Optimal Stopping Group Project Report

MATH-339SP-01 Stochastic Processes

Professor Tim Chumley

May 7th, 2018

Weijia Bao, Jessica Feng, Sunan Qian

Contents

1. Optimal Stopping of Markov Chains

1.1 Overview

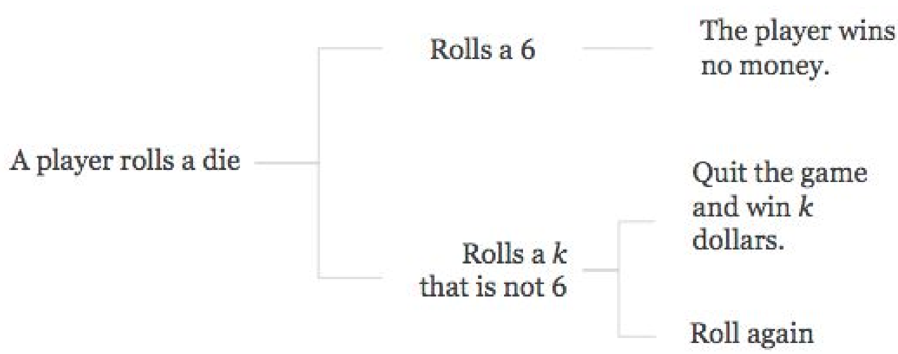
The theory of optimal stopping solves the type of problems where we need to choose the right time to take action in order to maximize the payoff or minimize the cost. It also has many real-world applications in different areas like economics and the secretary/marriage problem is one of the most famous. In order to further study this problem, we first start with an intuitional example.

1.2 An Intuitional Problem

Suppose a gambler plays a game and she rolls a die. If she rolls a 6, her payoff is 0 and she has to quit the game. If she rolls 1, 2, … 5, she can choose to quit and leave with 1, 2, … 5 dollars respectively, or roll again. The goal is to maximize the payoff. Some questions we need to answer are:

(1) What is the optimal strategy of the game? That is to say, when should she stop playing to ensure that she can maximize the probability that she gains the largest amount of money?
(2) What is the expected payoff the gambler will get if she follows the optimal strategy?

Therefore, the game continues until either a 6 is rolled or the gambler chooses to quit the game.



To find the optimal strategy, we need to first define several variables:

(1) ***f*(*k*):** the payoff associated with each roll

$f(k) = k$ if $k \leq 5$. For example, $f(1) = 1, f(2) = 2, \ldots f(5) = 5$
$f(k) = 0$ if $k = 6$.
$f = \{1, 2, 3, 4, 5, 0\}$

The total payoff of the game is always k dollars, with k being the value of the last roll.

(2) ***v*(*k*):** the expected winnings of the gambler given that the first roll is k, assuming that she

takes the optimal strategy (even though we still don't know what it is)

When the first roll is 6, $v(6) = 0$, because the player has to quit from here and there will not be another round.
When the first roll is 5, $v(5) = 5$, since the maximum amount of money she can get is 5 dollars and it doesn't pay for her to roll again.
When $k \leq 4$, we still don't know what $v(k)$ is for now.

(3) ***u*(*k*):** the expected payoff if the gambler does not stop after rolling a k ($k \leq 5$), but from then on follows the optimal strategy when she plays.

Therefore, since the probability of rolling 1 to 6 is each $\frac{1}{6}$, we can get

$$u(k) = \frac{1}{6}(v(1) + v(2) + v(3) + v(4) + v(5) + v(6))$$

Thus, we can generate the following optimal strategy:
If $f(k) > u(k)$, the gambler should stop and take the money.
If $f(k) < u(k)$, the gambler should roll again.
Hence we know that $v(k) = \max\{f(k), u(k)\}$, since the expected winnings given that the first roll is k is the bigger one between the payoff if the gambler stops and if she does not.

Since we have $u(k) = \frac{1}{6}(v(1) + v(2) + v(3) + v(4) + v(5) + v(6))$ and v(k) is always greater than f(k), we get

$$u(k) \geq \frac{f(1) + ... + f(6)}{6} = \frac{1 + ... + 5 + 0}{6} = \frac{5}{2}, fork \leq 5$$

Now we take the conclusion and see what happens for each situation when the first roll is 1, 2, 3 and 4. There is no need to discuss about the first roll being 5 or 6, since the gambler will quit the game from here.

If the first roll is 1, $v(1) = \max\{f(1), u(1)\} = \max\{1, u(1)\} = u(1)$, because $u(k)$ is always greater than 2.5.

If the first roll is 2, $v(2) = \max\{f(2), u(2)\} = \max\{2, u(1)\} = u(2)$. Therefore, the gambler should roll again if the first roll is 1 or 2.

If the first roll is 3, $v(3) = \max\{f(3), u(3)\}$. Suppose the gambler rolls again whenever a 3 comes up but follows the optimal strategy otherwise, then we get

$$u(3) = P(roll \leq 3)u(3) + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 = \frac{1}{2}u(3) + \frac{2}{3} + \frac{5}{6}$$

By solving the equation, we get $u(3) = 3$, which equals $f(3)$. In this case, the gambler can either choose to stop or continue playing because the amount of payoff will be the same.

If the first roll is 4, $v(4) = \max\{f(4), u(4)\}$. Suppose the optimal strategy is to roll another round, then the game would continue until the gambler rolls a 5 or 6 with equal probability.

$$u(4) = \frac{f(5) + f(6)}{2} = \frac{5 + 0}{2} < f(4) = 4$$

Thus, it doesn't pay to continue rolling and the gambler should stop at 4.

The general optimal strategy for this intuitional game is to roll again when the first roll is 1 or 2, stop when the first roll is 4, 5, or 6, and either continue or stop when the first roll is 3. In order to generalize an optimal strategy for similar problems, we develop an algorithm.

2. The Algorithm, Example and Intuition behind

2.1 Overview of Algorithm

To generalize the analytical idea that we talked above, there is an algorithm that finds v(x) iteratively. The main goal is to find the smallest superharmonic function with respect to P (the transition matrix) that is greater or equal to f.

We mentioned before, $v(x) = max\{u(x), f(x)\}$, with u(x) being the expected payoff if we continue playing, and f(x) is the option that we stops playing and take the payoff associate with the roll. A superharmonic function is one that satisfies $u(x) \geq \mathbf{P}u(x)$. The textbook in reference[1] proved that every superharmonic function that is larger or equal to f is greater or equal to v. So v is the smallest superharmonic function that is greater or equal to f. Then we get $v(x) = inf\ u(x)$ where the infimum is over over all the superharmonic functions u with $u(x) \geq f(x)$. We then have the algorithm below to find v iteratively, and we will explain the intuition behind the algorithm once we illustrate with an example.

2.2 The Algorithm

$$u_1(x) = \begin{cases} f(x) & \text{if } x = absoribing\ state \\ max\{f(x)\} & \text{otherwise} \end{cases}$$

$$\mathbf{P}v(x) = \sum_{y \epsilon S} p(x,y)v(y)$$

$$\mathbf{P}\ is\ the\ transition\ matrix$$

$$u_2(x) = max\{\mathbf{P}u_1(x), f(x)\}$$

$$u_n(x) = max\{\mathbf{P}u_{n-1}(x), f(x)\}$$

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

2.3 Example and Intuition

Let's consider the example that we talked before. Player rolls a die and quit and get $0 if the roll is 6. If roll is other numbers k for k=1,2..5, player can choose to take $k or roll again.

f is the payoff associate with the die, so f=(1 2 3 4 5 0). Follow the algorithm we get u1(6)=0 because we definitely quit and get 0 when roll is 6. u1(5)=5 because we quit and take

$5. The other entries in u1 will be max{f(x)}=5. So we have $u_1 = (5\ 5\ 5\ 5\ 5\ 0)$. Then we calculate u2 using $u_2(x) = max\{\mathbf{P}u_1(x),\ f(x)\}$.

$$u_2(1) = max\{\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 0,\ f(1) = 1\} = max\{\frac{25}{6}, 1\} = \frac{25}{6}.$$

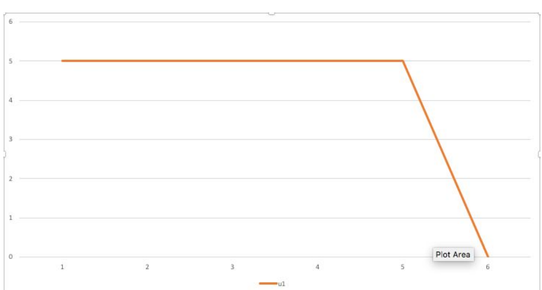
If we continue doing this for all entries in u2, we can get $u_2 = (\frac{25}{6}\ \frac{25}{6}\ \frac{25}{6}\ \frac{25}{6}\ 5\ 0)$.
Then follow the algorithm iteratively with $u_n(x) = max\{\mathbf{P}u_{n-1}(x), f(x)\}$, we get
$u_9 = (3.004\ 3.004\ 3.004\ 4\ 5\ 0)$, $u_{10} = (3.002\ 3.002\ 3.002\ 4\ 5\ 0)$

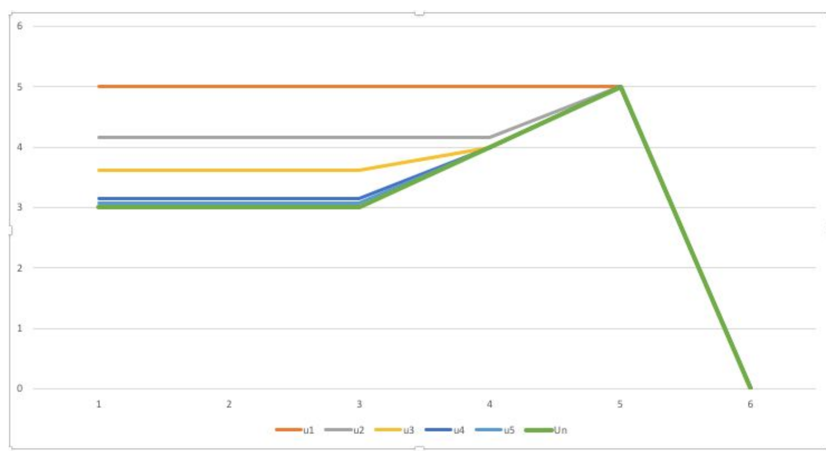
Eventually we have $v = \lim_{n\to\infty} u_n = (3\ 3\ 3\ 4\ 5\ 0)$.

The expected payoff for roll 1 and 2 are 3, which is larger than f(1) and f(2), so we must have chosen u(1) instead of f(x), which we continues to roll. For roll 3,4,5,6, v=f, so we quit and take the payoff associate with the roll.

Intuition:

The idea of this algorithm is first to make a big guess that satisfies u1 is superharmonic and
u1 ≥ f, which looks like the graph on the left. u1 is significantly larger than f, which is like an upper bound for v. We then use the algorithm to calculate iteratively until u_n converges as $n \to \infty$, which is the v we want to get that is the smallest superharmonic function that is larger or equal to f . Orange line is u1, grey is u2, yellow is u3, blue is u4 and u5, and green is u_n, which is v.





3. Optimal Stopping with Cost and Discounting

With Cost:

If we modify the original problem to that we need to pay $1 with each additional roll, what is the optimal strategy? This kind of problem gets harder analytically, but luckily we can still apply the same algorithm with a little modification.

g(x): cost associate with each state. In this problem g=(1 1 1 1 1 0) since we are forced to quit if roll is 6.

We then include g(x) in our algorithm, which in fact all we need to do is subtract the cost in the option that we continues to play, $\mathbf{P}u_{n-1}(x) - g(x)$.

$$u_1(x) = \begin{cases} f(x) & \text{if } x = absorbing\ state \\ max\{f(x)\} & \text{otherwise} \end{cases}$$

$$u_2(x) = max\{f(x), \mathbf{P}u_1(x) - g(x)\}$$

$$u_n(x) = max\{f(x), \mathbf{P}u_{n-1}(x) - g(x)\}$$

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

With this algorithm we get $u_{10} = (1.6\ 2\ 3\ 4\ 5\ 0)$. Only the first entry is larger than the one in f, so we should only continues playing when roll is 1 and stop for other rolls.

This is reasonable since there is a cost associated with continue playing, making us leaning towards stopping and take the money, instead of pay to play another round.

With Discounting:

Another modified version is that the value of money is decreasing, like inflation. This should make us want to stop for more rolls and take the money before the value decrease. Similarly we can make a simply change in the algorithm to show this.

$\alpha : \alpha$ is the discounting factor with $\alpha < 1$.

If $\alpha = .8$, then \$1 now will only worth \$.8 after one time unit.

In the algorithm, we just need to multiply the discounting factor to the option of continue playing. We get $u_{10} = (\frac{24}{11}\ \frac{24}{11}\ 3\ 4\ 5\ 0)$.

So we should continue playing when roll is 1 or 2, and stop otherwise.

$$u_1(x) = \begin{cases} f(x) & \text{if } x = absorbing\ state \\ max\{f(x)\} & \text{otherwise} \end{cases}$$

$$u_2(x) = max\{f(x), \alpha\mathbf{P}u_1(x)\}$$

$$u_n(x) = max\{f(x), \alpha\mathbf{P}u_{n-1}(x)\}$$

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

4. Secretary Problem

Problem:

You want to hire the best secretary from N candidates. You can either make an offer or you move on. If you don't make an offer, no going back. Once you make an offer, the game stopes. We want to find the selection strategy that will maximize the probability of selecting the candidate that is, in fact, best out of all potential N. The optimal stopping point is $\frac{1}{e} \approx 0.37$. Reject the first 37%, then take first one that is "best so far".

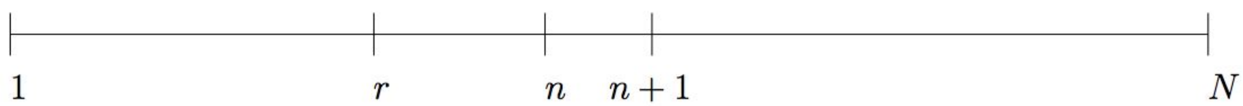
Solution:

Let r be the last applicant you will see before you actually start considering hiring anyone (the last one you're going to reject no matter what). Let the best applicant of all N, i $*$ occur arbitrarily through 1 to N, and N is the total number of applicants you have the potential to interview.

$$P(success) = P(r)$$

We will now show that the optimal solution is found by optimizing $P(r)$ by the standard route of solving: $P'(r) = 0$

The diagram below can help visualize the problem.



$P(k) = \Sigma P(being\ in\ position\ n) \times P(being\ selected\ given\ in\ position\ is\ the\ best\ candidate)$

(1) $n \leq r$, the best occurs within the first r candidates

The probability being in position i $*$ is $\frac{1}{N}$, and the probability of choosing this candidate given n is the best candidate is 0 (you reject, no matter what). Sum over all possible $n \leq r$, so we have

$$P(n \leq r) = \sum \frac{1}{N} \times 0 = 0$$

(2) $n > r$, the best occurs after r candidates

Similarly, the probability being in position i $*$ is $\frac{1}{N}$, and the the probability of choosing this candidate given n is the best candidate $= 1 - P(not\ choosing\ this\ candidate)$

which equals to $1 - \frac{i}{n+i} = \frac{n}{n+i}$, where $n+i$ goes to $N-1$. Sum over all possible $n > r$, so we have

$$P(n > r) = \frac{1}{N}\left[\frac{r}{r} + \frac{r}{r+1} + \frac{r}{r+2} + \ldots + \frac{r}{N-1}\right] = \frac{r}{N}\sum_{n=r}^{N-1}\frac{1}{n}$$

Adding up (1) and (2),

$$P(r) = P(n \leq r) + P(n > r) = \frac{r}{N}\sum_{n=r}^{N-1}\frac{1}{n}$$

P(r) is in fact a Riemann approximation to an integral. By inspecting the expression in the limit as N → ∞, letting $lim_{N\to\infty}\frac{r}{N} = x$, and $lim_{N\to\infty}\frac{n}{N} = T$ we find the following:

$$P(r) = \lim_{N\to\infty}\frac{r}{N}\sum_{n=r}^{N-1}\frac{N}{n}\frac{1}{N} = x\int_x^1 \frac{1}{t}dt = -x\ln x$$

As N grows infinitely large, we find that the ratio of applicants reviewed and rejected to the number of total applicants approaches x. By solving $P'(r) = 0$ for r, we have the optimal ratio and the probability of success $P_(r_{optimal})$

$$P'(r) = -\ln x - 1 = 0 \Rightarrow x = \frac{1}{e}$$
$$P(\frac{1}{e}) = \frac{1}{e}$$
$$\frac{1}{e} \approx 0.37$$

The ratio of r to N is optimal at $\frac{1}{e}$ outputting a probability of success of, coincidentally, $\frac{1}{e}$ as well. So for N >> 1 the $r_{optional}$ is nearly $\frac{N}{e}$, otherwise it can be found by computing P(r) directly.

$\square$

5. References

[1] Lawler, Gregory F. "Introduction to Stochastic Processes."

https://www.math.ucla.edu/~tom/Stopping/sr1.pdf. Accessed 25 March 2018.

[2] The secretary problem solution details.

https://thebryanhernandezgame.files.wordpress.com/2010/05/secretary-problem.pdf.

Accessed 25 March 2018.