

作业 6

1. 三种模型的实际例子与建模

1) 马尔可夫过程示例：天气变化

- 状态空间 S

$S = \{\text{晴天, 多云, 雨天}\}$

- 状态转移矩阵 P

	晴天	多云	雨天
晴天	0.7	0.2	0.1
多云	0.3	0.5	0.2
雨天	0.2	0.3	0.5

其中 $P[i][j]$ 表示从状态 i 转移到状态 j 的概率。

2) 马尔可夫回报过程示例：股票投资

- 状态空间 S

$S = \{\text{上涨, 下跌, 震荡}\}$

- 状态转移矩阵 P

	上涨	下跌	震荡
上涨	0.6	0.2	0.2
下跌	0.3	0.5	0.2
震荡	0.2	0.3	0.5

- 状态期望回报 r

$r = [0.05, -0.03, 0.01]$ // 分别对应上涨、下跌、震荡的期望收益率

- 折现因子 γ

$\gamma = 0.95$

在股票投资的例子中，折现因子表示未来收益对当前决策的重要性， γ 越大，说明投资者越重视长期收益。

3) 马尔可夫决策过程示例：学习时间分配

- 状态空间 **S**

$S = \{\text{精力充沛, 一般, 疲惫}\}$

- 行动集合 **A**

$A = \{\text{学习, 休息, 运动}\}$

- 状态转移矩阵 **P**

对于每个行动 a ，都有一个转移矩阵 $P(a)$ ：

- $P(\text{学习})$ ：

	精力充沛	一般	疲惫
精力充沛	0.4	0.4	0.2
一般	0.3	0.5	0.2
疲惫	0.2	0.3	0.5

- $P(\text{休息})$ ：

	精力充沛	一般	疲惫
精力充沛	0.7	0.2	0.1
一般	0.6	0.3	0.1
疲惫	0.5	0.3	0.2

- $P(\text{运动})$ ：

	精力充沛	一般	疲惫
精力充沛	0.6	0.3	0.1
一般	0.5	0.4	0.1
疲惫	0.4	0.4	0.2

- 行动期望回报 **R**

例如：

状态	学习	休息	运动
精力充沛	0.8	0.3	0.6
一般	0.5	0.4	0.5
疲惫	0.2	0.6	0.4

其中， $R(\text{状态}, \text{行动})$ 表示在某一状态下采取某一行动的期望回报。

- 折现因子 γ

$$\gamma = 0.9$$

在学习时间分配的例子中，折现因子表示未来学习或休息带来的收益对当前决策的影响， γ 越大，说明更重视长期的精力和学习效果。

2. 甲乙比赛问题

1) 状态空间和状态转移矩阵

- 状态空间 S

$S = \{-2, -1, 0, 1, 2\}$ // 表示甲的得分

- 状态转移矩阵 P

当前分数	-2	-1	0	1	2
-2	1	0	0	0	0
-1	q	r	p	0	0
0	0	q	r	p	0
1	0	0	q	r	p
2	0	0	0	0	1

其中 p 为甲胜概率， q 为乙胜概率， r 为平局概率。

2) 甲积 1 分时，恰好再赛两局结束比赛的概率

- 甲当前 1 分，恰好两局结束比赛的情况：

第一局平局（概率 r ），第二局甲胜（概率 p ）

- 总概率：

$$P = rp$$