1 00:00:03,230 –> 00:00:06,420 you must have gotten examinations this

2 00:00:06,420 –> 00:00:07,790 week

3 00:00:07,790 –> 00:00:13,440 studying for examinations undergraduate

4 00:00:13,440 –> 00:00:20,640 oh I think I worry even worse okay okay

你们这周得考试

1 AGGREGATION

1.1 PROBLEM APPROXIMATION - AGGREGATION

5 00:00:20,640 –> 00:00:25,230 so everything so far has been in the

6 00:00:25,230 –> 00:00:26,970 context of projected equations now we're

7 00:00:26,970 –> 00:00:28,830 going to discuss a different approach

8 00:00:28,830 –> 00:00:32,420 that seems to be fundamentally unrelated aggregation

关于投影方程我们已经讲了很多内容了，下面我要将一个不同的方法，他看起来与之前讲的都没有关系，也就是聚合

9 00:00:32,420 –> 00:00:35,070 however we will see that

10 00:00:35,070 –> 00:00:36,899 it's also connected with a projection equation approach

实际上我们一会就能看到聚合和投影方程是有联系的

11 00:00:36,899 –> 00:00:39,739 in aggregation the

12 00:00:39,739 –> 00:00:44,040 major idea is to approximate either the

13 00:00:44,040 –> 00:00:47,670 optimal cost function or the the cost

14 00:00:47,670 –> 00:00:50,160 function associated with policies with a

15 00:00:50,160 –> 00:00:53,879 cost go function of a simpler problem

聚合方法主要的想法是近似更简单的问题的最优成本函数或者策略相关的成本函数

16 00:00:53,879 –> 00:00:57,300 so we have a complicated problem we will

17 00:00:57,300 –> 00:00:59,100 try to look at a simpler version of this

18 00:00:59,100 –> 00:01:02,730 problem and solve that in place of the original

我们现在有一个复杂的问题，我们用一个更简单的视角来看这个问题，并用求解它来取代求解原始问题

19 00:01:02,730 –> 00:01:07,229 and the approach is aggregation

20 00:01:07,229 –> 00:01:11,760 and the main elements are

聚合方法的主要元素如下：

21 00:01:11,760 –> 00:01:16,350 we introduce a few aggregate states which are viewed as

22 00:01:16,350 –> 00:01:19,619 the states of an aggregate system our

23 00:01:19,619 –> 00:01:22,110 objective is to formulate an aggregate

24 00:01:22,110 –> 00:01:28,320 problem involving just a few states

我们介绍了一点聚合状态，状态聚合后会得到一个新系统，新系统的目标函数是使用更少的状态 (聚合状态) 表示的

25 00:01:28,320 –> 00:01:30,509 to do this we need to define transition

26 00:01:30,509 –> 00:01:32,310 probabilities and cost functions for

27 00:01:32,310 –> 00:01:35,790 this aggregate system

为了完成聚合我们需要定义新系统的状态转移概率与聚合系统的成本函数

28 00:01:35,790 –> 00:01:37,829 and in order to make the aggregate problem related to

29 00:01:37,829 –> 00:01:39,450 the original the transition

30 00:01:39,450 –> 00:01:41,759 probabilities and cost have to be

31 00:01:41,759 –> 00:01:44,700 related to the transition probabilities

32 00:01:44,700 –> 00:01:48,540 and costs of the of the original problem

为了让聚合系统与原系统相关，聚合系统的状态转移概率与成本需要与原系统的状态转移概率与成本相关 (或者说由原系统得到)

33 00:01:48,540 –> 00:01:51,600 and we have to relate original system

34 00:01:51,600 –> 00:01:55,110 states with aggregate States

我们必须将原系统与聚合系统联系起来

35 00:01:55,110 –> 00:01:57,420 after we do that we solve exactly or approximately

36 00:01:57,420 –> 00:02:00,860 the aggregate problem by any kind of

37 00:02:00,860 –> 00:02:02,750 value iteration or policy direction

38 00:02:02,750 –> 00:02:05,600 perhaps even with simulation based

39 00:02:05,600 –> 00:02:09,258 methods to obtain approximations of J star and J mu

联系起来之后我们需要使用任意形式的值迭代或者策略迭代甚至是基于仿真的方法来求聚合问题的精确解或者近似解来得到 J^* 和 J_μ

40 00:02:09,258 –> 00:02:15,650 and let's say that

41 00:02:15,650 –> 00:02:20,060 aggregate state Y has an optimal cost or

42 00:02:20,060 –> 00:02:23,060 cost of policy R hat Y so this is a

43 00:02:23,060 –> 00:02:26,570 number associated with aggregate state Y

44 00:02:26,570 –> 00:02:29,920 and there several of those Y

我们说聚合系统中状态 y 的最优成本或者某策略的成本 $\hat{R}(y)$ 是一个与聚合状态 y 相关的数

45 00:02:29,920 –> 00:02:33,560 we use the cost function approximation which is

46 00:02:33,560 –> 00:02:40,070 a weighted sum of this our hats

我们用所有聚合状态 y 的最优成本 $\hat{R}(y)$ 的加权和来近似状态 j 的成本

47 00:02:40,070 –> 00:02:43,280 so we have a big state space States J in them

48 00:02:43,280 –> 00:02:46,790 in it and we solve the aggregate problem

49 00:02:46,790 –> 00:02:48,500 a different problem that involves this

50 00:02:48,500 –> 00:02:52,010 few but sort of bigger states we get the

51 00:02:52,010 –> 00:02:57,100 cost of those states and then we use a

52 00:02:57,100 –> 00:03:01,970 weighted sum of these costs of the

53 00:03:01,970 –> 00:03:04,400 approximate of the aggregate problem to

54 00:03:04,400 –> 00:03:06,500 approximate the optimal cost of the original problem

我们有一个很大的状态空间，状态 j 在这个空间中，我们求解这个与原问题不同 (规模比原问题小) 的聚合问题得到这些状态的最优成本，然后对这些成本进行加权累加来近似原问题的最优成本

55 00:03:06,500 –> 00:03:08,180 now what are these

56 00:03:08,180 –> 00:03:10,220 weights here first of all their

57 00:03:10,220 –> 00:03:12,739 probabilities they add up to one okay

这些权重是什么，首先他们是概率，累加等于一

58 00:03:12,739 –> 00:03:14,540 and we call them the aggregation probabilities

我们叫他们聚合概率

59 00:03:14,540 –> 00:03:17,840 and basically they encode

60 00:03:17,840 –> 00:03:21,320 the degree of membership of J into the

61 00:03:21,320 –> 00:03:24,650 aggregate state y it's how J relates

62 00:03:24,650 –> 00:03:28,130 to y in some sense that this far here encodes

基本的，他们的编码是把 j 聚合进了状态 y，这就是 j 和 y 的联系

63 00:03:28,130 –> 00:03:32,480 now notice that you can also

64 00:03:32,480 –> 00:03:36,640 view this Phi terms as features okay

你可以把 ϕ 这项当作一个特征

65 00:03:36,640 –> 00:03:39,500 basically this is of the form Phi

66 00:03:39,500 –> 00:03:44,510 capital Phi R capital R so these are

67 00:03:44,510 –> 00:03:46,670 matrix Phi here of the aggregation

68 00:03:46,670 –> 00:03:50,120 probabilities weighted by a vector of

69 00:03:50,120 –> 00:03:52,940 R hats a smaller dimensional version

70 00:03:52,940 –> 00:03:54,140 vote for five hats

这种形式可以理解成一个矩阵 Φ 和一个向量 $\hat{R}$，聚合概率矩阵 Φ 被向量 $\hat{R}$ 加权求和这样就可以从一个比较小的维度出发进行求解了

71 00:03:54,140 –> 00:03:57,019 so this is a linear architecture with a

72 00:03:57,019 –> 00:03:59,980 few J y are the features of state J

这是一个线性结构，ϕ_{jy} 是状态 j 的特征

73 00:03:59,980 –> 00:04:03,110 one feature for aggregate state one one

74 00:04:03,110 –> 00:04:08,050 figure for market state two and so on

聚合状态 1 的特征，聚合状态 2 的特征，之类的

75 00:04:08,050 –> 00:04:10,550 so at a higher level we're not doing

76 00:04:10,550 –> 00:04:13,670 anything different if it is again

77 00:04:13,670 –> 00:04:17,029 a linear feature based architecture but

78 00:04:17,029 –> 00:04:19,608 the training method is different because

79 00:04:19,608 –> 00:04:21,079 now we're looking at the simplified

80 00:04:21,079 –> 00:04:28,370 problem to get these coefficients

从更高的角度来看，如果还使用线性结构，聚合其实和之前讲的方法没有任何区别，除了训练的方法有一点不同，因为我们想要求解一个更简单的问题来获得近似结构的系数

81 00:04:28,370 –> 00:04:30,710 so now let's look at an example and see how we
82 00:04:30,710 –> 00:04:32,780 can get this aggregate problem how we can define it
我们来看一个例子，如何定义能得到一个聚合问题

1.2 HARD AGGREGATION EXAMPLE

83 00:04:32,780 –> 00:04:37,760 the simplest example is hard aggregation
最简单的例子是硬聚合
84 00:04:37,760 –> 00:04:42,620 in hard aggregation we
85 00:04:42,620 –> 00:04:46,130 have this big original system and we
86 00:04:46,130 –> 00:04:48,229 group various states of the system together
硬聚合中我们有一个非常大规模得原问题，把原问题的状态分组
87 00:04:48,229 –> 00:04:52,430 as an example let's say we have
88 00:04:52,430 –> 00:04:55,070 nine states of the original system we
89 00:04:55,070 –> 00:04:57,139 group this four together into a single
90 00:04:57,139 –> 00:05:00,410 aggregate state here we group to here 1 2 and so on
这是一个例子，原系统有九个状态，我们把这四个分到一组里，然后这两个一组，这两个一组，这一个一组
91 00:05:00,410 –> 00:05:05,600 so we have nine original
92 00:05:05,600 –> 00:05:09,340 system States and four aggregate States
所以这是一个九个状态的原系统和四个状态的聚合系统
93 00:05:09,340 –> 00:05:12,620 and the aggregation probabilities in
94 00:05:12,620 –> 00:05:14,770 hard aggregation is what you see here
硬聚合的聚合概率在这里
95 00:05:14,770 –> 00:05:17,419 for state 1 these are the aggregation probabilities
这是状态 1 的聚合概率
96 00:05:17,419 –> 00:05:21,919 state 1 is going to be
97 00:05:21,919 –> 00:05:23,870 related to the aggregate cost of
98 00:05:23,870 –> 00:05:31,700 aggregate state 1 exactly and so on
原状态 1 与聚合状态 1 的聚合成本相关
99 00:05:31,700 –> 00:05:33,620 so if I solve the aggregate problem somehow
100 00:05:33,620 –> 00:05:35,660 I define this aggregate problem I get
101 00:05:35,660 –> 00:05:38,900 four numbers ok the cost of aggregate
102 00:05:38,900 –> 00:05:41,240 state one cost at least eight two three and four
如果我求解这个聚合问题，可以得到四个聚合状态的聚合成本
103 00:05:41,240 –> 00:05:47,000 and my original system cost
104 00:05:47,000 –> 00:05:50,200 function is a piecewise constant
105 00:05:50,200 –> 00:05:53,539 function where all the states in here
106 00:05:53,539 –> 00:05:56,390 have the same cost the aggregate cost of
107 00:05:56,390 –> 00:05:58,580 a state all the states in here the two
108 00:05:58,580 –> 00:06:00,650 states have the idea of the state here and so on
原始问题的成本函数是一个分段常值函数，聚合状态 1 的四个原始状态的成本等于聚合状态 1 的聚合成本，其他聚合成本以此类推
109 00:06:00,650 –> 00:06:04,190 and this is what this phi
110 00:06:04,190 –> 00:06:08,570 matrix implies the the the features are
111 00:06:08,570 –> 00:06:12,110 lined up like this the columns are the
112 00:06:12,110 –> 00:06:14,919 basis functions
矩阵 ϕ 中，行表示特征，列表示基函数
113 00:06:16,220 –> 00:06:18,780 so after I solve the aggregate problem
114 00:06:18,780 –> 00:06:20,720 that's the way I'm going to recover
115 00:06:20,720 –> 00:06:22,620 approximations to the cost of the original problem
所以我求解聚合问题之后，就可以用结果来恢复原问题的近似成本
116 00:06:22,620 –> 00:06:26,460 however how do i define subproblems
但是我该如何定义子问题 (聚合问题) 呢
117 00:06:26,460 –> 00:06:28,410 in particular i have to
118 00:06:28,410 –> 00:06:31,200 define transition probabilities from X 1
119 00:06:31,200 –> 00:06:35,580 to X 2 to X 4 2 X 3 what makes sense here
我们必须定义状态转移概率

120 00:06:35,580 –> 00:06:38,820 well to define such transition
121 00:06:38,820 –> 00:06:40,590 probabilities basically we have to
122 00:06:40,590 –> 00:06:42,570 relate them to the transition
123 00:06:42,570 –> 00:06:44,220 probabilities between these states and
124 00:06:44,220 –> 00:06:46,110 these states sort of some kind of an
125 00:06:46,110 –> 00:06:48,750 aggregate probability
定义状态转移概率首先必须把聚合状态之间的联系弄清楚
126 00:06:48,750 –> 00:06:50,610 and the way to do that so that's the question what should
127 00:06:50,610 –> 00:06:51,840 be the aggregate transition probabilities of X
这样做的时候就会遇到一个问题，聚合状态 x 的转移概率是多少呢
128 00:06:51,840 –> 00:06:56,550 we may select at
129 00:06:56,550 –> 00:07:00,510 random some state here look at how
130 00:07:00,510 –> 00:07:02,430 transitions whether it transitions
131 00:07:02,430 –> 00:07:05,040 within the same aggregate state what
132 00:07:05,040 –> 00:07:06,690 whether it transitions out of that
133 00:07:06,690 –> 00:07:12,030 aggregate state and define define view
134 00:07:12,030 –> 00:07:13,830 this state single state as
135 00:07:13,830 –> 00:07:17,100 representative of of the aggregate state
136 00:07:17,100 –> 00:07:20,070 and in the defined range problems that way
我们在聚合状态 x_1 中随机选一个状态，看他转移之后留在这个聚合状态和离开这个聚合状态的概率，然后把这个状态的概率代表这个聚合状态的概率，就这样定义所有的聚合状态的概率
137 00:07:20,070 –> 00:07:24,230 but which states should I use okay
但是我们该用哪一个状态呢
138 00:07:24,230 –> 00:07:27,450 the simplest probability is to randomize
最简单的概率计算方式是随机选
139 00:07:27,450 –> 00:07:30,090 and to say I'm going to select us when I
140 00:07:30,090 –> 00:07:32,010 am in aggregate state that's one I'm
141 00:07:32,010 –> 00:07:34,890 going to take by at random any one state
142 00:07:34,890 –> 00:07:36,750 of the state's here look at the
143 00:07:36,750 –> 00:07:38,820 transition probabilities and consider
144 00:07:38,820 –> 00:07:42,930 the corresponding to could be provided
145 00:07:42,930 –> 00:07:46,620 probabilistic way describing how from
146 00:07:46,620 –> 00:07:48,900 this state I'll go back into the state or that state
我现在处于聚合状态时，随机选一个状态，然后看他留在这个状态和离开这个状态的概率，用这个概率来描述聚合状态的转移概率就可以了
147 00:07:48,900 –> 00:07:52,590 so the simplest
148 00:07:52,590 –> 00:07:54,600 probability is to assume that all the
149 00:07:54,600 –> 00:07:57,510 states in within the aggregate state X are
150 00:07:57,510 –> 00:08:00,050 equally likely
所以最简单的概率是假设某一个聚合状态中的原始状态的概率是相同的
151 00:08:00,110 –> 00:08:02,600 however I can consider more general randomization
现在我要考虑一个更一般性的随机方式
152 00:08:02,600 –> 00:08:07,610 whereby I use
153 00:08:07,610 –> 00:08:11,100 disaggregation probabilities for
154 00:08:11,100 –> 00:08:16,620 aggregate state X D X I is a probability
155 00:08:16,620 –> 00:08:19,830 by which I select I within here to
156 00:08:19,830 –> 00:08:21,870 define the transition probability
157 00:08:21,870 –> 00:08:23,850 mechanism to aggregation
158 00:08:23,850 –> 00:08:24,880 States
用分解概率来表示聚合概率，有聚合状态 x，d_{xi} 是在状态 x 中选择 i 的概率，用这种方式来定义聚合状态的状态转移概率
159 00:08:24,880 –> 00:08:29,690 roughly speaking DXi is the degree to
160 00:08:29,690 –> 00:08:34,789 which I is representative of X
简单地说，d_{xi} 是使用 i 表示 x 的概率
161 00:08:34,789 –> 00:08:37,760 okay if there was a dominant state in here I'd
162 00:08:37,760 –> 00:08:40,309 like to use that state more more
163 00:08:40,309 –> 00:08:42,049 generally if there are more states that
164 00:08:42,049 –> 00:08:44,210 are equal that are important I want to

165 00:08:44,210 –> 00:08:46,100 give positive probabilities to goals as
166 00:08:46,100 –> 00:08:48,400 well
如果这个聚合状态有一个主导的原状态，我想要用这个主导的原状态来表示聚合状态，如果有更多相同的主导状态，我想要给他们正概率来达到目标 (用一个原始状态表示聚合状态)
167 00:08:54,060 –> 00:08:57,090 so if I have this DX i I solve the
168 00:08:57,090 –> 00:08:59,430 problem of transition probabilities
169 00:08:59,430 –> 00:09:02,520 between x1 and x2 a randomized within
170 00:09:02,520 –> 00:09:05,400 here consider the transition probability
171 00:09:05,400 –> 00:09:08,460 from any one of those and see what part
172 00:09:08,460 –> 00:09:10,020 of the transition probability takes me
173 00:09:10,020 –> 00:09:14,960 to states in the aggregate state two
如果我有了 d_{xi}，我要计算聚合状态 x_1 到聚合状态 x_2 的转移概率。从聚合状态 x_1 中随机选出一个原始状态，考虑聚合状态 x_1 中任意一个原始状态转移到聚合状态 x_2 中原始状态的概率
174 00:09:15,890 –> 00:09:18,330 intuitively that makes sense for hard aggregation
175 00:09:18,330 –> 00:09:21,270 and it is this mekinese that
176 00:09:21,270 –> 00:09:24,690 were going to generalize
直观上这种方法对于硬聚合是有效的，下面我要做一个总结

1.3 AGGREGATION/DISAGGREGATION PROBS

177 00:09:24,690 –> 00:09:28,160 so here's the general framework for aggregation
这就是聚合的通用框架
178 00:09:28,160 –> 00:09:30,840 the original system is described by states I
179 00:09:30,840 –> 00:09:33,420 and J with transition probabilities P IJ
180 00:09:33,420 –> 00:09:37,050 that depends on control
原系统被状态 i 和 j 还有依赖于控制的状态转移概率 $p_{ij}(u)$ 描述
181 00:09:37,050 –> 00:09:39,930 we introduced this aggregate stage we want to define
182 00:09:39,930 –> 00:09:44,010 the transition mechanism from X to Y
这是聚合状态，定义从聚合状态 x 到聚合状态 y 的转移机制
183 00:09:44,010 –> 00:09:47,310 we introduce now these two transition these
184 00:09:47,310 –> 00:09:49,130 two probability types the disaggregation
185 00:09:49,130 –> 00:09:52,650 probabilities defined by a matrix D with
186 00:09:52,650 –> 00:09:55,490 elements d X I and this aggregation
187 00:09:55,490 –> 00:09:58,440 probabilities which are defined by this matrix Phi
我来介绍一下这两个转换，这个分解概率被矩阵 d_{xi} 定义，这个聚合概率被矩阵 ϕ 定义
188 00:09:58,440 –> 00:10:04,800 and the matrices D and Phi
189 00:10:04,800 –> 00:10:08,610 are arbitrary except for the factor that
190 00:10:08,610 –> 00:10:11,010 their rows have to be transition probabilities
这两个矩阵 d 和 ϕ 是任意行表示转移概率的矩阵
191 00:10:11,010 –> 00:10:17,310 these are the aggregation
192 00:10:17,310 –> 00:10:21,660 probabilities and the meaning of Phi J Y
193 00:10:21,660 –> 00:10:25,860 is the degree of membership of J within
194 00:10:25,860 –> 00:10:29,250 the aggregate state Y
这些是聚合概率，ϕ_{jy} 表示聚合状态 y 下原状态是 j 的概率
195 00:10:29,250 –> 00:10:31,950 so if J is sort of relates to more than one aggregate
196 00:10:31,950 –> 00:10:34,400 States then I will have more than one
197 00:10:34,400 –> 00:10:38,310 nonzero probability spy
所以如果 j 与不止一个聚合状态有关，那么 ξ 就会有超过一个非零概率
198 00:10:38,310 –> 00:10:42,630 in hard aggregation only one ϕ is a one and all the others are 0
硬聚合中，每列只有一个元素是 1，剩下的都是 0
199 00:10:42,630 –> 00:10:45,540 more generally however
200 00:10:45,540 –> 00:10:48,450 you could have a fine matrix involving a
201 00:10:48,450 –> 00:10:53,880 different kind of choice
更一般地，你会有一个包括很多不同选择的矩阵
202 00:10:53,880 –> 00:10:57,410 now the disaggregation probability
203 00:10:58,510 –> 00:11:02,780 given that we are at state X it
204 00:11:02,780 –> 00:11:06,790 randomizes within this piece state of

205 00:11:06,790 –> 00:11:09,260 states I according to these
206 00:11:09,260 –> 00:11:12,800 probabilities selects the state this way
207 00:11:12,800 –> 00:11:16,160 in here and then jumps off to J and then down to Y
分解概率表示现在我们在聚合状态 x，在这些状态中随机选一个状态 i，然后按照状态转移概率(原状态转移概率) 跳转到状态 j，再映射到聚合状态 y
208 00:11:16,160 –> 00:11:20,330 and the meaning intuitive
209 00:11:20,330 –> 00:11:22,610 meaning of this reaction probability is
210 00:11:22,610 –> 00:11:28,850 a degree which is representative of X
直观地解释，这个概率是状态 i 代表状态 x 的概率
211 00:11:28,850 –> 00:11:31,070 so once I give you these two matrixes D and
212 00:11:31,070 –> 00:11:33,940 phi you have an aggregation scheme
一旦我给顶了矩阵 d 和 ϕ，你就得到了一个聚合方案
213 00:11:33,940 –> 00:11:36,830 different choices of D and Phi gives you
214 00:11:36,830 –> 00:11:38,840 different types of aggregation schemes
215 00:11:38,840 –> 00:11:41,900 but the theory of all those is quite
216 00:11:41,900 –> 00:11:44,170 calm
d 和 ϕ 不同的选择给出不同的聚合方案，但是这些理论都很**冷静**

1.4 AGGREGATE SYSTEM DESCRIPTION

217 00:11:47,990 –> 00:11:51,149 okay now suppose I have given you the
218 00:11:51,149 –> 00:11:52,560 disaggregation probabilities in the
219 00:11:52,560 –> 00:11:54,899 aggregation probabilities you can define
220 00:11:54,899 –> 00:11:56,700 the transition probabilities between x
221 00:11:56,700 –> 00:12:00,060 and y the transition probability from X
222 00:12:00,060 –> 00:12:03,029 to Y is obtained as the product of the
223 00:12:03,029 –> 00:12:06,660 probabilities of going to I then P IJ to
224 00:12:06,660 –> 00:12:10,170 go to J and then down to Y according to
225 00:12:10,170 –> 00:12:12,600 this probability
假定给定了聚合概率和分解概率，我们可以定义 x 和 y 之间的转移概率如下：x 以概率分解到 i，i 以概率 $p_{ij}(u)$) 转移到 j，然后根据概率 ϕ_{jy} 映射到聚合状态 y
226 00:12:12,600 –> 00:12:14,430 so that's the transition probability and it's well
227 00:12:14,430 –> 00:12:16,050 defined it's a transition probability matrix
这就是状态转移概率矩阵
228 00:12:16,050 –> 00:12:20,940 in fact it is given by it's given
229 00:12:20,940 –> 00:12:23,579 by this expression and in matrix form
230 00:12:23,579 –> 00:12:28,490 it's given like so
事实上这个概率被这个表达式定义，如果用矩阵形式定义的话，是这样的
231 00:12:28,490 –> 00:12:31,620 the matrix D multiplies the matrix P of the original
232 00:12:31,620 –> 00:12:34,410 system and multiplies the matrix phi of
233 00:12:34,410 –> 00:12:38,190 the disaggregation matrix phi
矩阵 D 乘以原系统的转移概率矩阵矩阵 P，再乘以分解矩阵 ϕ
234 00:12:38,190 –> 00:12:40,529 and that's the compact form of the transition
235 00:12:40,529 –> 00:12:42,510 probability matrix of the aggregate
236 00:12:42,510 –> 00:12:45,209 problems then it depends on u because
237 00:12:45,209 –> 00:12:51,750 this p IJ h depends on u
这是聚合问题状态转移矩阵的简写形式，这个转移矩阵依赖于控制 u，因为原问题的状态转移矩阵依赖于控制 u
238 00:12:51,750 –> 00:12:55,709 this process of going like this defines also costs
这个过程的成本函数是这样定义的
239 00:12:55,709 –> 00:12:59,730 the cost of using control u at aggregate
240 00:12:59,730 –> 00:13:03,510 State X is obtained by randomization to
241 00:13:03,510 –> 00:13:08,100 go into here then cost GI u of j to go
242 00:13:08,100 –> 00:13:12,120 from here to here and then and then in
243 00:13:12,120 –> 00:13:13,890 that and then that's that that's the
244 00:13:13,890 –> 00:13:15,810 cost that u incur for a transition out of X

在聚合状态 x 执行控制 u 的成本包括 x 与 i 的对应关系，然后是原状态 i 下执行控制 u 由概率转移到 j 产生的成本，聚合系统的新状态是 j 根据聚合概率转移到的聚合状态 y，这就是聚合状态 x 执行控制 u 产生的成本

245 00:13:15,810 –> 00:13:20,279 so the expected transition cost is this

这就是一次转移的期望成本

246 00:13:20,279 –> 00:13:27,510 and in matrix form it is d times P

247 00:13:27,510 –> 00:13:31,470 actually piece of u here times G G

248 00:13:31,470 –> 00:13:36,510 being the cost of the one stage costs so

249 00:13:36,510 –> 00:13:39,649 the vector of one stage cost

这是他的矩阵形式，D 乘以控制 u 下的 P，再乘以 g，g 是一次转移成本，这个成本是一个向量

250 00:13:40,120 –> 00:13:42,580 so if you have a fixed policy for the

251 00:13:42,580 –> 00:13:44,860 aggregate problem you can set up bellman

252 00:13:44,860 –> 00:13:48,240 equation for it involving this(P 的矩阵形式) in that(g 的矩阵形式)

如果你有了一个聚合问题的不动策略，你可以根据 P 和 g 写出 bellman 方程

253 00:13:48,240 –> 00:13:51,520 and you can also consider the optimal

254 00:13:51,520 –> 00:13:53,260 cost of the aggregate problem which

255 00:13:53,260 –> 00:13:57,600 involves the corresponding operator T

你同样可以考虑聚合问题相关算子 T 的最优成本

256 00:13:59,640 –> 00:14:03,400 here is balance equation for the

257 00:14:03,400 –> 00:14:06,130 aggregate problem it is the unique

258 00:14:06,130 –> 00:14:14,230 solution R hat of this equation which

259 00:14:14,230 –> 00:14:18,280 involves G hat the one stage cost of the

260 00:14:18,280 –> 00:14:20,560 aggregate problem and the transition

261 00:14:20,560 –> 00:14:23,610 probabilities of the aggregate problem

262 00:14:23,610 –> 00:14:27,610 this R hat is a vector with one entry

263 00:14:27,610 –> 00:14:31,660 per aggregate state ordinarily a small

264 00:14:31,660 –> 00:14:33,760 dimensional vector if a number of

265 00:14:33,760 –> 00:14:41,530 aggregate states is small

这是一个聚合问题的 bellman 方程，这个是这个方程的唯一解 $\hat{R}$，方程中包括聚合问题的一阶段成本与转移概率，这个 $\hat{R}$ 是一个包括所有聚合状态的向量，如果聚合状态很少的话，这个向量维度也会很少

266 00:14:41,530 –> 00:14:44,380 suppose that I solve this equation and I solve the

267 00:14:44,380 –> 00:14:46,170 aggregate problem for the optimal cost

268 00:14:46,170 –> 00:14:49,390 then I can get an approximation of the

269 00:14:49,390 –> 00:14:52,750 optimal cost of the original by using the phi of features

假设我求解了这个聚合问题的方程，得到了最优成本，然后我就可以用这个最优成本向量和特征矩阵 Φ 对原问题的最优成本进行近似

270 00:14:52,750 –> 00:14:55,630 in these parts are as weights

这个向量 R 就是权重

271 00:14:55,630 –> 00:15:00,310 there is also a bellman equation

272 00:15:00,310 –> 00:15:04,870 associated with a policy

这也是一个关于某策略的 bellman 方程

273 00:15:04,870 –> 00:15:06,850 it's linear if you can solve it then you obtain an

274 00:15:06,850 –> 00:15:09,790 approximation to the cost of a policy by

275 00:15:09,790 –> 00:15:13,240 again a linear combination of the course

276 00:15:13,240 –> 00:15:18,490 of the corresponding policy we aggregate

277 00:15:18,490 –> 00:15:20,910 problem

这是一个线性方程组，如果你求解之后就可以使用一个线性组合的策略成本进行近似了

278 00:15:21,700 –> 00:15:26,440 okay so this is in principle the idea

279 00:15:26,440 –> 00:15:30,560 simplify the problem define the data of

280 00:15:30,560 –> 00:15:33,410 the problem based on the data of the

281 00:15:33,410 –> 00:15:36,170 original problem solve it solve the

282 00:15:36,170 –> 00:15:38,510 simpler problem and then obtain an

283 00:15:38,510 –> 00:15:40,600 approximation to the cost of the

284 00:15:40,600 –> 00:15:43,670 original problem by means of a linear combination

这个方法的原则是把问题简化，根据原问题的数据定义简化问题的数据，然后求解这个比较简单的问题并且用求得的解的线性组合来近似原问题的成本

1.5 EXAMPLE I: HARD AGGREGATION

285 00:15:43,670 –> 00:15:50,870 now let's go back and look
286 00:15:50,870 –> 00:15:54,890 at various special cases
现在我们回头看一下这个特殊的例子
287 00:15:54,890 –> 00:15:57,200 we know that the case of hard aggregation earlier and
288 00:15:57,200 –> 00:15:59,780 here we group the original system states
289 00:15:59,780 –> 00:16:03,410 to subsets view each subset has an
290 00:16:03,410 –> 00:16:07,010 aggregate state this partition is
291 00:16:07,010 –> 00:16:09,440 exhaustive it covers all the states each
292 00:16:09,440 –> 00:16:12,580 state has to belong to one and only one
293 00:16:12,580 –> 00:16:15,620 aggregate state the aggregation
294 00:16:15,620 –> 00:16:18,230 probabilities are either ones or zeros
295 00:16:18,230 –> 00:16:20,600 depending on the membership of original
296 00:16:20,600 –> 00:16:23,270 system States into aggregation States again
297 00:16:23,270 –> 00:16:26,660 the same figure I had before
我们之前讲过这个聚合问题，我们把原问题的状态分组，每一个子集合都对应一个聚合状态，这个划分很详细，它包括了所有状态而且每一个状态只对应一个聚合状态，也就是说从原状态到聚合状态的聚合概率不是 1 就是 0，还用我们之前用过的那个图来表示
298 00:16:26,660 –> 00:16:28,580 now for this aggregation probabilities there are
299 00:16:28,580 –> 00:16:32,300 many possibilities for example all the
300 00:16:32,300 –> 00:16:34,340 states within a aggregation state having equal probability
这个聚合概率有很多可能的情况，比如所有状态到聚合状态的概率相等
301 00:16:34,340 –> 00:16:37,280 so if there are let's
302 00:16:37,280 –> 00:16:41,300 say M states within navigate state this
303 00:16:41,300 –> 00:16:43,640 probability would be 1 over m in this
304 00:16:43,640 –> 00:16:47,600 particular case with 1 over 4 1 half 1 1 half
如果某一组有 M 个原状态，我们说每一个状态的聚合概率是 $\frac{1}{m}$，这个图中的概率就是 $\frac{1}{4}$，$\frac{1}{2}$，1 和 $\frac{1}{2}$
305 00:16:47,600 –> 00:16:55,270 however I may use more general
306 00:16:55,270 –> 00:17:00,530 probabilities as long as as long as they
307 00:17:00,530 –> 00:17:02,600 are positive only within the
308 00:17:02,600 –> 00:17:08,599 corresponding aggregation States
我可能会用更一般的概率，只要他们是正数并且只与相应的聚合状态有关就行
309 00:17:08,599 –> 00:17:10,369 ok now I know that earlier that because of the
310 00:17:10,369 –> 00:17:12,079 nature of these of these phi matrix
311 00:17:12,079 –> 00:17:14,630 what you get in the end is a piecewise
312 00:17:14,630 –> 00:17:17,420 constant approximation over the
313 00:17:17,420 –> 00:17:19,459 aggregate States so there's a single
314 00:17:19,459 –> 00:17:21,589 value of course associated with each
315 00:17:21,589 –> 00:17:24,680 with all the states within here all the
316 00:17:24,680 –> 00:17:26,150 states within here all the states within
317 00:17:26,150 –> 00:17:26,599 here
318 00:17:26,599 –> 00:17:28,920 it's a piecewise constant approximation
319 00:17:28,920 –> 00:17:31,530 in particular if the optimal cost factor
320 00:17:31,530 –> 00:17:34,410 of the original problem is piecewise
321 00:17:34,410 –> 00:17:37,290 constant over the aggregate states then
322 00:17:37,290 –> 00:17:40,560 hard aggregation is exact okay there is
323 00:17:40,560 –> 00:17:43,440 no approximation error
这点我们之前就知道，由于这个矩阵 ϕ 的形式，原问题的最优成本是所有聚合状态的分段常数，硬聚合的近似就没有近似误差了
324 00:17:43,440 –> 00:17:45,570 and this suggests how I should be grouping States into
325 00:17:45,570 –> 00:17:48,720 aggregate states similar states with
326 00:17:48,720 –> 00:17:51,480 similar costs should go together okay it
327 00:17:51,480 –> 00:17:53,940 should be grouped together
这就促使我们把成本大小相近的状态分到同一组
328 00:17:53,940 –> 00:17:56,790 if I have some insight about what are similar
329 00:17:56,790 –> 00:17:59,130 States I would be able to group them
330 00:17:59,130 –> 00:18:06,000 together into aggregate States

如果我们关注相似的状态，我们就可以把他们凑成同一个聚合状态

331 00:18:06,000 –> 00:18:08,040 there is also a variant of the scheme called soft

332 00:18:08,040 –> 00:18:11,550 aggregation where the phi entries within

333 00:18:11,550 –> 00:18:14,160 this matrix are not all of them 1 or 0

334 00:18:14,160 –> 00:18:16,680 but for states that are sort of at the

335 00:18:16,680 –> 00:18:19,590 boundary between to to aggregate states

336 00:18:19,590 –> 00:18:22,950 it may be it may involve a randomization

337 00:18:22,950 –> 00:18:26,010 between the the aggregate States and the boundary

同样有另一个叫做软聚合的方案，软聚合的矩阵 ϕ 不全是由 1 和 0 构成的，在聚合状态之间的边界可能具有随机性

338 00:18:26,010 –> 00:18:29,640 so for example for state 5 this

339 00:18:29,640 –> 00:18:35,070 may involve a 1/2 1/2 here ok

比如对于状态 5 的概率向量可能是 $\frac{1}{2}$ 和 $\frac{1}{2}$

340 00:18:35,070 –> 00:18:37,590 or there may be there may be no other entries

341 00:18:37,590 –> 00:18:39,450 here nonzero entries in place of the zeros

或者这一个向量全都是非零数

342 00:18:39,450 –> 00:18:41,850 this is called soft aggregation

343 00:18:41,850 –> 00:18:44,700 provides soft boundaries between the aggregate States

这就叫软聚合，在聚合状态之间提供软边界

344 00:18:44,700 –> 00:18:47,010 and it's a it's

345 00:18:47,010 –> 00:18:49,830 something that that's that that that has

346 00:18:49,830 –> 00:18:52,620 been used and I'm just mentioning it

347 00:18:52,620 –> 00:18:54,990 just to give you an idea of a variety of possibilities

这种方法之前被用过，我在这里提到它只是为了告诉你还有很多其他类型的聚合方法

348 00:18:54,990 –> 00:19:00,950 so that hard aggregation

349 00:19:00,950 –> 00:19:05,690 each state belonging to an aggregate

350 00:19:05,690 –> 00:19:10,860 state but there's the question of how do

351 00:19:10,860 –> 00:19:13,500 I pick the boundaries of these states

352 00:19:13,500 –> 00:19:17,720 I should be grouping states with similar

353 00:19:17,720 –> 00:19:21,810 cost into aggregates but how do I decide

354 00:19:21,810 –> 00:19:25,260 that states have similar cost well I may

355 00:19:25,260 –> 00:19:28,200 have good features of the states and

356 00:19:28,200 –> 00:19:30,600 therefore I may make the supposition

357 00:19:30,600 –> 00:19:33,330 that states with similar features

358 00:19:33,330 –> 00:19:34,690 features have

359 00:19:34,690 –> 00:19:37,030 similar costs and should be grouped

360 00:19:37,030 –> 00:19:40,960 together this is called feature based our aggregation

这就是硬聚合，每一个状态都属于一个聚合状态，但是有一个问题是我该如何选择这些状态的边界。我可以把成本差不多的状态聚到一起，但是我怎么知道哪些状态的成本差不多呢，我可以选择比较好的状态特征，然后做出一个假设，特征相似的状态成本也相似，可以聚成一个状态，这种方法被叫做基于特征的聚合

1.6 EXAMPLE II: FEATURE-BASED AGGREGATION

361 00:19:40,960 –> 00:19:45,070 so questions how do we

362 00:19:45,070 –> 00:19:46,780 group States together into aggregate

363 00:19:46,780 –> 00:19:49,360 states so if we know good features then

364 00:19:49,360 –> 00:19:51,790 it makes sense to group together states

365 00:19:51,790 –> 00:19:55,780 that have similar features

问题是我们该如何聚合状态，如果我们知道比较好的特征，就可以根据这个特征来对相似的状态进行分组

366 00:19:55,780 –> 00:19:58,510 now this involves a partition that's based on the

367 00:19:58,510 –> 00:20:00,880 space that's based on a partition space of features

这是一个基于特征的划分空间

368 00:20:00,880 –> 00:20:06,700 from States I get features

369 00:20:06,700 –> 00:20:09,600 by a by a feature extraction mapping

我通过一个提取映射得到状态的特征

370 00:20:09,600 –> 00:20:12,820 suppose that I discretize the space of

371 00:20:12,820 –> 00:20:16,060 feature more or less evenly
假设我把状态的特征空间离散化
372 00:20:16,060 –> 00:20:18,750 and then I consider a partition of the state space
373 00:20:18,750 –> 00:20:22,090 according to which States map into which feature
然后我考虑状态与特征的对应关系来进行状态空间的划分
374 00:20:22,090 –> 00:20:24,850 so instead of bothering to
375 00:20:24,850 –> 00:20:27,010 discretize this space here I despise
376 00:20:27,010 –> 00:20:28,390 this space here that may be more convenient
这样可以避免把状态空间离散化然后聚集状态，用更方便的方式，直接离散化特征空间
377 00:20:28,390 –> 00:20:32,500 each one of these boxes
378 00:20:32,500 –> 00:20:34,720 within the feature space may be mapped
379 00:20:34,720 –> 00:20:37,600 into a unique aggregate state and that
380 00:20:37,600 –> 00:20:39,760 induces a partition of the set of original States
特征空间的每一个格子都对应一个唯一的聚合状态，对应原始状态空间的一个划分
381 00:20:39,760 –> 00:20:42,790 solve an idea that has
382 00:20:42,790 –> 00:20:44,560 been used many times it's an idea that
383 00:20:44,560 –> 00:20:49,300 makes sense if you have good features
如果你有比较好的特征提取方法的话，这种方法很有用并且被使用了很多次
384 00:20:49,300 –> 00:20:50,980 so it's a general approach for passing from
385 00:20:50,980 –> 00:20:53,380 a feature based State representation to
386 00:20:53,380 –> 00:20:56,670 a hard aggregation based representation
这是一种非常一般的方法，从基于特征的状态表达到基于硬约束的表达
387 00:20:56,670 –> 00:20:59,590 and essentially we discretize the space
388 00:20:59,590 –> 00:21:03,280 of features and we generate and the
389 00:21:03,280 –> 00:21:07,030 aggregate problem involves constant
390 00:21:07,030 –> 00:21:09,280 aprox bitwise constant approximation
391 00:21:09,280 –> 00:21:13,180 within each one of these boxes
实际上我们把特征空间离散化，然后对每一个格子生成聚合状态并进行近似
392 00:21:13,180 –> 00:21:15,550 states that have similar features in that group
393 00:21:15,550 –> 00:21:21,790 together obtain as the same cost and in
394 00:21:21,790 –> 00:21:24,850 the end I have a piecewise constant
395 00:21:24,850 –> 00:21:27,780 approximation
有相似特征的状态由于成本也相似被分成一组，近似后就得到了一个分段常数近似
396 00:21:31,470 –> 00:21:34,960 we discussed earlier linear feature
397 00:21:34,960 –> 00:21:37,750 based approximations given a set of
398 00:21:37,750 –> 00:21:40,150 features for a state weigh them linearly
399 00:21:40,150 –> 00:21:44,680 with a weight vector this is not linear
400 00:21:44,680 –> 00:21:48,040 this is piecewise constant approximation
401 00:21:48,040 –> 00:21:52,390 so it's nonlinear
我们之前讨论过线性特征近似，给定一个特征集合后通过权重向量进行线性近似，但是聚合方法不是线性的，而是分段常数近似
402 00:21:52,390 –> 00:21:54,550 it's a much more powerful architecture potentially than a
403 00:21:54,550 –> 00:21:56,800 linear architecture because it is
404 00:21:56,800 –> 00:21:59,650 nonlinear in the features but also it
405 00:21:59,650 –> 00:22:03,640 may require a lot of discretization here
406 00:22:03,640 –> 00:22:07,870 and a lot of aggregate states to to
407 00:22:07,870 –> 00:22:09,760 reach the same level of performance as
408 00:22:09,760 –> 00:22:13,870 the level of performance of linear
409 00:22:13,870 –> 00:22:17,890 feature based architectures
这是一种比线性结构更有效的结构，因为他是非线性的，但是这种方法需要把状态空间离散化并生成很多聚合状态来达到与线性特征结构相同等级的表现
410 00:22:17,890 –> 00:22:19,540 it's a different way to use features in
411 00:22:19,540 –> 00:22:22,140 approximation
这是一种不同的使用特征近似的方法
412 00:22:26,410 –> 00:22:29,070 this is still a hard aggregation method
413 00:22:29,070 –> 00:22:33,940 because every state participate into some aggregate state
这仍然是一种硬聚合，因为每一一个状态都被划分到相应的聚合状态中

1.7 EXAMPLE III: REP. STATES/COARSE GRID

414 00:22:33,940 –> 00:22:37,570 now let's look at

415 00:22:37,570 –> 00:22:39,370 something different which also has a long tradition

我们来看另一个很传统的例子

416 00:22:39,370 –> 00:22:46,360 okay here's the original

417 00:22:46,360 –> 00:22:48,310 state space in fact it could be

418 00:22:48,310 –> 00:22:52,420 continuous has many many states

这是原问题的连续状态空间，有很多状态

419 00:22:52,420 –> 00:22:55,480 I look at a small subset of states which I call representative

我要看这个具有代表性的小的状态子集合

420 00:22:55,480 –> 00:23:01,230 this black circles are representative states

这些黑色圆圈是有代表性的状态

421 00:23:01,230 –> 00:23:05,380 and I associate

422 00:23:05,380 –> 00:23:08,710 with each one of these states an aggregate state

我把这些状态叫做聚合状态

423 00:23:08,710 –> 00:23:14,520 the these aggregation

424 00:23:14,520 –> 00:23:19,060 probabilities are 1 for 0 for the state

425 00:23:19,060 –> 00:23:20,800 the original system state that happens

426 00:23:20,800 –> 00:23:23,740 to be also a representative it is 1 and

427 00:23:23,740 –> 00:23:28,510 all the others are 0

这些聚合状态的概率是 1 或者 0，对于原系统的状态来说，如果它是一个表现状态，值是 1，否则是 0

428 00:23:28,510 –> 00:23:30,460 so so here that's the these aggregation problems are very simply defined

这些聚合问题的定义非常简单

429 00:23:30,460 –> 00:23:33,790 the aggregation

430 00:23:33,790 –> 00:23:36,070 probabilities which correspond to states

431 00:23:36,070 –> 00:23:39,400 that are not representative involve some

432 00:23:39,400 –> 00:23:42,360 kind of randomization to the representative States

与非代表性状态相关的聚合概率包括他们与代表性状态的随机性

433 00:23:42,360 –> 00:23:49,810 so for every state

434 00:23:49,810 –> 00:23:53,310 of the original system there are these

435 00:23:53,310 –> 00:23:56,950 probabilities that relate them the

436 00:23:56,950 –> 00:24:00,640 representative States

所以对于原系统的每一个状态，概率都与代表性状态相关

437 00:24:00,640 –> 00:24:03,340 so for example you might have a grid discretization of the

438 00:24:03,340 –> 00:24:05,980 original space and four points that fall

439 00:24:05,980 –> 00:24:09,010 outside the grid you have a

440 00:24:09,010 –> 00:24:11,530 probabilistic mechanism to assign the

441 00:24:11,530 –> 00:24:15,570 two states in the grid

比如你有一个格离散的原始状态空间和四个格子外的点，这时候就有一种概率机制来分配两个格子外的状态的概率

442 00:24:17,910 –> 00:24:22,380 the aggregate system works as follows

聚合系统是这样工作的

443 00:24:22,380 –> 00:24:27,870 you start at some representative state

从一个代表性状态开始

444 00:24:27,870 –> 00:24:32,830 you generate another state J according

445 00:24:32,830 –> 00:24:35,830 to this green arrows according to the

446 00:24:35,830 –> 00:24:37,780 probable transition probabilities of the original system

你根据这条绿线，也就是原系统的状态转移概率生成了另一个状态 j

447 00:24:37,780 –> 00:24:42,460 and then you go back to

448 00:24:42,460 –> 00:24:45,010 the represented States by means of this

449 00:24:45,010 –> 00:24:48,430 red probabilistic mechanism

然后你根据红线的概率机制回到了这些代表性状态

450 00:24:48,430 –> 00:24:50,770 so in the end you go from representative state to

451 00:24:50,770 –> 00:24:52,420 representative state according to this

452 00:24:52,420 –> 00:24:54,490 probabilistic mechanism and that defines

453 00:24:54,490 –> 00:24:58,420 the aggregate problem

所以你遵循某种概率机制从一个代表性节点到了另一个代表性节点，这就是聚合问题的定义
454 00:24:58,420 –> 00:25:00,940 this is what's used when you discretize a continuous
455 00:25:00,940 –> 00:25:02,920 state space that's a general way to
456 00:25:02,920 –> 00:25:05,500 discretize a continuous state space you
457 00:25:05,500 –> 00:25:08,350 generate a grid and then a mechanism of
458 00:25:08,350 –> 00:25:11,500 passing from the grid to the non grid bigger space
这就是你离散化连续状态空间的方法，这是一种离散化连续状态空间很常用的方法，生成一个格子，然后设计从一个格子到格子外更大空间的机制
459 00:25:11,500 –> 00:25:16,570 it's very well-suited for
460 00:25:16,570 –> 00:25:18,310 discrete Euclidian space discretization
461 00:25:18,310 –> 00:25:23,580 as in control problems for example
它非常适合于离散连续欧几里得空间，比如很多控制问题
462 00:25:23,580 –> 00:25:27,190 also substantial area of application is for
463 00:25:27,190 –> 00:25:30,190 POMDP I don't know if you've heard the
464 00:25:30,190 –> 00:25:37,150 term POMDP POMDP but POMDP starts
465 00:25:37,150 –> 00:25:40,870 for partially observable Markov decision
466 00:25:40,870 –> 00:25:43,930 problems in partially observable macro
467 00:25:43,930 –> 00:25:47,380 Markov decision problems you do not have
468 00:25:47,380 –> 00:25:49,660 exact knowledge of the state the
469 00:25:49,660 –> 00:25:52,000 controller does not observe the state
470 00:25:52,000 –> 00:25:55,320 itself but rather receives measurements
471 00:25:55,320 –> 00:25:58,900 about the possible whereabouts of the
472 00:25:58,900 –> 00:26:03,060 state and formulates a belief state
473 00:26:03,060 –> 00:26:07,680 which is a probability distribution of
474 00:26:07,680 –> 00:26:13,210 the different states of the problem
另一个大量应用的领域是 POMDP，不知道你们听没听说过 POMDP，也就是部分可观测马尔可夫决策过程，特殊地，POMDP 只能观测部分状态信息，控制器无法观测系统状态而是估算当前状态的分布，建立一个不同状态的概率分布
475 00:26:13,210 –> 00:26:16,570 so from one belief state I receive a
476 00:26:16,570 –> 00:26:19,269 measurement and I moved to another
477 00:26:19,269 –> 00:26:21,669 belief state that depends on that
478 00:26:21,669 –> 00:26:25,269 measurement and this defines a markov
479 00:26:25,269 –> 00:26:27,909 decision problem on the space of beliefs
480 00:26:27,909 –> 00:26:31,690 a higher dimensional space a continuous state space
从一个置信状态 (belief state)，我计算了它们的分布然后移动到另一个状态，这就定义了一个在高维连续状态空间内的置信空间马尔可夫决策问题
481 00:26:31,690 –> 00:26:36,489 because beliefs live into the unit simplex
因为 beliefs 出现在单元格里
482 00:26:36,489 –> 00:26:43,869 okay so to solve this
483 00:26:43,869 –> 00:26:46,570 POMDP type problem you have this
484 00:26:46,570 –> 00:26:48,789 unit simplex you introduce a
485 00:26:48,789 –> 00:26:51,039 discretization of the simplex and you
486 00:26:51,039 –> 00:26:53,169 define an aggregate transition
487 00:26:53,169 –> 00:26:55,659 probability mechanism that involves just
488 00:26:55,659 –> 00:27:00,119 the discrete subset of beliefs
为了求解这个 POMDP 问题，你对它进行了离散化，然后定义了一种只包括 belief 的离散子集合的聚合状态转移概率机制
489 00:27:00,119 –> 00:27:02,289 formulate the corresponding Markov
490 00:27:02,289 –> 00:27:05,349 decision problem that involves just a
491 00:27:05,349 –> 00:27:09,849 finite state Markov chain solve that in
492 00:27:09,849 –> 00:27:12,339 any way you want even approximately and
493 00:27:12,339 –> 00:27:14,889 that gives you by means of this formula
494 00:27:14,889 –> 00:27:17,679 gives you an approximation the solution
495 00:27:17,679 –> 00:27:23,499 of the POMDP which is piecewise linear
496 00:27:23,499 –> 00:27:26,469 according to this proper list this
497 00:27:26,469 –> 00:27:29,739 probabilistic mechanism
对这个有限状态马尔科夫决策问题建模，可以使用任何近似方法进行求解，然后根据这个可以得到它的分段线性近似解

498 00:27:29,739 –> 00:27:31,779 I don't want to get into this but this is typically the
499 00:27:31,779 –> 00:27:33,759 kind of discretization you use for
500 00:27:33,759 –> 00:27:37,440 continuous type of state space
我不想深入讲这个问题，但是他确实是你用来对连续状态空间进行离散化的工具

1.8 EXAMPLE IV: REPRESENTATIVE FEATURES

501 00:27:43,560 –> 00:27:46,320 okay now here's another scheme that sort
502 00:27:46,320 –> 00:27:48,210 of contains all the previous one are special cases
这里有一种包括了之前所有内容的聚合方法
503 00:27:48,210 –> 00:27:52,440 instead of having the
504 00:27:52,440 –> 00:27:55,740 representative states I have a representative subsets
这种方法中没有代表性状态而是代表性子集合
505 00:27:55,740 –> 00:27:59,910 so an aggregate
506 00:27:59,910 –> 00:28:02,910 state having a state one is associated
507 00:28:02,910 –> 00:28:05,610 with a subset of states having state two
508 00:28:05,610 –> 00:28:07,710 is associate another subset of states I
509 00:28:07,710 –> 00:28:09,500 have all the subsets
一个聚合状态表示原状态的一个子集合，另一个聚合状态表示另一个子集合，我有所有的聚合状态
510 00:28:09,500 –> 00:28:12,720 however the subset may involve more than
511 00:28:12,720 –> 00:28:16,680 one state and they need not cover the entire space
然而一个子集合可能包括不止一个状态，同时所有的子集合可能没有包括状态空间中的所有状态
512 00:28:16,680 –> 00:28:24,480 to give you an example with
513 00:28:24,480 –> 00:28:27,660 aggregate state X I may associate a
514 00:28:27,660 –> 00:28:29,640 group of states that has similar
515 00:28:29,640 –> 00:28:33,540 features but some states that do not are
516 00:28:33,540 –> 00:28:37,050 not clearly identified with any one of
517 00:28:37,050 –> 00:28:39,630 the set of features that define this
518 00:28:39,630 –> 00:28:44,580 this set are sort of left outside of the active States
举个例子，聚合状态可以与一组原系统中特征相似的状态相联系，但是其他状态不与其他任何聚合状态相关，这样就出现了几个聚合状态外的原始状态
519 00:28:44,580 –> 00:28:50,670 so in this scheme I start
520 00:28:50,670 –> 00:28:53,400 out with some aggregate state I
521 00:28:53,400 –> 00:28:56,220 randomize within here obtain the
522 00:28:56,220 –> 00:28:58,500 transition probabilities to go to other
523 00:28:58,500 –> 00:29:01,830 states and then from there I go to
524 00:29:01,830 –> 00:29:04,230 aggregate States according to the phi probabilities
用这种方法我从一个聚合状态开始，随机选一个原始状态，然后根据状态转移概率跳转到另一个状态，再根概率据 ϕ 从这个状态跳转到另一个聚合状态
525 00:29:04,230 –> 00:29:10,080 there are some researchers
526 00:29:10,080 –> 00:29:13,730 with the scheme the aggregate states
527 00:29:13,730 –> 00:29:19,230 this subset should be disjoint
关于这个方法有一些研究，聚合状态表示的原始状态子集合必须不相关
528 00:29:19,230 –> 00:29:21,120 the disaggregation probabilities are
529 00:29:21,120 –> 00:29:25,830 positive only within here but zero outside
分解概率只有在原始状态在聚合状态内的时候是正数，其他时候都是 0
530 00:29:25,830 –> 00:29:29,880 and the aggregation
531 00:29:29,880 –> 00:29:33,060 probabilities we will here are equal to one
聚合状态内的原始状态的聚合概率一定等于 1
532 00:29:33,060 –> 00:29:37,230 so if you end up with a state J in
533 00:29:37,230 –> 00:29:44,160 here you stay in here
如果你停止在状态 j，那么你一定处于 j 所在的聚合状态
534 00:29:44,160 –> 00:29:47,310 hard aggregation is the special case
硬聚合是一种特殊情况
535 00:29:47,310 –> 00:29:50,700 where these sets cover the entire space
所有的子集合刚好覆盖整个状态空间

536 00:29:50,700 –> 00:29:51,750 aggregation with representative
537 00:29:51,750 –> 00:29:56,850 is a special case where these aggregate
538 00:29:56,850 –> 00:30:00,090 states consist of a single element
带有表示性的聚合也是一种特例，每一个聚合状态只包括一个原始状态
539 00:30:00,090 –> 00:30:02,460 so more general scheme in quite generally applicable
更一般的方法会用在一般的应用上
540 00:30:02,460 –> 00:30:11,750 any questions so far
541 00:30:11,750 –> 00:30:16,140 about how we formulate these aggregate
542 00:30:16,140 –> 00:30:18,030 problems and the different types of
543 00:30:18,030 –> 00:30:21,680 aggregations that we can use
到这有什么问题吗

1.9 APPROXIMATE PI BY AGGREGATION

544 00:30:31,250 –> 00:30:34,730 okay so now assume that we have
545 00:30:34,730 –> 00:30:37,039 formulate an aggregate problem we can
546 00:30:37,039 –> 00:30:39,919 try to do value iteration policy
547 00:30:39,919 –> 00:30:42,440 duration or approximate versions of those
假设我们有一个聚合问题，可以用值迭代和策略迭代，也可以用他们的近似方法来求解
548 00:30:42,440 –> 00:30:45,740 let's focus on approximate policy
549 00:30:45,740 –> 00:30:49,549 duration we define which we are the
550 00:30:49,549 –> 00:30:51,559 typical iteration we have a policy that
551 00:30:51,559 –> 00:30:59,179 map states X into controls
我们来看一下近似策略迭代，我们定义一个从状态 x 到控制的映射
552 00:30:59,179 –> 00:31:01,960 and we solve the corresponding aggregation equation
553 00:31:01,960 –> 00:31:05,480 associated with that policy this is this
554 00:31:05,480 –> 00:31:09,169 equation here R is a fixed point of this
555 00:31:09,169 –> 00:31:12,370 equation here these the matrix like this
556 00:31:12,370 –> 00:31:15,169 T mu is the mapping that is
557 00:31:15,169 –> 00:31:20,539 high-dimensional phi is the matrix its
558 00:31:20,539 –> 00:31:22,880 allegation matrix and R is the vector of
559 00:31:22,880 –> 00:31:25,159 cost of the aggregate States for mu
我们来求解这个给定策略的聚合方程，这就是这个方程 ($\bar{J} = \Phi R$)
560 00:31:25,159 –> 00:31:27,890 and then we evaluate the policy
561 00:31:27,890 –> 00:31:31,880 according to this equation
然后我们用这个方程 ($\bar{J} = \Phi R$) 评价这个策略
562 00:31:31,880 –> 00:31:34,520 and this can be done by simulation
评价可以用仿真来完成
563 00:31:34,520 –> 00:31:37,220 if you can do simulation of the original system you
564 00:31:37,220 –> 00:31:39,020 can also do simulation of the aggregate system
如果你可以对原问题仿真，那么你也可以对聚合问题进行仿真
565 00:31:39,020 –> 00:31:41,090 you just use the aggregation
566 00:31:41,090 –> 00:31:43,340 probabilities to map to happen to go up this way
你可以用聚合概率来完成仿真
567 00:31:43,340 –> 00:31:45,530 use the simulator of the
568 00:31:45,530 –> 00:31:47,750 original system to go to J
用仿真让原系统状态到达 j
569 00:31:47,750 –> 00:31:49,039 and then use the disagree the aggregation
570 00:31:49,039 –> 00:31:52,000 probabilities to go to Y
然后用聚合概率转移到 y
571 00:31:52,000 –> 00:31:55,730 back into i across down and so on
从 i 转移过来，然后继续转移
572 00:31:55,730 –> 00:31:59,659 the simulator is very easy to implement what you once you
573 00:31:59,659 –> 00:32:01,250 have a simulator of the original system
一旦有了原系统的仿真，聚合系统的仿真也会很简单
574 00:32:01,250 –> 00:32:05,000 and there is there are simulation based
575 00:32:05,000 –> 00:32:08,690 versions for now for for solving this equation
这是基于仿真的方法来求解这些方程

576 00:32:08,690 –> 00:32:12,590 in these equations actually are
577 00:32:12,590 –> 00:32:15,890 exact we solve exactly this problem by
578 00:32:15,890 –> 00:32:17,059 simulation there is no approximation
579 00:32:17,059 –> 00:32:20,020 involved
我们要使用仿真求这个问题的精确解而不是近似解
580 00:32:21,840 –> 00:32:26,580 now look at this equation here
让我们来看一下这个方程 ($R = DT_\mu(\Phi R)$)
581 00:32:26,580 –> 00:32:29,430 it looks also like a projected equation
这看起来和投影方程很像
582 00:32:29,430 –> 00:32:34,070 the projected equation has this form
投影方程是这样的 ($R = \Pi T_\mu(\Phi R)$)
583 00:32:34,070 –> 00:32:40,430 so if phi R is equal to phi d phi T mu phi R and phi d is a projection
$\Phi R = \Phi D \Phi T_\mu(\Phi R)$，这时候 ΦD 是一个投影
584 00:32:40,430 –> 00:32:46,230 then I also have a
585 00:32:46,230 –> 00:32:50,180 projected equation type of approximation
这时候同同样能活得一个近似投影方程
586 00:32:51,110 –> 00:32:53,640 so it looks like the project an equation
587 00:32:53,640 –> 00:32:56,190 approach but it has the advantage there
588 00:32:56,190 –> 00:32:57,950 is no problem with oscillations
所以它看起来是一个投影方程的方法，但是有一个好处是不会出现震荡的情况
589 00:32:57,950 –> 00:33:00,000 why is there no problem with oscillations
为什么不会震荡呢
590 00:33:00,000 –> 00:33:03,630 because I'm applying policy
591 00:33:03,630 –> 00:33:06,030 Direction exactly with no approximations
592 00:33:06,030 –> 00:33:08,730 to this smaller problem and I know that
593 00:33:08,730 –> 00:33:10,590 this converges in a finite number of steps
因为我没有近似求解而是用精确策略迭代来求解这个更小的问题，我知道他会在有限次的迭代后收敛
594 00:33:10,590 –> 00:33:16,950 a relative to the projected
595 00:33:16,950 –> 00:33:18,840 equation approach there's a disadvantage
596 00:33:18,840 –> 00:33:22,290 in that there is a restriction for Phi
597 00:33:22,290 –> 00:33:24,450 and for D to be probability distribution
这种投影方法的一个缺点就是 Φ 和 D 的概率分布是受到限制的
598 00:33:24,450 –> 00:33:30,030 I cannot use arbitrary features
我不能用任意特征来计算
599 00:33:30,030 –> 00:33:32,010 they have to have they have to be defined by
600 00:33:32,010 –> 00:33:35,180 probability distributions and that is
601 00:33:35,180 –> 00:33:39,470 this significant resection
他们必须被定义成概率分布的形式，这就是这种方法的缺陷
602 00:33:44,200 –> 00:33:47,650 so this approximate policy direction
603 00:33:47,650 –> 00:33:50,860 there is also an approximate value
604 00:33:50,860 –> 00:33:54,460 iteration method associated with this
605 00:33:54,460 –> 00:33:58,730 with this aggregation and it is exact
606 00:33:58,730 –> 00:34:01,460 value iteration for the approximate problem
所以这就是近似策略迭代，同样还有聚合问题的近似值迭代方法，用精确值迭代来求解近似问题，
607 00:34:01,460 –> 00:34:06,080 followed by this feature based
608 00:34:06,080 –> 00:34:08,270 weighting that gives you an
609 00:34:08,270 –> 00:34:11,890 approximation to the original problem
用这个基于特征的权重 (ΦR) 对原问题进行近似
610 00:34:15,100 –> 00:34:17,409 now there are some issues associated
611 00:34:17,409 –> 00:34:24,580 with with this procedure and there are
612 00:34:24,580 –> 00:34:26,710 too many for me to cover what I'd like
613 00:34:26,710 –> 00:34:30,520 to do is cover them selectively in the
614 00:34:30,520 –> 00:34:33,639 last part of this lecture so let's take
615 00:34:33,639 –> 00:34:37,120 a break now and come back and discuss these

有一些与这个方法相关的话题，太多了，我想要在最后一次课程有选择性地降一点，我们现在先休息一会

(someone asking questions

问题：状态聚合对探索有帮助么

答案：有，因为聚合状态太少了，所以很容易就可以进行很充分地探索，就不用仿真选择轨迹了，可以对聚合状态采样然后进行一步转移

616 00:34:37,120 –> 00:34:49,239 have any questions yes okay so the
617 00:34:49,239 –> 00:34:55,389 question is does aggregation help with
618 00:34:55,389 –> 00:34:59,650 the issue of exploration and the answer
619 00:34:59,650 –> 00:35:03,690 is yes it is much easier to implement
620 00:35:03,690 –> 00:35:06,700 exploration in a simulation based
621 00:35:06,700 –> 00:35:09,510 contact context within this setting
622 00:35:09,510 –> 00:35:12,910 because you simply choose you choose a
623 00:35:12,910 –> 00:35:15,010 sampling scheme that involves
624 00:35:15,010 –> 00:35:16,780 exploration adequate exploration of
625 00:35:16,780 –> 00:35:18,970 these states here that's much easier to
626 00:35:18,970 –> 00:35:19,360 do
627 00:35:19,360 –> 00:35:21,490 for example they don't have to be
628 00:35:21,490 –> 00:35:24,190 generated by a single trajectory they
629 00:35:24,190 –> 00:35:27,430 can be generated by just collecting a
630 00:35:27,430 –> 00:35:30,490 set of sample States aggregate States
631 00:35:30,490 –> 00:35:33,220 and using that in one step transitions
632 00:35:33,220 –> 00:35:39,360 this will work and it will any pool and
633 00:35:39,360 –> 00:35:42,100 and it will have absolutely no problem
634 00:35:42,100 –> 00:35:45,580 with exploration so yes it's not only
635 00:35:45,580 –> 00:35:47,560 easier to deal with oscillations within
636 00:35:47,560 –> 00:35:49,600 this context but also exploration is an
637 00:35:49,600 –> 00:35:53,340 additional area where you get a benefit
638 00:35:53,340 –> 00:35:56,020 but the disadvantage of course is that
639 00:35:56,020 –> 00:36:00,360 is this restriction that's
640 00:36:02,730 –> 00:36:05,040 okay so let's take a break and get back
641 00:36:05,040 –> 00:00:00,000 in about ten minutes