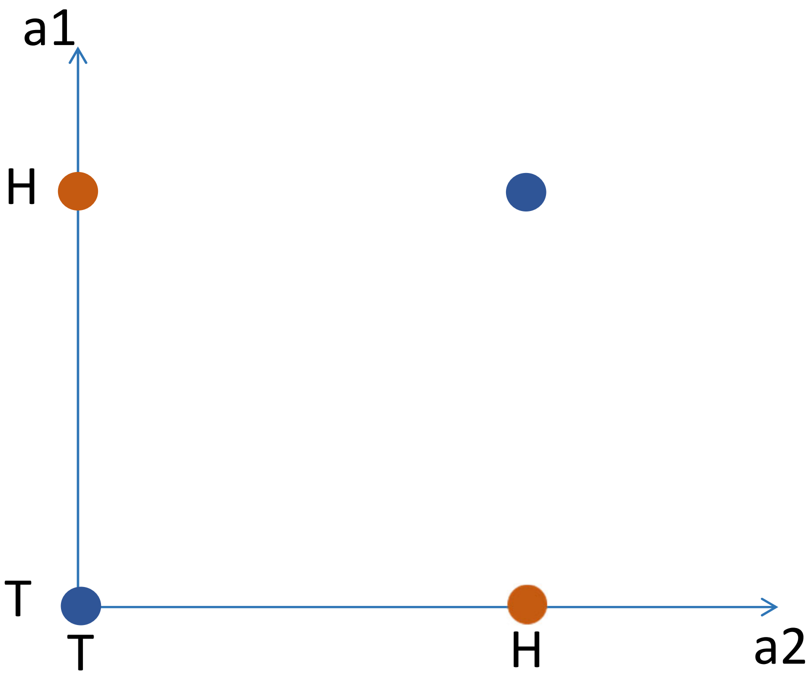
  $q = \frac{\mathrm{P}}{\mathrm{V+P}} \quad p = \frac{\mathrm{R}}{\mathrm{D+R}}$

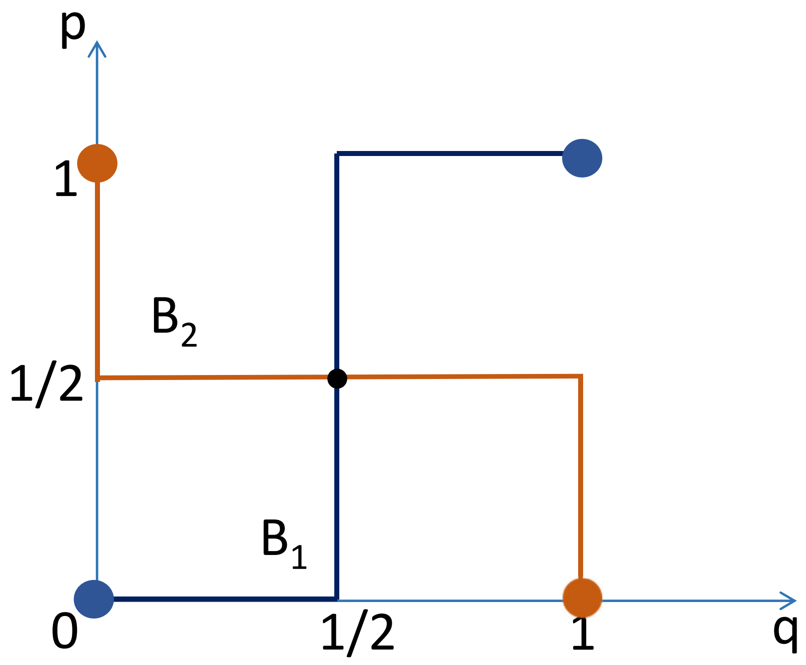
- 加重对于小偷的惩罚，即提高P，在**均衡**中
  - 并不会减少小偷的作案率 p；只是增加了警察睡觉的概率 q
  - 在短期内，小偷减少作案，但这导致警察增加睡觉，导致小偷增加作案，直到达成新的均衡
- 加重对于警察偷懒睡觉的处罚，即增大D，在**均衡**中
  - 并不会减少警察睡觉的概率 q；只是减少了小偷作案的概率 p
  - 在短期内，警察减少睡觉，但这导致小偷减少作案，导致警察增加睡觉，直到达成新的均衡

纳什定理

- 如果只考虑单纯策略，很多博弈中都不存在纳什均衡；但是混合策略的引入，使得纳什均衡的存在成为相当普遍的现象
- **纳什定理**：在任意一个有限博弈（博弈者的数目和每个博弈者可选的行动数是有限的）中，存在至少一个（混合策略的）纳什均衡
  - 我们不给出具体证明细节
- 直观上：允许使用混合策略时，最优反应函数是连续的；或者说将最优反应连续化了

		q	1-q
		H	T
p	H	1, -1	-1, 1
1-p	T	-1, 1	1, -1





纳什定理的推广

- 纳什定理适用于有限博弈
- 纳什均衡的存在性定理可以推广到（并非所有）无限博弈中，关键的条件（之一）仍然是保证最优反应映射的连续性
- 例如（仅供了解）

  定理(Glicksberg, 1952)：在n个博弈者的策略型博弈中，如果每个博弈者的行动集是欧几里得空间(R^m)中的一个非空、有界且闭合(compact)的子集，而且每个人的收益函数是连续的，那么该博弈存在（混合策略的）纳什均衡

  - 比如前面的库诺特竞争模型，若限定产量q_i属于某一闭区间如[0, 10000]，那么可适用该定理

作业2：请查教学网，3月26日课堂上交