

## 4. OPTIMAL STOPPING TIME

4.1. **Definitions.** On the first day I explained the basic problem using one example in the book. On the second day I explained how the solution to the problem is given by a “minimal superharmonic” and how you could find one using an iteration algorithm. Also, a simple geometric construction gives the solution for fair random walks. On the third day I explained the variations of the game in which there is a fixed cost per move and on the fourth day I did the payoff with discount. I skipped the continuous time problem.

4.2. **The basic problem.** I started with the example given in the book: You roll a die. If you get a 6 you lose and get nothing. But if you get any other number you get the value on the die (1,2,3,4 or 5 dollars). If you are not satisfied with what you get, you can roll over and you give up your reward. For example, if you roll a 1 you probably want to go again. But, if you roll a 6 at any time then you lose: You get nothing. The question is: When should you stop? The answer needs to be a strategy: “Stop when you get 4 or 5.” or maybe “Stop when you get 3,4 or 5.” You want to choose the best “stopping time.”

4.2.1. *stopping time.*

**Definition 4.1.** In a stochastic process  $T$  is called a *stopping time* if you can tell when it happens.

Basically, a stopping time is a formula which, given  $X_1, X_2, \dots, X_n$  tells you whether to stop at step  $n$ . (Or in continuous time, given  $X_t$  for  $t \leq T$ , tells you whether  $T$  is the stopping time.)

Some examples of stopping time are:

- (1) the 5th visit to state  $x$
- (2) 10 minutes after the second visit to  $y$ .
- (3) the moment the sum  $X_1 + X_2 + \dots + X_n$  exceeds 100.

You cannot stop right before something happens. In class we discussed the scenario where you are driving on the highway and you are about to have an accident. Is the second before the moment of impact a stopping time? Even if the probability is 1, you are not allowed to call it a stopping time because “probability one” is not good enough. You have to use the information you have until that moment to decide if this is stopping time. For example, you could say,  $T$  is the moment your car gets within 1 *cm* of the car in front of you. That would be a stopping time.

4.2.2. *payoff function.* The *payoff function* is a function

$$f : S \rightarrow \mathbb{R}$$

which assigns to each state  $x \in S$  a number  $f(x) \in \mathbb{R}$  representing how much you get if you stop at state  $x$ . To figure out whether to stop you need to look at what you can expect to happen if you don't stop.

- (1) If you stop you get  $f(x)$ .
- (2) If, starting at  $x$ , you take one step and then stop you expect to get

$$\sum p(x, y)f(y)$$

You need to analyze the game before you play and decide on an algorithm when to stop. (Or you have someone play for you and you give them very explicit instructions when to stop and take the payoff.) This stopping time is  $T$ .

$X_T$  is the state that you stop in.

$f(X_T)$  is the payoff that you will get.

You want to maximize  $f(X_T)$ .

4.2.3. *value function.* The *value function*  $v(x)$  is the expected payoff using the optimal strategy starting at state  $x$ .

$$v(x) = \mathbb{E}(f(X_T) | X_0 = x)$$

Here  $T$  is the optimal stopping time. You need to remember that this is given by an algorithm based on the information you have up to and including that point in time.

**Theorem 4.2.** *The value function  $v(x)$  satisfies the equation*

$$v(x) = \max(f(x), \sum_y p(x, y)v(y))$$

In this equation,

$f(x)$  = your payoff if you stop.

$\sum_y p(x, y)v(y)$  = your expected payoff if you continue.

Here you assume you are going to use the optimal strategy if you continue. That is why you will get  $v(y)$  instead of  $f(y)$ . When you compare these two ( $f(x)$  and  $\sum p(x, y)v(y)$ ), the larger number tells you what you should do: stop or play.

The basic problem is to find the optimal stopping time  $T$  and calculate the value function  $v(x)$ .

**Example 4.3.** Suppose that you toss a die over and over. If you get  $x$  your payoff is

$$f(x) = \begin{cases} x & \text{if } x \neq 6 \\ 0 & \text{if } x = 6 \end{cases}$$

And: if you roll a 6 you lose and the game is over. I.e., 6 is recurrent. If  $X_0$  is your first toss,  $X_1$  your second, etc. the probability transition matrix is:

$$P = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Since  $v(6) = 0$ , the second number in the boxed equation is the product of the matrix  $P$  and the *column vector*  $v$ :

$$\sum_y p(x, y)v(y) = Pv(x) = \frac{1}{6}(v(1) + v(2) + v(3) + v(4) + v(5))$$

(for  $x < 6$ ). I pointed out that, since the first 5 rows of  $P$  are the same, the first 5 entries in the column vector  $Pv$  are the same (and the 6th entry is 0).

**4.3. Solutions to basic problem.** On the second day I talked about solutions to the optimal stopping time problem and I explained:

- (1) minimal superharmonics
- (2) the iteration algorithm
- (3) solution for random walks

4.3.1. *minimal superharmonic.*

**Definition 4.4.** A *superharmonic* for the Markov chain  $X_n$  is a real valued function  $u(x)$  for  $x \in S$  so that

$$u(x) \geq \sum_{y \in S} p(x, y)u(y)$$

In matrix form the definition is

$$u(x) \geq (Pu)(x)$$

where  $u$  is a column vector.

**Example 4.5.** Roll one die and keep doing it until you get a 6. (6 is an absorbing state.) The payoff function is:

states $x$	payoff $f(x)$	probability $\mathbb{P}$
1	150	1/6
2	150	1/6
3	150	1/6
4	300	1/6
5	300	1/6
6	0	1/6

The transition matrix in this example is actually  $6 \times 6$  as in the first example. But I combined these into 3 states<sup>1</sup>:  $A = 1, 2$  or  $3$ ,  $B = 4$  or  $5$  and  $C = 6$ :

states $x$	payoff $f(x)$	probability $\mathbb{P}$
$A$	150	1/2
$B$	300	1/3
$C$	0	1/6

Then, instead of a  $6 \times 6$  matrix,  $P$  became a  $3 \times 3$  matrix:

$$P = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 0 & 0 & 1 \end{pmatrix}$$

The best payoff function you can hope for is (the column vector)

$$u = (300, 300, 0)^t$$

where the  $^t$  means transpose. (But later I dropped the  $^t$ .) Then

$$Pu = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 300 \\ 300 \\ 0 \end{pmatrix} = \begin{pmatrix} 250 \\ 250 \\ 0 \end{pmatrix}$$

Since  $300 \geq 250$ ,  $u = (300, 300, 0)$  is superharmonic.

**Theorem 4.6.** *The value function  $v(x)$  is the minimal superharmonic so that  $v(x) \geq f(x)$  for all states  $x$ .*

This doesn't tell us how to find  $v(x)$ . It is used to prove that the iteration algorithm converges to  $v(x)$ .

<sup>1</sup>You can combine two states  $x, y$  if:

- (1)  $f(x) = f(y)$  and
- (2) the  $x$  and  $y$  rows of the transition matrix  $P$  are identical.

4.3.2. *iteration algorithm.* This gives a sequence of superharmonics which converge to  $v(x)$ . You start with  $u_1$  which is the most optimistic. This the best payoff you can expect to get:

$$u_1(x) = \begin{cases} 0 & \text{if } x \text{ is absorbing} \\ \max f(y) & \text{if } x \text{ is transient} \end{cases}$$

In the example,  $\max f(y) = 300$  and  $C$  is absorbing. So,

$$u_1 = \begin{pmatrix} u_1(A) \\ u_1(B) \\ u_1(C) \end{pmatrix} = \begin{pmatrix} 300 \\ 300 \\ 0 \end{pmatrix}$$

Next,  $u_2$  is given by

$$u_2(x) = \max(f(x), (Pu_1)(x))$$

We just figured that  $Pu_1 = (250, 250, 0)$ . So,

$$u_2 = \max \begin{pmatrix} 150 \\ 300 \\ 0 \end{pmatrix} \begin{pmatrix} 250 \\ 250 \\ 0 \end{pmatrix} = \begin{pmatrix} 250 \\ 300 \\ 0 \end{pmatrix}$$

Keep doing this using the recursive equation:

$$u_{n+1}(x) = \max(f(x), (Pu_n)(x))$$

You get:

$$\begin{aligned} u_1 &= (300, 300, 0) \\ u_2 &= (250, 300, 0) \\ u_3 &= (225, 300, 0) \\ u_4 &= (212.5, 300, 0) \end{aligned}$$

When you do this algorithm you get an approximate answer since

$$\lim_{n \rightarrow \infty} u_n(x) = v(x)$$

To get an exact answer you need to realize that only the first number is changing. So, you let  $z = v(A)$  be the limit of this first number. Then:

$$z = v(A) = \max(f(A), Pv(A)) = \max(150, Pv(A)) = Pv(A)$$

(The calculation shows that  $z \approx 200 > 150$ .) Once you get rid of the “max” you can solve the equation:

$$z = Pv(A) = \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) \begin{pmatrix} z \\ 300 \\ 0 \end{pmatrix} = \frac{z}{2} + \frac{300}{3} = \frac{z}{2} + 100$$

So,

$$z = 200$$

and

$$v = (200, 300, 0)$$

The optimal strategy is to stop if you get 4 or 5 and play if you get 1, 2 or 3.

4.3.3. *concave-down value function.*<sup>2</sup> Suppose you have a simple random walk with absorbing walls. Then, for  $x$  not one of the walls, you go left or right with probability  $1/2$ :

$$p(x, x+1) = \frac{1}{2}$$

$$p(x, x-1) = \frac{1}{2}$$

and  $p(x, y) = 0$  in other cases. A function  $u(x)$  is superharmonic if

$$u(x) \geq \sum_y p(x, y)u(y) = \frac{u(x-1) + u(x+1)}{2}$$

This equation says that the graph of the function  $u(x)$  is *concave down*. In other words, the point  $(x, u(x))$  is above the point which is midway between  $(x-1, u(x-1))$  and  $(x+1, u(x+1))$ .

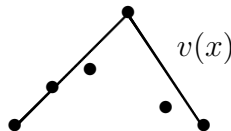
So, the theorem that the value function  $v(x)$  is the minimal superharmonic so that  $v(x) \geq f(x)$  means that the graph of  $v(x)$  is the convex hull of the graph of  $f(x)$ .

**Example 4.7.** Suppose that we have a random walk on  $S = \{0, 1, 2, 3, 4, 5\}$  with absorbing walls and payoff function:

$$\begin{array}{rcccccc} x = & 0 & 1 & 2 & 3 & 4 & 5 \\ f(x) = & 0 & 2 & 3 & 6 & 1 & 0 \end{array}$$



Then  $v(x)$  is the convex hull of this curve:



<sup>2</sup>Students were correct to point out that “convex” means *concave up*.