

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 :

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

It satisfies the equation:

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

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.

$$\begin{array}{rcccccccccccc} p = & 1 & 2 & 3 & 4 & 5 & 6 & 5 & 4 & 3 & 2 & 1 & \cdot 1/36 \\ x = & 2 & 3 & 4 & 5 & 6 & \mathbf{7} & 8 & 9 & 10 & 11 & 12 \\ f(x) = & 2 & 3 & 4 & 5 & 6 & \mathbf{0} & 8 & 9 & 10 & 11 & 12 \\ g(x) = & 2 & 2 & 2 & 2 & 1 & \mathbf{12} & 1 & 1 & 1 & 1 & 1 \end{array}$$

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 $v(7) = 0$):

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

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

$$E_1, E_2, E_3, \dots$$

All you need is the single number

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

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 :

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

Repeating this process you get:

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

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

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

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

$$36E = 15E + 140$$

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

So, the value function is the vector:

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

With cost

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

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

$$E_3 = 6.096$$

$$E_4 = 5.960$$

...

$$E_{10} = 5.939393$$

$$E_{11} = 5.939393$$

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,

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

$$E = 196/33 = 5\frac{31}{33} = 5\frac{93}{99} = 5.939393 \dots$$

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

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

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

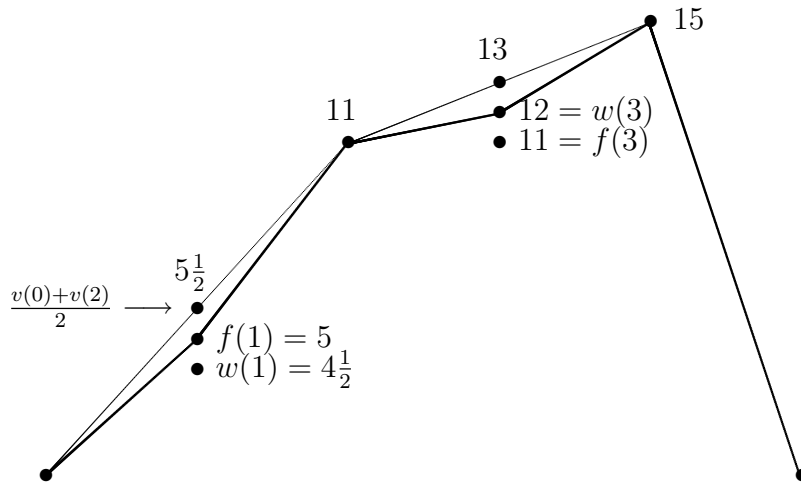
In the case of the random walk, this part \uparrow is what we have before. So, $v(x)$ is the smallest function so that

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

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

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:

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

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.