MATH 56A SPRING 2008
STOCHASTIC PROCESSES

KIYOSHI IGUSA

Contents

Date: March 17, 2008.

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 chose 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 $X_1, X_2, \cdots, X_n$ 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 $X_1 + X_2 + \cdots + X_n$ 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

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

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 an take the payoff.) This stopping time is T.

$$X_T \text{ is the state that you stop in.}$$

$$f(X_T) \text{ is the payoff that you will get.}$$

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.

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

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*

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

In this equation,

$$f(x) = \text{ your payoff if you stop.}$$

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

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 6×6 as in the first example. But I combined these into 3 states[1]: $A = 1, 2$ 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 6×6 matrix, P became a 3×3 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 $v(x)$.

---

[1]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,

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

Keep doing this using the recursive equation:

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

You get:

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

When you do this algorithm you get an approximate answer since

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

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:

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

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.* [2] 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

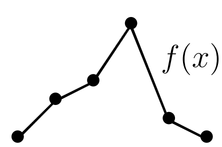
$$u(x) \geq \sum_y p(x, y)u(y) = \frac{u(x-1) + u(x+1)}{2}$$

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 $(x+1, u(x+1))$.

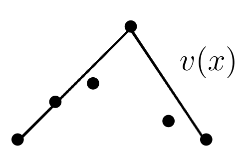
So, the theorem that the value function $v(x)$ is the minimal superharmonic so that $v(x) \geq f(x)$ means that the graph of $v(x)$ is the convex hull of the graph of $f(x)$.

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

$x =$	0	1	2	3	4	5
$f(x) =$	0	2	3	6	1	0



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



[2]Students were correct to point out that “convex” means *concave up*.

4.4. **Cost functions.** The *cost function* $g(x)$ gives the price you must pay to continue from state x. If T is your *stopping time* then X_T is your *stopping state* and $f(X_T)$ is your *payoff*. But your *cost* to play to that point was

$$\text{cost } = g(X_0) + g(X_1) + \cdots + g(X_{T-1}) = \sum_{j=0}^{T-1} g(X_j)$$

So, your net gain is

$$\text{net } = f(X_T) - \sum_{j=0}^{T-1} g(X_j)$$

The *value function* $v(x)$ is the expected net gain when using the optimal stopping time starting at state x:

$$v(x) = \mathbb{E}(f(X_T) - \sum_{j=0}^{T-1} g(X_j) \,|\, X_0 = x)$$

It satisfies the equation:

$$\boxed{v(x) = \max(f(x), (Pv)(x) - g(x))}$$

Proof: In state x you should either *stop* or *continue.*

(1) If you stop you get $f(x)$.
(2) If you continuous and use the optimal strategy after that then you get $v(y)$ with probability $p(x, y)$ but you have to pay $g(x)$. So, you would expect to get

$$\sum_{y \neq x} p(x, y)v(y) - g(x)$$

You should pick the one which gives you a higher expected net. So, $v(x)$ is the maximum of these two numbers.

4.4.1. *nonstochastic case.* The first example was a dice problem. You toss two dice and let $x =$ the sum. Then $2 \leq x \leq 12$ with probabilities indicated below.

$p =$	1	2	3	4	5	6	5	4	3	2	1	$\cdot 1/36$
$x =$	2	3	4	5	6	**7**	8	9	10	11	12	
$f(x) =$	2	3	4	5	6	**0**	8	9	10	11	12	
$g(x) =$	2	2	2	2	1	**12**	1	1	1	1	1	

There was a question of what should be $g(7)$. It does not make sense to talk about the cost of continuing from 7 if you are not allowed to continue. So, I decided that, in order to have a well defined function ($g(x)$ needs to be defined for every $x \in S$ in order for g to be a function), we should allow the player to pay $\max f(x) = 12$ to continue from 7. It doesn't make sense to pay 12 to play a game where the maximum gain is 12. So, this has the effect of making 7 recurrent.

The problem is to find the value function $v(x)$ and the optimal strategy (the formula for T). I pointed out that this Markov chain is actually *not stochastic* in the sense that the probabilities do not change with time. This implies that the value function $v(x)$ which is a vector with 10 unknown coordinates (11 coordinates of which we know only $v(7) = 0$):

$$v = (v(2), v(3), \cdots, v(6), v(7) = 0, v(8), \cdots, v(12)$$

is determined by one number

$$E = \text{ the expected payoff if you continue}$$

Then, your expected net if you continue is $E - g(x)$ so

$$v(x) = \max(f(x), E - g(x)) \text{ if } x \neq 7.$$

And E is given in terms of v by:

$$E = \sum_{x \neq 7} p_x v(x)$$

When you do the iteration algorithm you compute

$$E_n = \sum_{x \neq 7} p_x u_n(x)$$

and you get a sequence of numbers

$$E_1, E_2, E_3, \cdots$$

All you need is the single number

$$E = E_\infty = \lim_{n \to \infty} E_n.$$

No cost First I did this in the no cost case.

When $n = 1$ you take the most optimistic view: Hope to get $x = 12$. But you have a probability $p_7 = 6/36 = 1/6$ of getting $x = 7$ and losing. So,

$$E_1 = \sum_{x \neq 7} p_x u_1(x) = \sum_{x \neq 7} p_x 12 = (5/6)12 = 10.$$

Then

$$u_2(x) = \max(f(x), E) \text{ if } x \neq 7$$

(But make sure to put $u_2(7) = 0$):

$x =$	2	3	4	5	6	**7**	8	9	10	11	12
$f(x) =$	2	3	4	5	6	**0**	8	9	10	**11**	**12**
$E =$	10	10	10	10	10	10	10	10	10	10	10
$u_2(x) =$	10	10	10	10	10	**0**	10	10	10	11	12

If you take the average value of $u_2(x)$ you get E_2:

$$E_2 = \sum_{x \neq 7} p_x u_2(x) = 8.4444...$$

Repeating this process you get:

$$\begin{aligned} E_3 &= 7.4691 \\ E_4 &= 7.001 \\ E_5 &= 6.806 \\ E_6 &= 6.7247 \\ &\cdots \\ E_{25} &= 6.6667 \end{aligned}$$

Once you realize that E_∞ is somewhere between 6 and 7 you know the *winning strategy*: You need to continue if you get 6 or less and stop if you get 8 or more. So,

$$v(x) = (E, E, E, E, E, 0, 8, 9, 10, 11, 12)$$

which makes

$$E = \frac{1}{36}\left(E + 2E + 3E + 4E + 5E + 5(8) + 4(9) + 3(10) + 2(11) + 12\right)$$

$$= \frac{1}{36}(15E + 140)$$

$$36E = 15E + 140$$

$$E = 140/21 = 20/3 = 6\tfrac{2}{3}$$

So, the value function is the vector:

$$v = (\tfrac{20}{3}, \tfrac{20}{3}, \tfrac{20}{3}, \tfrac{20}{3}, \tfrac{20}{3}, 0, 8, 9, 10, 11, 12)$$

<u>With cost</u>

The iteration algorithm starts with (I don't remember what I said but you have to remember that $u_n(x) \geq f(x)$ all the time):

$$u_1(x) = \begin{cases} 0 & \text{if } x = 7 \\ f(x) & \text{if } f(x) \geq \max f(y) - g(x) \\ \underbrace{\max f(y)}_{\text{hope for best}} - \underbrace{g(x)}_{\text{cost}} & \text{otherwise} \end{cases}$$

This gives:

$$u_1 = (10, 10, 10, 10, 11, 0, 11, 11, 11, 11, 12)$$

The average of these numbers is:

$$E_1 = \sum_{x \neq 7} p_x u_1(x) = 8.917$$

Then

$$u_2(x) = \max(f(x), E_1 - g(x))$$

$$u_2 = (6.917, 6.917, 6.917, 6.917, 7.917, 0, 8, 9, 10, 11, 12)$$

with average

$$E_2 = \sum_{x \neq 7} p_x u_2(x) = 6.910$$

$$\begin{aligned} E_3 &= 6.096 \\ E_4 &= 5.960 \\ &\cdots \\ E_{10} &= 5.939393 \\ E_{11} &= 5.939393 \end{aligned}$$

We just needed to know that E_∞ is between 5 and 6. This tells us that the optimal strategy is to continue if you get 2 or 3 and stop if you get 4 or more.

After you determine the optimal strategy, you can find the exact value of both E and the value function $v(x)$. First you find $v(x)$ in terms of E:

$$v(x) = (E - 2, E - 2, 4, 5, 6, 0, 8, 9, 10, 11, 12)$$

The average of these numbers is E. So,

$$E = (E - 2)(3/36) + 202/36$$

$$E = 196/33 = 5\tfrac{31}{33} = 5\tfrac{93}{99} = 5.939393\cdots$$

$$v(x) = (3\tfrac{31}{33}, 3\tfrac{31}{33}, 4, 5, 6, 0, 8, 9, 10, 11, 12)$$

4.4.2. *random walk.* with absorbing walls. In the general (stochastic) case the value function is the solution of the equation:

$$v(x) = \max(f(x), \underbrace{\sum_y p(x,y)v(y)}_{\frac{v(x-1)+v(x+1)}{2}} - g(x))$$

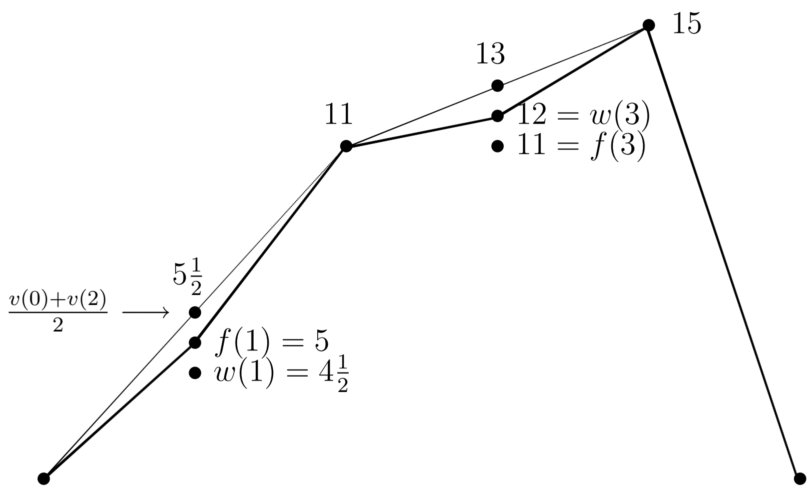
In the case of the random walk, this part ↑ is what we have before. So, $v(x)$ is the smallest function so that

$$v(x) \geq f(x) \quad \text{and}$$

$$v(x) \geq \frac{v(x-1)+v(x+1)}{2} - g(x)$$

Example 4.8. Suppose the payoff and cost functions are:

states $x =$	0	1	2	3	4	5
$f(x) =$	0	5	11	11	15	0
$g(x) =$	0	1	1	1	1	0



In the graph, the thin lines gives the convex hull of the function $f(x)$. This would be the answer if there were no cost. Since the cost is 1, we have to go one step below the average. I called this function $w(x)$:

$$w(x) := \frac{f(x-1)+f(x+1)}{2} - g(x)$$

Since $v(x) \geq f(x)$, it must also be $\geq w(x)$:

$$v(x) \geq \frac{v(x-1)+v(x+1)}{2} - g(x) \geq \frac{f(x-1)+f(x+1)}{2} - g(x) = w(x)$$

This is a simple case in which the gaps have length 2. So, we can just compare $f(x)$ and $w(x)$ to get the value function $v(x)$. If the gap is more than two, the equation becomes more complicated.

I think I forgot to say this: For the iteration algorithm we can start with the value function that we know how to calculate when there is no cost:

$$u_1(x) = (0, 5\tfrac{1}{2}, 11, 13, 15, 0)$$

Then

$$u_2(x) = \max\left(f(x), \frac{u_1(x-1) + u_1(x+1)}{2} - g(x)\right)$$

$$u_2 = (0, 5, 11, 12, 15, 0)$$

If you do it again, you get the same thing: $u_3 = u_2$. So, this is also equal to the value function:

$$v = (0, 5, 11, 12, 15, 0)$$

So, the optimal strategy is to continue when $x = 3$ (since that is the only point where $v(x) > f(x)$) but stop at any other point.

4.5. **discounted payoff.** Here we assume that the payoff is losing value at a fixed rate so that after T steps it will only be worth $\alpha^T f(x)$ where α is the discount rate, say $\alpha = .90$. If there is no cost, the value function will satisfy the equation

$$v(x) = \max(f(x), \alpha(Pv)(x))$$

Again there is a recursive formula converging to this answer:

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

where you start with

$$u_1(x) = \begin{cases} 0 & \text{if } x \text{ is absorbing} \\ f(x) & \text{if } f(x) \geq \alpha f(y) \text{ for all } y \\ \max \alpha f(y) & \text{otherwise} \end{cases}$$

Example 4.9. Random walk with absorbing walls:

$$\begin{array}{rcccccc} x = & 0 & 1 & 2 & 3 & 4 & 5 \\ f(x) = & 0 & 10 & 6 & 8 & 10 & 0 \\ \alpha = & 90\% & & & & & \end{array}$$

If $X_0 = 2$ then you get $f(x) = 6$ if you stop. But, if you continue you expect to get

$$\frac{f(x-1) + f(x+1)}{2} = \frac{10+8}{2} = 9$$

which will be worth

$$\alpha 9 = (.9)9 = 8.1 > 6$$

So, you should play even if there is a cost of $g(x) = 2$.

If $X_0 = 3$ then you get $f(x) = 8$ if you stop and

$$\alpha\left(\frac{f(x-1) + f(x+1)}{2}\right) = .9\left(\frac{6+10}{2}\right) = (.9)(8) = 7.2$$

if you continue. It looks like you should stop at $x = 3$. But we saw in class that this was not right.

Iteration algorithm: If you use $u_1 = (0, 10, 8.1, 7.2, 10, 0)$ then the iteration gives an sequence which increases and converges to $v(x)$ from below. That is because this vector is *subharmonic*. The superharmonic starting point is that $u_1(x)$ is either $f(x)$ or

$$\alpha \max f(y) = (.9)10 = 9$$

whichever is large. So

$$u_1(x) = (0, 10, 9, 9, 10, 0)$$

The vector Pu_1 is given by taking the average of the numbers on both sides:

Pu_1	αPu_1	$f(x)$	u_2
0	0	0	0
<10	<10	10	10
9.5	8.55	6	8.55
9.5	8.55	8	8.55
<10	<10	10	10
0	0	0	0

Pu_2	αPu_2	$f(x)$	u_3
0	0	0	0
<10	<10	10	10
9.275	8.3475	6	8.3475
9.275	8.3475	8	8.3475
<10	<10	10	10
0	0	0	0

Now you see that only the middle two coordinates of u change:

$$u_n = (0, 10, z, z, 10, 0)$$

The middle number converges:

$$\begin{array}{l} 8.3475 \\ 8.2564 \\ 8.2154\cdots \\ 8.1818 \end{array}$$

The important point is that it never goes below 8. So, the optimal strategy is to keep playing until you reach $x = 1$ or 4 and get $f(x) = 10$.

During the game you always have a 1/2 chance of stopping on the next turn. Your probability of stopping in two turns is 1/4, in 3 turns it is 1/8 and so on. Each time you play the final payoff decreases in value by a factor of $\alpha = .9$ So, your expected payoff at 2 or 3 is

$$\begin{aligned} v(2) = v(3) &= \frac{1}{2}\alpha \cdot 10 + \frac{1}{4}\alpha^2 \cdot 10 + \frac{1}{8}\alpha^3 \cdot 10 + \cdots \\ &= 10\left(\frac{\alpha}{2} + \left(\frac{\alpha}{2}\right)^2 + \left(\frac{\alpha}{2}\right)^3 + \left(\frac{\alpha}{2}\right)^4 + \cdots\right) \\ &= \frac{\text{first term}}{1 - \text{ratio}} = \frac{5\alpha}{1 - \alpha/2} = \frac{10\alpha}{2 - \alpha} = \frac{9}{1.1} \\ &= 8\tfrac{2}{11} = 8\tfrac{18}{99} = 8.181818\cdots \end{aligned}$$

So,

$$v = (0, 10, 8\tfrac{2}{11}, 8\tfrac{2}{11}, 10, 0).$$

Homework 4
Optimal Stopping Time

Two problems due Wednesday, March 19. Answers will be posted the following week.

Quiz 2 postponed until Thusday, March 20.

4.6 with the following additional questions.

(c) Suppose there is a discount α but no cost. What is the smallest value of α so that you should stop no matter what you get?

(d) Suppose that $g(x) = 5$ and $\alpha = .8$ Then what is the optimal strategy? Calculate the value function $v(x)$.

Hint: The iteration algorithm should start with:

$$u_1(x) = \begin{cases} 0 & \text{if recurrent} \\ f(x) & \text{if } f(x) \geq \max(\alpha f(y) - g(x)) \\ \max(\alpha f(y) - g(x)) & \text{otherwise} \end{cases}$$

Second problem: This is random walk with absorbing wall on the left and reflecting wall on the right.

$x =$	0	1	2	3	4	5	6
$f(x) =$	0	1	2	5	6	21	19
$g(x) =$	22	2	2	2	1	1	1

a) Draw a graph of $f(x)$ and connect the dots.

b) Find the optimal strategy and value function $v(x)$ if there is no cost. Describe geometrically what is the algorithm when the right wall is reflecting. [Hint: consider the mirror image of the function on the reflecting wall.]

c) Find the optimal strategy and value function if the cost $g(x)$ is given as above.

d) Find the optimal strategy and value function if there is a discount $\alpha = .9$ and no cost.

e) Find the optimal strategy and value function if there is a discount $\alpha = .9$ and cost $g(x)$ as given above.