

# APMA 1930X: Probability, Optimization, and Stochastic Calculus

Milan Capoor

Fall 2024

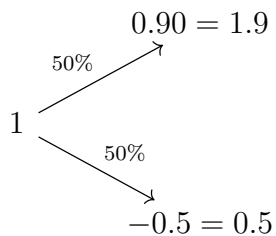
# Chapter 1

## Introduction and Examples

Sept 04

### Kelly Betting

Imagine a game where

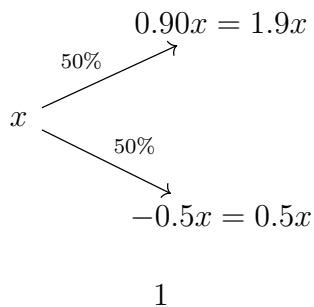


So our expected payoff is

$$\mathbb{E}(\text{payoff}) = 0.5 \cdot 0.9 + 0.5 \times (-0.5) = 0.2$$

so the game is favorable to you.

Now consider the game

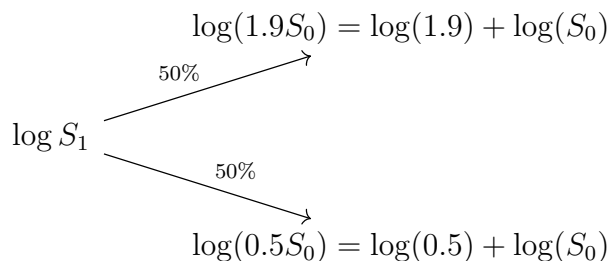


The expected payoff is

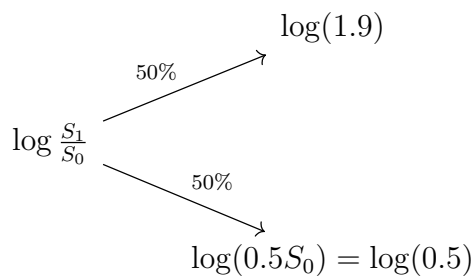
$$\mathbb{E}[\text{payoff}] = 0.2x$$

which is decidedly unfavorable.

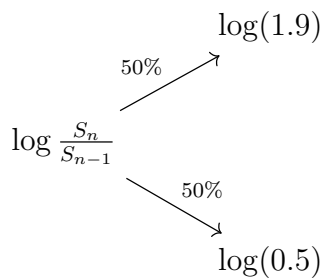
Now betting all-in is not a good idea. Suppose your current wealth is  $S_0$  and wealth after one bet is  $S_1$ . Consider the log of the second scenario:



so



Which tells us that after  $n$  bets



Suppose that each bet is iid with  $x_i = \log \frac{S_i}{S_{i-1}}$  with  $p(x_i = \log 0.5) = 0.5$  and  $p(x_i = \log 1.9) = 0.5$ . Then

$$\log \frac{S_n}{S_{n-1}} + \log \frac{S_{n-1}}{S_{n-2}} + \cdots + \log \frac{S_1}{S_0} = X_n + X_{n-1} + \cdots + X_1$$

hence

$$\log S_n = \log S_0 + (X_1 + \cdots + X_n)$$

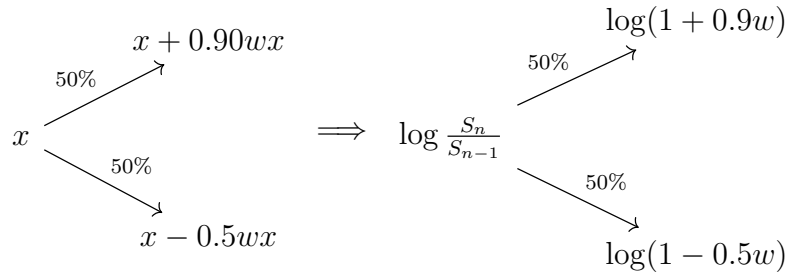
By LLN

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_i x_i &= \mathbb{E}[X_1] \\ &= 0.5 \log 1.9 + 0.5 \log 0.5 \\ &= 0.5 \log(0.95) < 0 \end{aligned}$$

so  $S_n \rightarrow 0$  Clearly, all-in betting is not a good idea.

What if we bet a very small fraction of our wealth? By LLN, we will *eventually* realize our edge but we will have to wait a long time. Can we do better?

Now bet  $w$  where  $0 \leq w \leq 1$  Then



so

$$\mathbb{E}[\log \frac{S_n}{S_{n-1}}] = 0.5 \log(1 + 0.9w) + 0.5(\log 1 - 0.5w)$$

which we want to maximize for  $w \in [0, 1]$

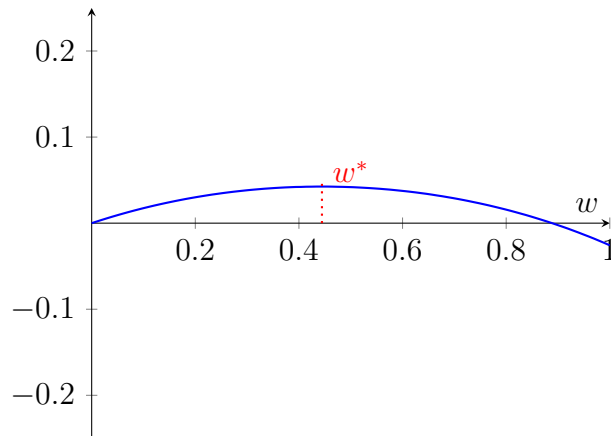
Taking the derivative, we get

$$\frac{0.9}{1 + 0.9w} - \frac{0.5}{1 - 0.5w} = 0 \implies 0.9 - 0.45w = 0.5 + 0.45w$$

so the best betting fraction is  $w^* = \frac{4}{9}$  and

$$\mathbb{E}(\log \frac{S_n}{S_{n-1}}) = 0.5 \log \frac{9.8}{9} \approx 0.05$$

We can graph our expected value function



## Probability Review

**Example:** Toss a fair coin  $p(H) = p(T) = 0.5$   $n$  times.

What is the probability of getting an even number of heads?

One method is to argue by partitioning the sample space.

Alternatively, we can argue from the first toss: Let  $x_n = p(\text{even \# of H after } n \text{ tosses})$ . Then after the first toss,

- if we got heads: we need an odd number of heads in the next  $n - 1$  tosses which happens with probability  $1 - x_{n-1}$
- if we got tails: we need an even number of heads in the next  $n - 1$  tosses which happens with probability  $x_{n-1}$

Therefore, by the Law of Total Probability,

$$x_n = 0.5(1 + x_{n-1}) + 0.5(x_{n-1}) = 0.5$$

We call this method **First step analysis**

## 1.1 Sept 06

Last time we ended with an example of a fair coin tossed  $n$  times. How can we compute

$$x_n = P(\text{even \# H with } n \text{ tosses})$$

**Method 1:** Make a partitioning argument, taking advantage of the equal probabilities of heads and tails.

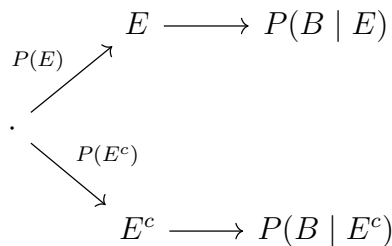
**Method 2:** First step analysis.

We argue that on the first toss, if we get heads, then the final probability we have an even number of heads is the same as the probability we have an odd number of heads after  $n - 1$  tosses. If we get tails, then the probability we have an even number of heads is the same as the probability we have an even number of heads after  $n - 1$  tosses.

Using the law of total probability, we get

$$x_n = \frac{1}{2}(1 - x_{n-1}) + \frac{1}{2}x_{n-1} = \frac{1}{2}$$

**The Law of Total Probability:** Consider the example



Then

**Law of total Probability:**

$$P(B) = P(E)P(B | E) + P(E^c)P(B | E^c)$$

*Proof:*

$$\begin{aligned} RHS &= P(E)P(B | E) + P(E^c)P(B | E^c) \\ &= P(E)\frac{P(B \cap E)}{P(E)} + P(E^c)\frac{P(B \cap E^c)}{P(E^c)} \\ &= P(B \cap E) + P(B \cap E^c) \\ &= P(B) \\ &= LHS \end{aligned}$$

Consider a scenario with a biased coin with  $p(H) = p$  and  $p(T) = q$  Now

$$\begin{aligned}x_n &= p(1 - x_{n-1}) + qx_{n-1} \\ &= p + (q - p)x_{n-1}\end{aligned}$$

We want a special solution to this equation which is a constant in  $n$ , i.e.  $x_n = c$

We know  $x_1 = P(T) = q$

Start with the guess  $c = \frac{1}{2}$  Then

$$\begin{aligned}x_n &= p + (1 - 2p)x_{n-1} \\ x_n - \frac{1}{2} &= p + (1 - 2p)x_{n-1} - \frac{1}{2} \\ &= (1 - 2p)(x_{n-1} - \frac{1}{2}) \\ &= (1 - 2p)(1 - 2p)(x_{n-2} - \frac{1}{2}) \\ &= \dots \\ &= (1 - 2p)^{n-1}(x_1 - \frac{1}{2}) \\ &= (1 - 2p)^{n-1}(p - \frac{1}{2})\end{aligned}$$

so

$$x_n = \frac{1}{2} + (1 - 2p)^{n-1}(q - \frac{1}{2}) = \frac{1}{2} + \frac{1}{2}(1 - 2p)^n$$

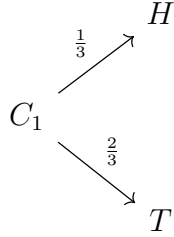
Thus, we can say the general method is 1) use first-step analysis to get a recursive formula and 2) guess a solution and prove it by induction.

**Example (Putnam 2001):** We have  $n$  coins  $C_1, C_2, \dots, C_n$  with  $P(C_k = H) = \frac{1}{2^{k+1}}$  What is the probability that the number of heads is odd after tossing all  $n$  coins?

Let's try the same approach as last time.

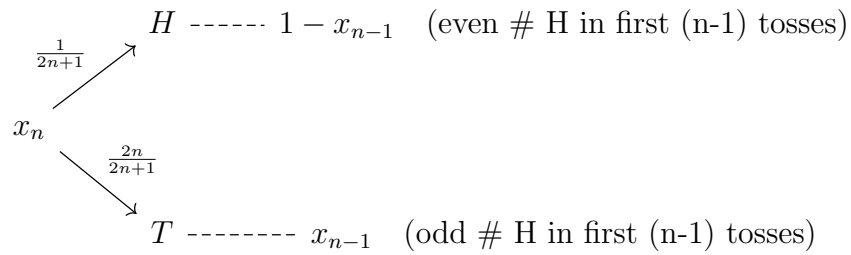
Let  $x_n = P(\text{odd \# of heads after } n \text{ coins})$  Then the first toss gives us

$$x_1 = P(C_1 = H) = \frac{1}{3}$$



But in fact this tells us nothing because the probability of getting an even number of heads in the next  $n - 1$  tosses is not  $1 - x_{n-1}$  but some function of  $\frac{1}{2^{n+1}}$

We actually need to start from the last toss:



So

$$\begin{aligned} x_n &= \frac{1}{2n+1}(1 - x_{n-1}) + \frac{2n}{2n+1}x_{n-1} \\ &= \frac{1}{2n+1} + \frac{2n-1}{2n+1}x_{n-1} \end{aligned}$$

Again guessing  $x_n = \frac{1}{2}$ ,

$$\begin{aligned} x_n - \frac{1}{2} &= \frac{2n-1}{2n+1}(x_{n-1} - \frac{1}{2}) \\ &= \frac{2n-1}{2n+1} \frac{2n-3}{2n-1}(x_{n-2} - \frac{1}{2}) \\ &= \frac{2n-1}{2n+1} \frac{2n-3}{2n-1} \cdots \frac{3}{5}(x_1 - \frac{1}{2}) &= \cdots \\ &= \frac{3}{2n+1}(\frac{1}{3} - \frac{1}{2}) \\ &= -\frac{1}{2(2n+1)} \\ x_n &= \frac{1}{2} - \frac{1}{2(2n+1)} = \boxed{\frac{n}{2n+1}} \end{aligned}$$



**Example (Law of Total Probability for Continuous RV):** When we have finitely many possible outcomes, we can simply sum them by the LTP. What if we have infinitely many?

Consider a coin where  $P(H) \sim \text{Unif}[0, 1]$  If we toss the coin  $n$  times and define  $X = \# \text{ of H}$ , what is  $P(X = k), k = 0, 1, \dots, n$ ?

(Easier variation) Define  $\theta = P(H) \sim \text{Unif}[0, 1]$  Suppose we are given  $\theta = p$  Now

$$P(X = k \mid \theta = p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

For the continuous case, we have

$$P(X = k) = \int_0^1 P(X = k \mid \theta = p) \cdot (\text{density of } \theta \text{ at } p) dp$$

But the density of  $\theta$  at  $p$  is simply 1 (uniform!). So

$$P(X = k) = \int_0^1 \binom{n}{k} p^k (1 - p)^{n-k} dp$$

Let's do some examples:

$$P(X = 0) = \int_0^1 (1 - p)^n dp = \frac{1}{n + 1}$$

(hint: change of variables  $u = 1 - p$ )

$$\begin{aligned} P(X = 1) &= \int_0^1 np(1 - p)^{n-1} dp \\ &\stackrel{x=1-p}{=} \int_1^0 n(1 - x)x^{n-1} (-dx) \\ &= \int_0^1 nx^{n-1} - nx^n dx \\ &= 1 - n \frac{1}{n + 1} = \frac{1}{n + 1} \end{aligned}$$

$$P(X = n) = \int_0^1 p^n dp = \frac{1}{n+1}$$

So we conclude  $P(X = k) = \frac{1}{n+1}$

## 1.2 Sept 09

Last time we introduced the *continuous version of the Law of Total Probability*:

$$P(A) = \int P(A \mid X = x) f_X(x) dx$$

**Example:** Suppose you have a fair die ( $X \in 1 \dots 6$ ).

1. What is the average number of tosses until 6 appears?
2. What is the average cumulative sum ( $\sum x_i$ ) until 6 appears?

Every time we toss the die, we know  $P(6) = \frac{1}{6}$  and  $P(\neg 6) = \frac{5}{6}$ . If  $X$  is the number of tosses until 6, we know

$$X \sim \text{geometric}(p) \implies \mathbb{E}[X] = \frac{1}{p}$$

Which we can see from

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=1}^n P(X = i) \cdot i \\ &= \sum_{i=1}^n (1-p)^{i-1} p \cdot i \\ &\vdots \\ &= \frac{1}{p} \end{aligned}$$

This method presents some difficulties in general: How do we explicitly know  $P(X = n)$ . Further, this method tells us nothing about question 2.

Let's try again with FSA: Let  $X$  be the average number of tosses until 6.

Look at the first toss:

| 1st Toss              |    |   |
|-----------------------|----|---|
| Outcome<br>first toss | on | Average num-<br>ber of tosses<br>to reach 6 |
| 1                     |    | $x$   |
| 2                     |    | $x$   |
| 3                     |    | $x$   |
| 4                     |    | $x$   |
| 5                     |    | $x$   |
| 6                     |    | 0   |

So

$$X = 1 + \frac{5}{6}x + \frac{1}{6} \cdot 0 = 1 + \frac{5}{6}X \implies \boxed{X = 6}$$

(the 1 is the first toss we already used)

Now let  $Y$  be the average cumulative sum until 6. We can draw exactly the same table because for each case the average depends on all the rest:

| 1st Toss              |    |   |
|-----------------------|----|---|
| Outcome<br>first toss | on | Average num-<br>ber of tosses<br>to reach 6 |
| 1                     |    | $Y$   |
| 2                     |    | $Y$   |
| 3                     |    | $Y$   |
| 4                     |    | $Y$   |
| 5                     |    | $Y$   |
| 6                     |    | 0   |

so

$$\begin{aligned}
 Y &= \frac{1}{6}(1 + y) + \frac{1}{6}(2 + y) + \cdots + \frac{1}{6}(5 + y) + \frac{1}{6}(6 + 0) \\
 &= \frac{5}{6}y + \frac{7}{2} \\
 &= 21
 \end{aligned}$$

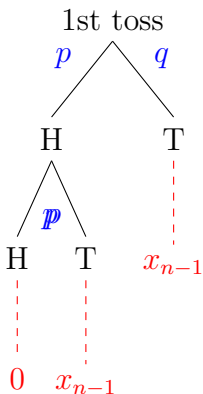
**Example:** Toss a coin  $n$  times and let  $P(H) = p$ . What is the probability that the pattern  $HH$  does not appear?

Let  $x_n$  be the probability that  $HH$  does not appear after  $n$  tosses. Then we know

$$\begin{aligned}x_1 &= 0 \\x_2 &= 1 - p^2\end{aligned}$$

but the rest become very complicated.

For simplicity, let  $P(T) = q = 1 - p$ .



So

$$x_n = qx_{n-1} + pqx_{n-2}, \quad x_1 = 1, x_2 = 1 - p^2$$

We can solve this using a method similar to 2nd order ODEs:

$$f'' + af' + bf = 0 \implies r^2 + ar + b = 0 \implies f = C_1 e^{r_1 t} + C_2 e^{r_2 t}$$

In our case

$$x_n = qx_{n-1} + pqx_{n-2} \implies r^2 = qr + pq$$

We can find roots  $r_1, r_2$  and then solve for  $C_1, C_2$  using the initial conditions, giving our solution:

$$\begin{cases} x_n = Ar_1^n + Br_2^n \\ x_1 = 1, x_2 = 1 - p^2 \end{cases}$$

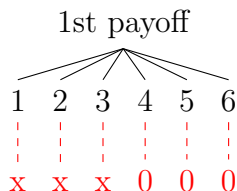
**Example:** We are playing a game of dice with the following rules:

1. If the toss is 1, 2, 3, you get the face value and continue
2. If the toss is 4, 5, you get the face value and the game ends

3. If the toss is 6, the game ends

*Example game:*  $3 \rightarrow 2 \rightarrow 2 \rightarrow 4$  gives score 11

What is the average payoff  $x$ ?



So

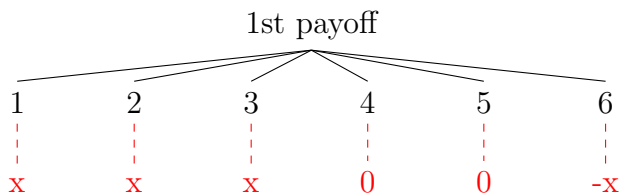
$$x = \frac{1}{6}(1 + x) + \frac{1}{6}(2 + x) + \frac{1}{6}(3 + x) + \frac{1}{6}(4 + 0) + \frac{1}{6}(5 + 0) + \frac{1}{6}(0) = 5$$

## 1.3 Sept 11

**Harder Variation:** now the game rules are

1. If the toss is 1, 2, 3, you get the face value and continue
2. If the toss is 4, 5, you get the face value and the game ends
3. If the toss is 6, the game ends and you lose all your points
4. You can stop the game whenever you want

Let  $y$  be the average cumulative sum. Now



Recall that in the previous case where the game ended on 6 but you retained your score,

$$x = \frac{1}{6}(1 + x) + \frac{1}{6}(2 + x) + \frac{1}{6}(3 + x) + \frac{1}{6}(4 + 0) + \frac{1}{6}(5 + 0) + \frac{1}{6}(0) = 5$$

However, in the new game this is not quite right: we are not guaranteed to keep our early points if we continue. What is the probability we keep our points?  $2/3$  because out of the three ending conditions, we keep our points in two of them.

So

$$y = \frac{1}{6}\left(\frac{2}{3} + y\right) + \frac{1}{6}\left(\frac{4}{3} + y\right) + \frac{1}{6}(2 + y) + \frac{1}{6}(4) + \frac{1}{6}(5) \implies \boxed{y = \frac{13}{3}}$$

But this really only works because the stopping condition is simple.

**Example (Prison Dilemma):** 100 prisoners are given a chance to play a game for their freedom. There are 100 boxes that each contain a paper with a number from 1 to 100. Each prisoner has a number 1-100 and is allowed to open 50 boxes. If one of these 50 boxes contain his number, he wins. If all 100 prisoners win, they all go free. If anybody loses, they are all executed.

We know that if each prisoner opens boxes at random, the probability of winning is  $\frac{1}{2^{100}}$ . Can they do better?

Here is one strategy, though we can't say if it is optimal:

1. Each player opens the box corresponding to their number
2. If they don't win, they open the box corresponding to the number on the paper in the box they just opened
3. Repeat until they win or they have opened 50 boxes

We are looking for cycles in the graph. If the cycle has length at most 50, everyone in the cycle wins. If the cycle is longer than 50, everyone loses.

As cycles are disjoint, we can partition the space. Then everyone will win if all the cycles are of length at most 50 so

$$\begin{aligned} \mathbb{P}(\text{win}) &= \mathbb{P}(\text{every cycle} \leq 50) \\ &= 1 - \mathbb{P}(\text{at least one cycle} > 50) \\ &= 1 - \mathbb{P}(\text{exactly one cycle} > 50) \end{aligned}$$

Let's look more generally: for  $2n$  boxes what is the probability one cycle is of length at least  $n$ ?

$$\mathbb{P}(\text{one cycle} > n) = \sum_{k=n+1}^{2n} \mathbb{P}(\text{one cycle} = k)$$

We can view the contents of the boxes in order as a permutation of  $2n$  elements. We know there are  $(2n)!$  permutations. How many give a cycle of length  $k$ ?

Clearly, we want to choose  $k$  elements to be in the cycle and then permute the rest. So

$$\binom{2n}{k}(2n-k)!$$

But how many ways are there to make a cycle with  $\binom{2n}{k}$  elements?  $(k-1)!$

All together,

$$\begin{aligned}\mathbb{P}(\text{one cycle} = k) &= \frac{1}{(2n)!} \binom{2n}{k} (2n-k)!(k-1)! \\ &= \frac{1}{(2n)!} \frac{(2n)!}{(2n-k)!k!} (2n-k)!(k-1)! \\ &= \frac{(k-1)!}{k!}\end{aligned}$$

So

$$\mathbb{P}(\text{win}) = 1 - \sum_{k=n+1}^{2n} \frac{1}{k}$$

We can approximate

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \log n + C$$

for a constant  $C$  so

$$\begin{aligned}\mathbb{P}(\text{win}) &= 1 - \sum_{k=n+1}^{2n} \frac{1}{k} \\ &= 1 - \left( \sum_{k=1}^{2n} \frac{1}{k} - \sum_{k=1}^n \frac{1}{k} \right) \\ &\approx 1 - (\log 2n + C - (\log n + C)) \\ &= 1 - \log 2 \approx 0.3\end{aligned}$$

## 1.4 Sept 13:

### Law of Total Expectation

**Example:** In America, 50.5% of the population is female, 49.5% is male. The average life expectancy is 79.1 for women and 73.2 for men. What is the average life expectancy of an American?

$$\begin{aligned}\frac{1}{N} \sum_{\text{Americans}} \text{life expectancy} &= \frac{1}{N} \sum_{\text{Women}} \text{life expectancy} + \frac{1}{N} \sum_{\text{Men}} \text{life expectancy} \\ &= \frac{1}{N} (0.505N \cdot 79.1) + \frac{1}{N} (0.495N \cdot 73.2) \\ &= 76.2\end{aligned}$$

or, in general,

$$P(A) = \sum_y P(A \mid Y = y)P(Y = y)$$

This is incredibly intuitive but let's generalize. Say  $Y$  is the average life expectancy of an American and  $X$  is their sex. Let  $X = 1$  if woman,  $X = 0$  if man. Then the situation is expressed

$$\begin{cases} \mathbb{P}(X = 1) = 50.5\%, & \mathbb{P}(X = 0) = 49.5\% \\ \mathbb{E}[Y \mid X = 1] = 79.1, & \mathbb{E}[Y \mid X = 0] = 73.2 \end{cases}$$

Therefore, if we define a random variable  $\mathbb{E}[Y \mid X]$  such that it takes value  $\mathbb{E}[Y \mid X = 1]$  on event  $X = 1$  and  $\mathbb{E}[Y \mid X = 0]$  on event  $X = 0$ , then

$$\mathbb{E}[Y] = \mathbb{E}[Y \mid X = 1]\mathbb{P}(X = 1) + \mathbb{E}[Y \mid X = 0]\mathbb{P}(X = 0) = \mathbb{E}[\mathbb{E}[Y \mid X]]$$

For  $Y$  discrete,

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]] = \sum_y \mathbb{E}[X \mid Y = y]\mathbb{P}(Y = y)$$

We call  $\mathbb{E}[Y \mid X]$  the *conditional expectation* of  $Y$  given  $X$ .

The Law of Total Expectation is a Generalization of the law of total probability.



Define for all  $A \subseteq \Omega$ , the *indicator random variable* by

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Then, straightforwardly,  $\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A]$  and the Law of Total Probability is

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{P}(A \mid X)]$$

for any family of random variables  $X$ .

When  $X$  is a single random variable:

$$\begin{aligned} \mathbb{P}(A) &= \sum_i \mathbb{P}(A \mid X = x_i) \mathbb{P}(X = x_i) \\ \mathbb{P}(A) &= \int_{\mathbb{R}} \mathbb{P}(A \mid X = x) f_X(x) dx \end{aligned}$$

depending on whether  $X$  is discrete or continuous.

**Example:** Sample  $X_1, X_2, \dots, X_n \sim \text{Unit}[0, 1]$ . Let  $X_{(1)} = \min\{X_i\}$ ,  $X_{(2)} = \min\{X_i \mid X_i \neq X_{(1)}\}$ , etc.

Then  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  is a rearrangement of  $\{X_1, X_2, \dots, X_n\}$  in ascending order, what we call *the order statistics*

These  $n$  points now partition  $[0, 1]$  into  $(n + 1)$  pieces. What is the average length of each piece?

It suffices to find the density of  $X_{(1)}$  and compute  $\mathbb{E}[X_{(1)}]$  and then compute the average length of each subinterval.

First, we calculate the CDF:

$$F_1(x) = \mathbb{P}(X_{(1)} \leq x) = 1 - \mathbb{P}(X_{(1)} > x) = 1 - (1 - x)^n$$

Then the density is

$$f_1(x) = F_1'(x) = n(1 - x)^{n-1} \cdot \mathbb{1}_{[0,1]}(x)$$

and

$$\begin{aligned} E[X_{(1)}] &= \int_{\mathbb{R}} x f_1(x) dx \\ &= \int_0^1 x n (1-x)^{n-1} dx \\ &= \int_0^1 n(1-t)t^{n-1} dt \\ &= \int_0^1 n(t^{n-1} - t^n) dt \\ &= 1 - \frac{n}{n+1} \\ &= \frac{1}{n+1} \end{aligned}$$

The first piece on  $[0, 1]$  is  $[0, X_{(1)}]$  which has average length  $\mathbb{E}[X_{(1)}] = \frac{1}{n+1}$ .

What about the second piece?

$$\begin{aligned} \mathbb{E}[X_{(2)}] &= \mathbb{E}[\mathbb{E}[X_{(2)} \mid X_{(1)}]] \\ &= \mathbb{E}\left[\frac{1}{n}(1 - X_{(1)})\right] \\ &= \frac{1}{n} - \frac{1}{n}\mathbb{E}[X_{(1)}] \\ &= \frac{1}{n} \left(1 - \frac{1}{n(n+1)}\right) \\ &= \frac{1}{n+1} \end{aligned}$$

Repeating this procedure, we can see that each interval has average length  $\frac{1}{n+1}$

## 1.5 Sept 9

### Optimization

**Classic Optimization:** Given a sequence or function, find its max/min

*Method:* Take the derivative and set it equal to zero.

**Constrained Optimization:** e.g. maximize  $f(x)$  given  $g(x) = 0$

*Method:* Lagrange Multipliers

**Infinite Dimensional Problems:** e.g.  $\max_f H(f)$  (where  $H$  is a function of functions or a *functional*)

*Method:* Calculus of Variations

**Example (Application of indicator RVs):** A company has  $n$  employees. Anyone can take the day off if it is the birthday of some employee. Every work day is 8 hours. How many employees should the company hire so that the total working hour from employees is maximized?

Denote  $a_n = \mathbb{E}[\text{total working hours with } n \text{ employees}]$ . We then want to find  $\max_n a_n$ .

Suppose  $X$  is the total number of working days from  $n$  employees. Then  $a_n = \mathbb{E}[8Xn] = 8n\mathbb{E}[X]$

But this distribution is very difficult to determine. Let's use indicator random variables. Let  $X_1, X_2, \dots, X_{365}$  be RVs defined by

$$X_k = \begin{cases} 1 & \text{if the } k\text{-th day is a working day} \\ 0 & \text{otherwise} \end{cases}$$

Then  $X = X_1 + X_2 + \dots + X_{365}$  is the total number of working days. So

$$\mathbb{E}[X] = \sum_{k=1}^{365} \mathbb{E}[x_k]$$

But  $x_k$  is a discrete RV so

$$\begin{aligned} \mathbb{E}[x_k] &= 1 \cdot P(x_k = 1) + 0 \cdot P(x_k = 0) \\ &= P(x_k = 1) \\ &= \mathbb{P}(k\text{-th day is a working day}) \\ &= \mathbb{P}(k\text{-th day is no one's birthday}) \\ &= \prod_{i=1}^n \mathbb{P}(k\text{-th day is not the birthday of employee } i) \\ &= \left(\frac{364}{365}\right)^n \end{aligned}$$

Therefore, our problem reduces to maximizing

$$a_n = 8n\mathbb{E}[X] = 8n \cdot 365 \left(\frac{364}{365}\right)^n$$

with respect to  $n$ .

Consider  $T_n = n \left(\frac{364}{365}\right)^n$ . Then

$$\frac{T_{n+1}}{T_n} = \frac{(n+1) \left(\frac{364}{365}\right)^n}{n \left(\frac{364}{365}\right)^n} = \frac{n+1}{n} \cdot \frac{364}{365}$$

Which is greater than 1 for  $n < 365$ ,  $< 1$  for  $n > 365$ , and equal to 1 for  $n = 365$ .

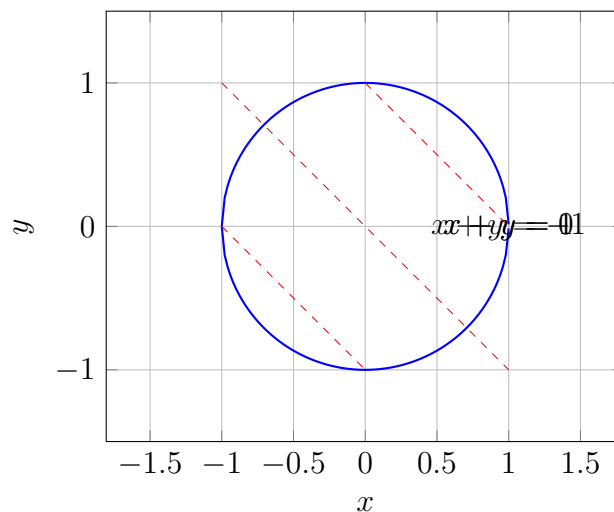
This tells us that  $T_n$  is increasing for  $n < 365$ , decreasing for  $n > 365$ , and constant for  $n = 364$  or  $365$ . Therefore, the optimal  $n^* = 364$  or  $365$ .

**Example (function maximization):**

1. max/min  $f(x)$  by solving  $f'(x^*) = 0$  for  $x^*$
2. max/min  $f(x, y)$  by solving  $f_x = f_y = 0$

**Example (Lagrange multipliers):** maximize  $x + y$  given  $x^2 + y^2 = 1$

Graphically:



We can easily solve this numerically, or we can use *Lagrange multipliers*.

Define

$$(x + y) - \lambda(x^2 + y^2 - 1) = H(x, y, \lambda)$$

so our problem involves maximizing  $H$  with no constraints. We can calculate

$$\begin{cases} H_x = 1 - 2\lambda x = 0 \\ H_y = 1 - 2\lambda y = 0 \\ H_\lambda = x^2 + y^2 - 1 = 0 \end{cases} \implies \begin{cases} x = y = \frac{1}{2\lambda} \\ x^2 + y^2 = 1 \end{cases} \implies x^* = y^* = \pm \frac{\sqrt{2}}{2}$$

**Example (Infinite dimensions):** Maximize  $\int_{-\infty}^{\infty} -f(x) \log f(x) dx$  over all non-negative functions  $f(x)$  such that

1.  $f$  is a probability density function
2.  $\int x f(x) dx = \mu$  for given  $\mu$
3.  $\int x^2 f(x) dx = \mu^2 + \sigma^2$  for given  $\sigma^2$

(the second equation is the condition  $\mathbb{E}[X] = \mu$  and the third is the condition  $\mathbb{E}[X^2] = \mu^2 + \sigma^2$  for  $X \sim f(x)$ )

Simply, we want the distribution with maximum entropy given mean and variance. Let's try to use Lagrange multipliers (as if it were finite dimensional)

Define

$$\begin{aligned} H(f, \lambda_1, \lambda_2, \lambda_3) = & \left( \int_{-\infty}^{\infty} -f(x) \log f(x) dx \right) - \lambda_1 \left( \int_{-\infty}^{\infty} f(x) dx - 1 \right) \quad (\text{from 1}) \\ & - \lambda_2 \left( \int_{-\infty}^{\infty} x f(x) dx - \mu \right) \quad (\text{from 2}) \\ & - \lambda_3 \left( \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 - \sigma^2 \right) \quad (\text{from 3}) \end{aligned}$$

and we subtract them off because we want each condition term to be zero.

We distribute

$$H = \int_{-\infty}^{\infty} [-f(x) \log f(x) - \lambda_1 f(x) - \lambda_2 x f(x) - \lambda_3 x^2 f(x)] dx + \lambda_1 + \lambda_2 \mu + \lambda_3 (\mu^2 + \sigma^2)$$

Maximizing with respect to the multipliers is easy. We just need to find a way to maximize the integral. Let  $t = f(x)$  and consider

$$\max_t (-t \log t - \lambda_1 t - \lambda_2 x t - \lambda_3 x^2 t)$$

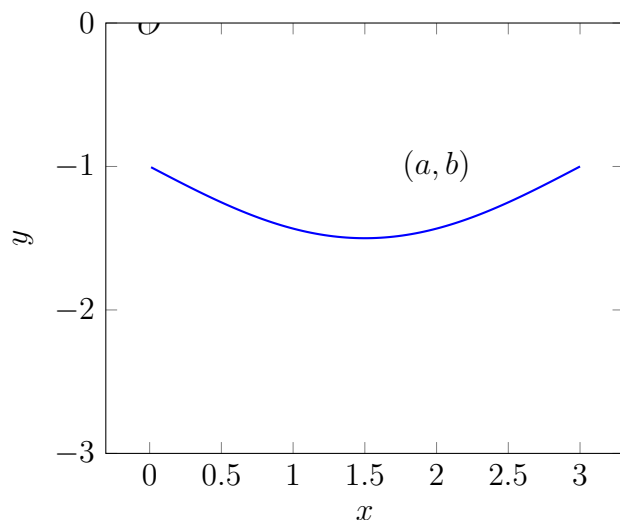
but we can just take the derivative and set to zero!

$$-\log t - 1 - \lambda_1 - \lambda_2 x - \lambda_3 x^2 = 0 \implies t = f(x) = e^{-1-\lambda_1-\lambda_2 x-\lambda_3 x^2}$$

## 1.6 Sept 18

### Calculus of Variations

**Example:** Suppose you have two points. Place a ball at the first point and let it roll to the second point. What shape should the path be to minimize the descent time?



Formally, let the curve be described by the function  $f$  such that  $f(0) = 0, f(a) = b$ . The descent time is given by  $L(f)$ . Our goal:

$$\min_f L(f)$$

Let  $u$  be any fixed function such that  $u(0) = 0, u(a) = 0$ .

Let  $\varepsilon$  be any small real number and define a *perturbation*

$$\begin{aligned} f_\varepsilon(x) &= f(x) + \varepsilon u(x) \\ f_\varepsilon(0) &= f(0) + \varepsilon u(0) = 0 \\ f_\varepsilon(a) &= f(a) + \varepsilon u(a) = b \end{aligned}$$

Suppose  $f^*(x)$  is the fastest descent curve. Then

$$L(f^*) \leq L(f_\varepsilon^*) = L(f^* + \varepsilon u)$$

The RHS is a function of  $\varepsilon$  so we can optimize it with respect to  $\varepsilon$ :

$$\left. \frac{d}{d\varepsilon} L(f^* + \varepsilon) \right|_{\varepsilon=0} = 0$$

From Physics for our rolling descent problem,

$$L(f) = \int_0^a \sqrt{\frac{1 + [f'(x)]^2}{2gf(x)}} dx$$

where  $g$  is the acceleration due to gravity.

The actual calculation is quite complicated but after taking the derivative,

$$x(\theta) = R(\theta - \sin \theta), \quad y(\theta) = R(1 - \cos \theta)$$

for  $0 \leq \theta \leq \theta^*$  where  $(R, \theta^*)$  are parameters determined by  $(a, b)$

## Discrete Deterministic Dynamic Programming

**Example (Knapsack Problem):** Consider a backpack with capacity  $C$  and some amount of items  $i$ . Each item has a weight  $w_i$  and a value  $v_i$ . What is the maximum value of items we can carry in the backpack?

Formally,

$$V^* = \max \left\{ \sum_{i \in A} v_i \mid A \subseteq \{1, 2, \dots, n\}, \sum_{i \in A} w_i \leq C \right\}$$

The obvious solution is brute force. But there are  $2^n$  possible subsets of  $A$  (each item is either in or out of the backpack). Let's try something else.

### Dynamic Programming:

1. Define a *value function*  $V_k^*(c)$  as the optimal value for capacity  $c$  and items  $\{1, 2, \dots, k\}$
2. Define the *optimal placement*  $A_k^*(c)$  as the optimal subset of  $\{1, 2, \dots, k\}$

3. Solve inductively. For  $(V_{k+1}^*(c), A_{k+1}^*(c))$ , consider the  $k + 1$ -th item.

(a)  $V_0^*(c) = 0$  and  $A_0^*(c) = \emptyset$

(b) If  $w_{k+1} > c$ , then  $A_{k+1}^*(c) = A_k^*(c)$  and  $V_{k+1}^*(c) = V_k^*(c)$

(c) If  $w_{k+1} \leq c$ , then

$$V_{k+1}^*(c) = \max\{V_k^*(c), V_k^*(c - w_{k+1}) + v_{k+1}\}$$

$$A_{k+1}^*(c) = \begin{cases} A_k^*(c) & \text{if } V_{k+1}^*(c) = V_k^*(c) \\ A_k^*(c - w_{k+1}) \cup \{k + 1\} & \text{otherwise} \end{cases}$$

## 1.7 Sept 20

**Example:** Consider  $\{a_1, \dots, a_n\}$  where  $a_i$  is the price of a stock. We have some conditions: you can buy and sell one share at a time, but must sell before you can buy again. Short selling is not allowed. Design a strategy to maximize profit.

For  $n = 2$ , the problem is easy: if  $a_1 \geq a_2$ , do nothing. Otherwise, buy at  $a_1$  and sell at  $a_2$  for profit  $a_2 - a_1$ .

Let's generalize. Let  $V_k^*$  be the maximum profit for all transactions between 1 and  $k$ . Let  $T_k^* = \{(B_k, S_k)\}$  be the corresponding optimal transactions where  $B_i$  is the time to buy the stock and  $S_i$  is the time to sell the stock so

$$1 \leq B_1 < S_1 < B_2 < S_2 < \dots \leq k$$

Clearly,

$$V_k^* = \sum_i (a_{S_i} - a_{B_i})$$

*Initial step:*  $k = 0$ ,  $V_0^* = 0$  and  $T_0^* = \emptyset$ .

*Inductive step:* Suppose we have  $\{V_0^*, \dots, V_k^*\}$  and  $\{T_0^*, \dots, T_k^*\}$ .

CASE 1: Do nothing at day  $k + 1$  so  $V_{k+1}^* = V_k^*$  and  $T_{k+1}^* = T_k^*$

CASE 2: Sell at day  $k + 1$  (last day so we cannot buy) so

$$\begin{cases} V_{k+1}^* = V_k^* + (a_{k+1} - a_j) \\ T_{k+1}^* = T_k^* \cup \{(j, k + 1)\} \end{cases}$$



where  $j \in \{1, 2, \dots, k\}$  is the corresponding buy day.

Thus, iterate

$$V_{k+1}^* = \max_j \{V_{j-1}^* + (a_{k+1} - a_j)\}$$

$$T_{k+1}^* = \begin{cases} T_k^* & \text{if } j^* = k+1 \\ T_{j^*-1}^* \cup \{(j^*, k+1)\} & j^* \leq k \end{cases}$$

## 1.8 Sept 23

### Discrete-time Deterministic Quadratic Control Problem

**Finite Horizon:**

Let

$$x_{n+1} = x_n(1 - u_n)$$

for  $n = 0, 1, \dots, N-1$

Our goal is to find  $\{u_0, u_1, \dots, u_{N-1}\}$  to minimize

$$\sum_{n=0}^N (x_n^2 + \lambda u_n^2 x_n^2) + a x_N^2 \quad \lambda, a > 0$$

with  $x_0$  given.

We want to write the Dynamic Programming Equation. Define  $V_k(x)$  as the minimal cost from time  $t = k$  forward:

$$V_k(x) = \min_{u_k, u_{k+1}, \dots, u_{N-1}} \sum_{n=k}^{N-1} (x_n^2 + \lambda u_n^2 x_n^2) + a x_N^2$$

given  $x_k = x$ .

We start at  $x_k = x$  and use any control  $u_k = u$ . Then

$$x_{k+1} = x(1 - u)$$

If we assume to act optimally for all  $n > k$ , then the cost  $V_{k+1}(x(1 - u))$  would be the cost at time  $t = k$  plus the cost paid for later times:

$$(x^2 + \lambda u^2 x^2) + V_{k+1}(x(1 - u))$$

And optimizing this for  $u$ ,

$$V_k(x) = \min_{0 \leq u \leq 1} (x^2 + \lambda u^2 x^2) + V_{k+1}(x(1-u))$$

and this is the DPE.

Because we have a finite horizon, we can solve this backwards:

$$V_N \rightarrow V_{N_1} \rightarrow V_{N-2} \rightarrow \cdots \rightarrow V_1 \rightarrow V_0$$

where  $V_N(x) = ax^2$ , the terminal cost, and

$$V_k(x) = \min_{0 \leq u \leq 1} (x^2 + \lambda u^2 x^2) + V_{k+1}(x(1-u))$$

### **Infinite Horizon:**

Just as before, let

$$x_{n+1} = x_n(1 - u_n)$$

but now  $n = 0, 1, 2, \dots$

Our goal is to find  $\{u_0, u_1, \dots\}$  to minimize

$$\sum_{n=0}^{\infty} (x_n^2 + \lambda u_n^2 x_n^2) \quad \lambda > 0$$

with  $x_0$  given.

First, we write the Dynamic Programming Equation:

$$V(x) = \min_{u_n, n \in \mathbb{N}} \sum_{n=0}^{\infty} x_n^2 + \lambda u_n^2 x_n^2$$

Notice that at time  $t = 1$ ,  $x_0 = x$  so  $x_1 = x_0(1 - u_0) = x(1 - u)$  which is exactly the same situation at every other timestep.

Suppose we act optimally for every point after this point. Then

$$V(x_1) = (x_0^2 + \lambda u_0^2 x_0^2) + V(x(1-u))$$

But we want to minimize this so say we want to optimize the control at  $t = 0$ :

$$\min_{0 \leq u \leq 1} [(x_0^2 + \lambda u_0^2 x_0^2) + V(x(1-u))]$$

but then this will be the optimal value for all  $u$ , so our DPE is

$$V(x) = \min_{0 \leq u \leq 1} [(x_0^2 + \lambda u_0^2 x_0^2) + V(x(1-u))]$$

How can we solve this?

One method is to argue that a solution exists based on system constraints. Alternatively, we can just approximate it numerically.

Let's try something different.

In the Finite Horizon case, each  $V_k$  is quadratic. Assume that this is also the case for the Infinite Horizon case:  $V(x) = cx^2$ .

Then, by definition,

$$\begin{aligned} cx^2 &= \min_{0 \leq 1} [(x^2 + \lambda u^2 x^2) + c(x(1-u))^2] \\ &= \min_{0 \leq 1} [1 + \lambda u^2 + c(1-u)^2] x^2 \\ c &= \min_{0 \leq u \leq 1} [1 + \lambda u^2 + c(1-u)^2] \end{aligned}$$

and this is just a quadratic function so we take

$$\begin{aligned} \frac{d}{du} [1 + \lambda u^2 + c(1-u)^2]_{u=0} &= 2\lambda u + 2c(1-u)(-1) = 0 \\ \implies u^* &= \frac{c}{\lambda + c} \end{aligned}$$

where  $u^*$  is the optimal control.

Then explicitly,

$$c = 1 + \lambda(u^*)^2 + c(1-u^*)^2 = 1 + \frac{\lambda c}{c + \lambda}$$

so

$$c = \frac{1 \pm \sqrt{1 + 4\lambda}}{2}$$

but  $c > 0$  by assumption so

$$c = \frac{1 + \sqrt{1 + 4\lambda}}{2}$$

and

$$V(x) = \frac{1 + \sqrt{1 + 4\lambda}}{2} x^2$$

will solve the dynamic programming equation.

And even better, we identified the optimal control in the process.

## 1.9 Sept 25

### Optimal Quadratic Control

**Recall:** let the system state evolve according to

$$x_{n+1} = x_n(1 - u_n)$$

*Goal:* find  $\mathcal{U} = \{u_1, u_2, \dots\}$  to minimize

$$\min_{u_i \in \mathcal{U}} \sum_{n=0}^{\infty} (x_n^2 + \lambda x_n^2 u_n^2)$$

We can write the dynamic programming equation:

$$V(x) = \min_{0 \leq u \leq 1} \left[ \underbrace{x^2(1 + \lambda u^2)}_{\text{current timestep cost}} + \underbrace{V(x(1 - u))}_{\text{next timestep value}} \right]$$

Which will give the solution

$$V(x) = cx^2 = \frac{1 + \sqrt{1 + 4\lambda}}{2} x^2$$

corresponding to  $u^* = \frac{c}{c+\lambda}$  for all  $x, n$

In most applications, we can stop here. But we haven't yet answered "is this provably optimal?"

**Theorem:**  $u^* = \frac{c}{c+\lambda}$  is the optimal control and  $V(x) = cx^2$  is the minimal cost if  $x_0 = x$

*Proof (Verification argument):* We want to show

1. Given any  $\mathcal{U} = \{u_0, u_1, \dots\}$ ,

$$\sum_{n=0}^{\infty} (x_n^2 + \lambda x_n^2 u_n^2) \geq cx^2$$

2. If  $u_n^* = u^*$  for all  $n$ , then

$$\sum_{n=0}^{\infty} (x_n^2 + \lambda x_n^2 u_n^{*2}) = cx^2$$

Part 1: Define

$$a_k = \sum_{n=1}^{k-1} (x_n^2 + \lambda x_n^2 u_n^2) + cx_k^2$$

(heuristically: use  $u_1, \dots, u_k$  then use  $u^*$  for the rest).

Intuitively,  $a_k$  should be increasing as we use more and more non-optimal controls. And indeed,

$$\begin{aligned} a_{k+1} - a_k &= \left[ \sum_{n=0}^k (x_n^2 + \lambda x_n^2 u_n^2) + V(x_k) \right] - \left[ \sum_{n=0}^{k-1} (x_n^2 + \lambda x_n^2 u_n^2) + V(x_{k-1}) \right] \\ &= (x_k^2 + \lambda x_k^2 u_k^2) + V(x_{k+1}) - V(x_k) \end{aligned}$$

but

$$\begin{aligned} V(x_k) &= \min_{0 \leq u \leq 1} [x_k^2(1 + \lambda u^2) + V(x_k(1 - u))] \\ &\leq x_k^2(1 + \lambda u_k^2) + V(x_k(1 - u_k)) \quad (\text{if } u = u_k) \\ &= x_k^2 + \lambda x_k^2 u_k^2 + V(x_{k+1}) \end{aligned}$$

so

$$\begin{aligned} a_{k+1} - a_k &= (x_k^2 + \lambda x_k^2 u_k^2) + V(x_{k+1}) - V(x_k) \\ &\geq (x_k^2 + \lambda x_k^2 u_k^2) + V(x_{k+1}) - (x_k^2 + \lambda x_k^2 u_k^2 + V(x_{k+1})) = 0 \end{aligned}$$

So

$$a_{k+1} \geq a_k \geq \dots \geq a_1 \geq a_0 = (x^2 + \lambda x^2 u_0^2) + V(x_1) \geq V(x)$$

Hence

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k (x_n^2 + \lambda x_n^2 u_n^2) + cx_{k+1}^2 \geq cx^2$$

so

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k (x_n^2 + \lambda x_n^2 u_n^2) \not\rightarrow 0 \implies \sum_{n=0}^{\infty} (x_n^2 + \lambda x_n^2 u_n^2) \geq cx^2$$

as desired.

Part 2: Omitted

## Deterministic Continuous Time Control

Let  $[0, T]$  be a time interval. Define a deterministic control process:  $u : [0, T] \rightarrow \mathbb{R}$  and a state process  $x : [0, T] \rightarrow \mathbb{R}$  such that  $x_t = x(t)$ .

We define the *linear regulator* state model:

$$\frac{dx_t}{dt} = ax_t + u_t$$

The goal is to find a control process  $\{u_t : t \in [0, T]\}$  to minimize

$$x_T^2 + \int_0^T (\nu x_t^2 + \lambda u_t^2) dt$$

with  $\nu, \lambda > 0$ .

Let  $V(t, x)$  be the minimal cost on  $[t, T]$  where  $x_t = x$ :

$$V(t, x) = \min \left[ x_T^2 + \int_t^T (\nu x_s^2 + \lambda u_s^2) ds \right]$$

So the minimal cost for the original problem given  $x_0 = x$  is just  $V(0, x)$ .

As above, we want to consider an arbitrary control  $u_t = u$  for a small time step  $[t, \varepsilon]$  and then optimally control the rest. So

$$V(0, x) \leq \int_0^\varepsilon (\nu x_t^2 + \lambda u^2) + V(\varepsilon, x_\varepsilon)$$

## 1.10 Sept 27

### Deterministic Continuous-Time Control

Let  $x_t$  be the state at time  $t$  and  $u_t \in \mathbb{R}$  be the control at time  $t$ .

Suppose the state evolves according to

$$\frac{dx_t}{dt} = ax_t + u_t$$

**Goal:** Find  $\{u_t : t \in (0, T)\}$  (finite horizon!) to minimize

$$\underbrace{\int_0^T (ux_t^2 + \lambda u_t^2) dt}_{\text{running cost}} + \underbrace{x_T^2}_{\text{terminal cost}}$$

Define  $V : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  as the minimal cost on  $[t, T]$  given  $x_t = x$ :

$$V(t, x) = \min_u \int_t^T (\nu x_s^2 + \lambda u_s^2) ds + x_T^2$$

One strategy to solve this is to consider an arbitrary control  $u_t = u$  for a small time step  $[t, \varepsilon]$  and then optimally control the rest:

$$V(0, x_0) \leq \int_0^\varepsilon (\nu x_t^2 + \lambda u^2) dt + V(\varepsilon, x_\varepsilon)$$

But in fact, for optimal  $u$ ,

$$V(0, x_0) = \min_u \left[ \int_0^\varepsilon (\nu x_t^2 + \lambda u^2) dt + V(\varepsilon, x_\varepsilon) \right]$$

If we were to start at any time  $t \geq 0$ , we can use a similar argument to show that

$$V(t, x_t) = \min_u \left[ \int_t^{t+\varepsilon} (\nu x_s^2 + \lambda u^2) ds + V(t + \varepsilon, x_{t+\varepsilon}) \right]$$

But for small  $\varepsilon$ ,  $x_t \approx x_0$  and

$$V(t + \varepsilon, x_{t+\varepsilon}) - V(t, x_t) \approx \frac{\partial V}{\partial t} \varepsilon + \frac{\partial V}{\partial dx} (x_{t+\varepsilon} - x_t)$$

(from Taylor Series expansion)

And using the state ODE,

$$\frac{\partial V}{\partial t}\varepsilon + \frac{\partial V}{\partial dx}(x_{t+\varepsilon} - x_t) \approx \frac{\partial V}{\partial t}\varepsilon + \frac{\partial V}{\partial dx}(ax_t + u)\varepsilon$$

Together,

$$\begin{aligned} V(t, x_t) &= \min_u \left[ \int_t^{t+\varepsilon} (\nu x_s^2 + \lambda u^2) ds + V(t + \varepsilon, x_{t+\varepsilon}) \right] \\ &\approx (\nu x^2 + \lambda u^2) \cdot \varepsilon + V(t, x) + \frac{\partial V}{\partial t}\varepsilon + \frac{\partial V}{\partial dx}(ax_t + u)\varepsilon \\ &= V(t, x) + \varepsilon \left[ \nu x^2 + \lambda u^2 + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}(ax + u) \right] \end{aligned}$$

Since  $V(t, x_t)$  is minimal, the LHS must be less or equal to the RHS. So the bracketed term is  $\geq 0$ . But in the optimal case, the bracketed term is exactly 0, which gives us our control:

$$0 = \min_u \left[ (\nu x^2 + \lambda u^2) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}(ax + u) \right]$$

In fact, this is the **Hamilton-Jacobi-Bellman Equation**. In general, there will be infinitely many solutions but we can do better.

Consider the terminal cost:  $V(T, x) = 0 + x_T^2 = x^2$ . So we now have a boundary condition!

$$\begin{cases} 0 = \min_u \left[ (\nu x^2 + \lambda u^2) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}(ax + u) \right] \\ V(T, x) = x^2 \end{cases}$$

But if we rearrange our terms,

$$0 = \min_u \left[ \lambda u^2 + \frac{\partial V}{\partial x} \cdot u + \frac{\partial V}{\partial t} + ax \frac{\partial V}{\partial x} + \nu x^2 \right]$$

And to minimize  $u$ , it suffices to take the derivative with respect to  $u$  and set to 0:

$$0 = -\frac{1}{4\lambda} \left( \frac{\partial V}{\partial x} \right)^2 + \frac{\partial V}{\partial t} + ax \frac{\partial V}{\partial x} + \nu x^2$$

This is a PDE which it just so happen we can solve.



*Conjecture:*  $V(t, x) = f(t)x^2$ .

Then

$$\begin{aligned} 0 &= -\frac{1}{4\lambda} \left( \frac{\partial V}{\partial x} \right)^2 + \frac{\partial V}{\partial t} + ax \frac{\partial V}{\partial x} + \nu x^2 \\ &= x^2 \left[ f'(t) + 2af(t) - \frac{1}{\lambda} f^2(t) + \nu \right] \end{aligned}$$

which implies

$$\begin{cases} f'(t) + 2af(t) - \frac{1}{\lambda} f^2(t) + \nu = 0 \\ f(T) = 1 \end{cases}$$

which is precisely the **Piccati Equation** from engineering and which happens to have an explicit solution.

## 1.11 Sept 30

### Discrete-time Stochastic Optimal Control

**Example (Secretary Problem):** Suppose you have  $N$  secretaries to interview. You can interview them one at a time and must decide to hire or reject them immediately. You can't go back. What is the optimal strategy to maximize the probability of hiring the best secretary?

Define  $\{x_n\}_{n=1}^N$  by

$$x_n = \begin{cases} 1 & \text{if the } n\text{-th secretary is the best (so far)} \\ 0 & \text{otherwise} \end{cases}$$

**Suggested Strategy:**

1. If  $x_n = 0$ , reject the  $n$ -th secretary (certainly they cannot be the best if they are not the best so far)
2. If  $x_n = 1$ , if  $n$  is large, hire the  $n$ -th secretary. If  $n$  is small, reject the  $n$ -th secretary.

But how do we define “large” and “small”?

Define  $V_n : \{0, 1\} \rightarrow [0, 1]$  as the maximal probability of hiring the absolute best secretary given that  $x_n = x$

But if  $x_n = 0$ , then we want to reject the  $n$ -th secretary and move on so

$$V_n(0) = \mathbb{E}[V_{n+1}(x_{n+1}) \mid x_n = 0]$$

If  $x_1 = 1$ , we have two options: hire the  $n$ -th secretary or continue. The expected value if we continue is exactly the same as if  $x_n = 0$ . If we hire them, the probability of hiring the best secretary given that the  $n$ -th is the best so far is  $\frac{n}{N}$ . Together,

$$V_n(1) = \max\left\{\frac{n}{N}, \mathbb{E}[V_{n+1}(x_{n+1}) \mid x_n = 1]\right\}$$

So

$$V_n(x) = \begin{cases} \mathbb{E}[V_{n+1}(x_{n+1}) \mid x_n = 0] & \text{if } x = 0 \\ \max\left\{\frac{n}{N}, \mathbb{E}[V_{n+1}(x_{n+1}) \mid x_n = 1]\right\} & \text{if } x = 1 \end{cases}$$

We have terminal condition:

$$\begin{cases} V_N(1) = 1 \\ V_N(0) = 0 \end{cases}$$

So it just remains to find the conditional distribution of  $x_{n+1}$  given  $x_n$ .

**Claim:**  $\{x_i\}_1^N$  are independent.

*Proof:* Rigorous proof by combinatorics. Heuristic proof:  $x_n$  is the relative ranking of the first  $n$  secretaries.  $x_{n+1}$  does not depend on the ranking of the first  $n$  secretaries at all – just the probability that  $x_{n+1}$  is better than all of them.

So, in fact,

$$V_n(x) = \begin{cases} \mathbb{E}[V_{n+1}(x_{n+1})] & \text{if } x = 0 \\ \max\left\{\frac{n}{N}, \mathbb{E}[V_{n+1}(x_{n+1})]\right\} & \text{if } x = 1 \end{cases}$$

We know

$$\begin{cases} \mathbb{P}(x_{n+1} = 1) = \frac{1}{n+1} \\ \mathbb{P}(x_{n+1} = 0) = \frac{n}{n+1} \end{cases}$$

So

$$\begin{aligned}\mathbb{E}[V_{n+1}(x_{n+1})] &= V_{n+1}(1) \cdot \mathbb{P}(x_{n+1} = 1) + V_{n+1}(0) \cdot \mathbb{P}(x_{n+1} = 0) \\ &= \frac{1}{n+1} V_{n+1}(1) + \frac{n}{n+1} V_{n+1}(0)\end{aligned}$$

And at last we have the full Dynamic Programming Equation:

$$V_n(x) = \begin{cases} \frac{1}{n+1} V_{n+1}(1) + \frac{n}{n+1} V_{n+1}(0) & \text{if } x = 0 \\ \max\{\frac{n}{N}, V_n(0)\} & \text{if } x = 1 \end{cases}$$

with terminal condition  $V_N(1) = 1$  and  $V_N(0) = 0$ .

In order to solve this explicitly, we make some observations:

1.  $V_n(1) \geq V_n(0)$  trivially for all  $n$
2.  $V_n(0)$  is decreasing in  $n$
3. Suppose  $V_n(1) = V_n(0)$  for some  $n$ . Then

$$V_n(1) = V_n(0) \geq \frac{n}{N} \implies V_k(1) = V_k(0) = V_n(1) = V_n(0)$$

for all  $1 \leq k \leq n$ .

Therefore, there exists a  $1 \leq n^* \leq N$  such that

- (a)  $V_n(1) = V_n(0) = a$  for all  $1 \leq n \leq n^*$  and some constant  $a$
- (b)  $V_n(1) > V_n(0)$  for all  $n \geq n^*$

Then we have  $V_n(1) = \frac{n}{N} > V_N(0)$  for  $n \geq n^*$  so

$$\begin{aligned}V_n(0) &= \frac{1}{n+1} \frac{n+1}{N} + \frac{n}{n+1} V_{n+1}(0) \\ &= \frac{1}{N} + \frac{n}{n+1} V_{n+1}(0)\end{aligned}$$

for  $n \geq n^* - 1$ .

Since  $V_N(0) = 1$ , we claim

$$V_n(0) = \frac{n}{N} \sum_{k=n}^{n-1} \frac{1}{k}$$

for  $n^* - 1 \leq n \leq N$ .

Further,

$$\begin{aligned}
a &= V_{n^*-1}(1) \\
&= \max \left\{ \frac{n^* - 1}{N}, V_{n^*-1}(0) \right\} \\
&= V_{n^*-1}(0) \\
&\geq \frac{n^* - 1}{N} \\
V_{n^*}(1) &= \frac{n^*}{N} > V_{n^*}(0)
\end{aligned}$$

So,

$$\sum_{k=n^*}^{N-1} \frac{1}{k} < 1 \leq \sum_{k=n^*-1}^{N-1} \frac{1}{k}$$

which uniquely determines  $n^*$ .

**Optimal strategy:** Stop at time

$$\tau^* = \inf\{n \geq n^* : X_n = 1\}$$

In this case, the maximal probability of finding the absolute best is

$$a = V_{n^*-1}(0) = \frac{n^* - 1}{N} \sum_{k=n^*-1}^{N-1} \frac{1}{k}$$

It just so happens that

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N \frac{1}{n} - \log N = \gamma \approx 0.577$$

( $\gamma$  is the *Euler-Mascheroni constant*). So

$$\sum_{k=n}^{N-1} \frac{1}{k} \approx (\log N + \gamma) - (\log n + \gamma) = \log \left( \frac{N}{n} \right)$$

So for large  $N$ ,

$$n^* \approx \frac{N}{e} \approx \frac{N}{3}, \quad a \approx \frac{1}{e} \approx \frac{1}{3}$$

## 1.12 Oct 02

### Utility Maximization

**Example:** Let  $S_0 = x_0 > 0$  be your initial wealth and  $S_n$  the wealth at the end of the  $n$ -th betting round subject to

$$S_{n+1} = S_n + x_{n+1}y_{n+1}$$

where

- $0 \leq y_{n+1} \leq S_n$  is the amount bet at the  $n + 1$  round
- $x_{n+1} \in \{\pm 1\}$  is the outcome of the  $n + 1$  round with  $\mathbb{P}(x_n = 1) = p$ ,  $\mathbb{P}(x_n = -1) = q$ ,  $p > q$
- We assume  $\{x_1, \dots, x_n, \dots\}$  are iid.

*Goal:* Find  $\{y_1, y_2, \dots, y_N\}$  that maximize  $\mathbb{E}[u(S_N)]$  where  $u$  is the concave *power utility*

$$u(x) = x^\alpha, \quad 0 < \alpha < 1$$

**Concave:** a function  $f$  is concave iff  $f'$  is non-increasing. Otherwise, it is *convex*.

Define

$$v_n(x) = \max \mathbb{E}[u(S_N) \mid S_n = x]$$

With terminal condition

$$v_N(x) = \max \mathbb{E}[u(S_N) \mid S_n = x] = u(x) = x^\alpha$$

Consider any time  $n < N$  corresponding to  $S_n = x$ . Arbitrarily pick  $y_{n+1} = y$  for  $0 \leq y \leq x$ . Then

- With probability  $p$ ,  $x_{n+1} = 1$  so  $S_{n+1} = x + y$
- With probability  $q$ ,  $x_{n+1} = -1$  so  $S_{n+1} = x - y$

Now suppose we choose optimal controls for  $\{y_{n+2}, \dots, y_N\}$ . Then if

$$\begin{cases} S_{n+1} = x + y \\ S_{n+1} = x - y \end{cases} \implies \begin{cases} v_{n+1}(x + y) \\ v_{n+1}(x - y) \end{cases}$$

So

$$v_n(x) = \max_{0 \leq y \leq x} [pv_{n+1}(x+y) + qv_{n+1}(x-y)]$$

subject to  $v_N(x) = x^\alpha$ .

This is a recursive equation with terminal condition but it is incredibly difficult to solve numerically because  $x \in \mathbb{R}^+$  and  $y \in [0, x]$ .

Let's take just the last step:

$$\begin{aligned} v_{N-1}(x) &= \max_{0 \leq y \leq x} [pv_N(x+y) + qv_N(x-y)] \\ &= \max_{0 \leq y \leq x} [p(x+y)^\alpha + q(x-y)^\alpha] \\ &\stackrel{y=tx}{=} \max_{0 \leq t \leq 1} [p(x+tx)^\alpha + q(x-tx)^\alpha] \\ &= \max_{0 \leq t \leq 1} x^\alpha [p(1+t)^\alpha + q(1-t)^\alpha] \end{aligned}$$

But the bracketed term is a constant (say  $c$ ) so

$$v_{N-1}(x) = cx^\alpha$$

and in fact  $v_{n-2}(x) = c^2x^\alpha, \dots$  so we are done.

With some ugly algebra,

$$t^* = \frac{p^\gamma - q^\gamma}{p^\gamma + q^\gamma}, \quad \gamma = \frac{1}{1-\alpha}$$

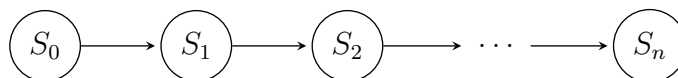
For all of these examples, we provide very little proof. That requires *Markov Chains and Martingales*

# Chapter 2

## Markov Chain & Martingales

### 2.1 Oct 02

**Markov chain:** a sequence of random variables  $\{S_n\}_{n=0}^{\infty}$  such that  $S_{n+1}$  depends only on  $S_n$



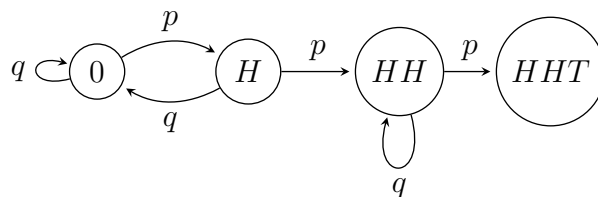
Formally,

$$[S_{n+1} \mid S_n, S_{n-1}, \dots, S_1] = [S_{n+1} \mid ]$$

**Example (Simple Random Walk):**  $S_{n+1} = S_n + X_{n+1}$  where  $\mathbb{P}(X_{n+1} = \pm 1) = \frac{1}{2}$

**Example:** Take a coin with  $\mathbb{P}(H) = p$ ,  $\mathbb{P}(T) = q$ . What is the expected number of tosses until the pattern HHT appears?

We can draw a Markov chain for a particular pattern:



## 2.2 Oct 04

### Background

**Recall:** A Markov chain is a sequence of random variables  $\{X_n\}_{n=0}^\infty$  such that

$$\mathbb{P}(X_{n+1} \in A \mid X_n = x_n, \dots, X_0 = x_0) = \mathbb{P}(X_{n+1} \in A \mid X_n = x_n)$$

**State space:** Let  $\mathcal{S}$  be all the possible values of  $X_n$ . In this class, assume  $\mathcal{S}$  is countable.

**Transition probability:** If  $\mathbb{P}(X_{n+1} = y \mid X_n = x)$  is time-invariant (does not depend on  $n$ ), then we call it **homogeneous** and write

$$P_{xy} = \mathbb{P}(X_{n+1} = y \mid X_n = x)$$

**Transition matrix:** We can enumerate all the transition probabilities by

$$[P_{xy}]_{xy} = \begin{bmatrix} P_{00} & P_{01} & \dots \\ P_{10} & P_{11} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

which will have two very nice properties:

1.  $P_{xy} \geq 0$  for all  $x, y \in \mathcal{S}$
2.  $\sum_{y \in \mathcal{S}} P_{xy} = 1$  for all  $x \in \mathcal{S}$  (each row sums to 1)

**Example (Simple Random Walk):** Consider

$$S_{n+1} = S_n + X_{n+1}$$

over  $\mathcal{S} = \mathbb{Z}$  where  $\mathbb{P}(X_{n+1} = 1) = p$  and  $\mathbb{P}(X_{n+1} = -1) = q$ .

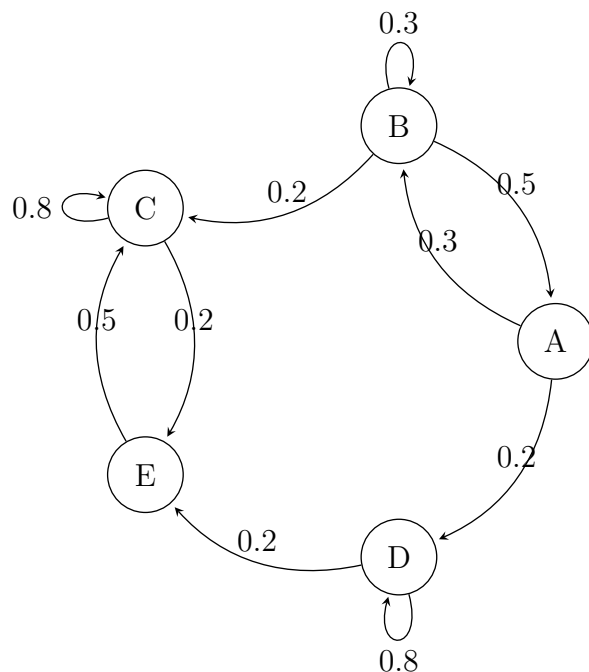
Clearly,

$$\begin{cases} P_{x,x+1} = p \\ P_{x,x-1} = q \\ P_{x,y} = 0 \end{cases} \quad \text{if } y \neq x \pm 1$$



## Computational Problems

Example:



So, for example,

- $P_{AD} = 0.2$
- $P_{EA} = 0.5$
- $P_{ED} = 0$

*Question 1:* What is the probability that starting from state  $B$ , the Markov Chain will hit state  $D$  eventually

Let's try first step analysis. Call  $x_B$  the probability of hitting  $D$  starting from  $B$ .

Then the first transition is just the probability of transitioning times the probability of hitting  $D$  from the new state:

$$\begin{aligned}x_B &= P_{BA}x_A + P_{BB}x_B + P_{BC}x_C + P_{BD}x_D + P_{BE}x_E \\&= 0.5x_A + 0.3x_B + 0.2x_C\end{aligned}$$

and if we repeat this for all states, we get a system of 5 linear equations with 5 unknowns – which we can solve with a matrix!

And to make things simpler,  $x_D = 1$ .

And after finding all the equations and solving, we get  $x_B = 1$

*Question 2:* What is the probability that starting from state  $B$ , the Markov Chain will hit state  $D$  before state  $E$ ?

Again call this probability  $x_B$  and notice that  $x_s$  for  $s \in \mathcal{S}$  is exactly the same as before (except with two special condition):

- $x_D = 1$
- $x_E = 0$

*Question 3:* What is the average number of steps to reach state  $D$  from state  $A$ ?

We can once again call this  $x_A$  and define

$$\begin{cases} x_s = 0 & \text{if } s = D \\ x_s = P_{sA}(1 + x_A) + P_{sB}(1 + x_B) + P_{sC}(1 + x_C) + P_{sD}(1 + x_D) + P_{sE}(1 + x_E) & \text{otherwise} \end{cases}$$

for  $s \in \mathcal{S}$ .

*Question 4:* Suppose the initial state is  $A$ . Let  $T$  be the number of steps to reach  $D$ . In Q3, we calculated  $\mathbb{E}[T]$ . Now calculate

$$\mathbb{E} \left[ \sum_{n=0}^{T-1} \beta^n \phi(x_n) \right]$$

where  $\phi(n)$  represent a payoff table

| $s$ | $\phi(x_s)$ |
|-----|-------------|
| A   | 1           |
| B   | 2           |
| C   | 0           |
| D   | -1          |
| E   | 4           |

and  $\beta$  is a discounting factor  $0 < \beta < 1$ .

As always, let

$$x_A = \mathbb{E} \left[ \sum_{n=0}^{T-1} \beta^n \phi(x_n) \right]$$

How do we generalize to all states?

Once again we have special case  $x_D = 0$ . Now consider  $x_A$ . On the first step, we reach  $A$  with probability  $P_{AA}$  and the payoff will be  $(\phi(A) + \beta x_A)$  so

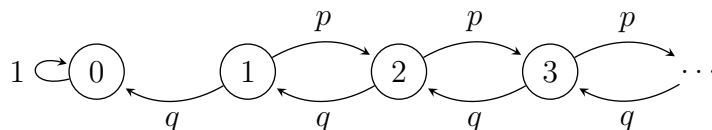
$$\begin{aligned} x_A &= P_{AA}(\phi(A) + \beta x_A) \\ &\quad + P_{AB}(\phi(A) + \beta x_B) \\ &\quad \vdots \\ &\quad + P_{AE}(\phi(A) + \beta x_E) \end{aligned}$$

and analogously for all other states.

## 2.3 Oct 07

**Example:** Let's see a case where first step analysis fails:

Let  $S_{n+1} = S_n + X_{n+1}$  with  $\mathbb{P}(X_{n+1} = 1) = p$  and  $\mathbb{P}(X_{n+1} = -1) = q$  and  $p > q$ .



What is the probability  $x_k$  that the random walk reaches 0 eventually given that  $S_0 = k$  and assuming  $p > q$ .

From first-step analysis,

$$x_k = px_{k+1} + qx_{k-1}$$

for all  $k \geq 1$  with  $x_0 = 1$ .

The characteristic equation is  $r^2 = pr^2 + q \implies r = \{1, \frac{q}{p}\}$

So the general solution is

$$x_k = Ar_1^k + Br_2^k = A + B \left(\frac{q}{p}\right)^k$$

where  $A, B$  are constants.

Plugging in the boundary condition,  $A + B = 1$  so

$$x_k = 1 + B \left[ \left(\frac{q}{p}\right)^k - 1 \right]$$

But clearly,  $0 \leq x_k \leq 1$  so  $0 \leq B \leq 1$ . But now  $B$  cannot be uniquely determined – *every* such  $B$  satisfies the boundary condition. What do we do now?

*Option 1:* Artificially fix a large number  $N$  and consider the  $\{x_k\}_0^N$ . The solution to the difference equation using first-step analysis will depend on  $N$  so just let  $N \rightarrow \infty$  and obtain  $x_k$ .

*Option 2:* Observe that (heuristically)  $x_k \rightarrow 0$  as  $k \rightarrow \infty$ . (Rigorous follows from  $p > q$  and some work).

In either approach, we find  $B = 1$  and

$$x_k = \left(\frac{q}{p}\right)^k$$

Both options work but both require some model-specific information which is not ideal.

Let's try to find a way to fix the first step analysis to determine the unique solution.

Consider a Markov chain with state space  $\mathcal{S}$  and transition matrix  $P$ . Let  $A \subseteq \mathcal{S}$ .

We define the *hitting time* of  $A$  as

$$T^A = \inf\{n \geq 0 : X_n \in A\}$$

with the convention that  $\inf \emptyset = \infty$ .

We want

$$h^A(x) = \mathbb{P}^x(T^A < \infty), d^A(x) = \mathbb{E}^x[T^A]$$

**Theorem:** The vector  $\{h^A(x) : x \in \mathcal{S}\}$  is the minimal nonnegative solution of

$$h(x) = \begin{cases} \sum_{y \in \mathcal{S}} P_{xy} h(y) & x \notin A \\ 1 & x \in A \end{cases}$$

Similarly,  $\{d^A(x) : x \in \mathcal{S}\}$  is the minimal nonnegative solution of

$$d(x) = \begin{cases} 1 + \sum_{y \in \mathcal{S}} P_{xy} h(y) & x \notin A \\ 0 & x \in A \end{cases}$$

*Proof:* By the LTP,  $\{h^A(x) : x \in \mathcal{S}\}$  is certainly a solution. Now suppose  $\{h(x) : x \in \mathcal{S}\}$  is an arbitrary nonnegative solution.

We want to show that  $h(x) \geq h^A(x)$  for all  $x \in \mathcal{S}$ .

If  $x \in A$ ,  $h(x) = h^A(x) = 1$ . So choose  $x \notin A$ .

Repeatedly applying the system of equations, we get:

$$\begin{aligned} h(x) &= \sum_{y \in \mathcal{S}} P_{xy} h(y) \\ &= \sum_{y \in A} P_{xy} h(y) + \sum_{y \notin A} P_{xy} h(y) \\ &= \sum_{y \in A} P_{xy} + \sum_{y \notin A} P_{xy} \left[ \sum_{z \in A} P_{yz} + \sum_{z \notin A} P_{yz} h(z) \right] \\ &= \sum_{y \in A} P_{xy} + \sum_{y \notin A} \sum_{z \in A} P_{xy} P_{yz} + \sum_{y \notin A} \sum_{z \notin A} P_{xy} P_{yz} h(z) \\ &= P^x(T^A = 1) + P^x(T^A = 2) + \sum_{y \notin A} \sum_{z \notin A} P_{xy} P_{yz} h(z) \\ &\vdots \end{aligned}$$

So

$$h(x) \geq \mathbb{P}^x(T^A \leq n)$$

for all  $n$ . Letting  $n \rightarrow \infty$ , we get

$$h(x) \geq \mathbb{P}^x(T^A < \infty) = h^A(x)$$

For  $\{d^A(x)\}$ , by the tower property of conditional expectation, it is a solution to the corresponding system of equations. We take  $\{d(x) : x \in \mathcal{S}\}$  to be an arbitrary nonnegative solution and want to show that  $d(x) \geq d^A(x)$ .

Again  $x \in A$  is trivial and for  $x \notin A$ ,

$$\begin{aligned} d(x) &= 1 + \sum_{y \notin A} P_{xy} d(y) \\ &= 1 + P_{xy} \left[ 1 + \sum_{z \notin A} P_{yz} h(z) \right] \\ &= P^x(T^A \geq 1) + P^x(T^A \geq 2) + \sum_{y \notin A} \sum_{z \notin A} P_{xy} P_{yz} d(z) \end{aligned}$$

So, for all  $N$ ,

$$d(x) \geq \sum_{n=1}^N P^x(T^A \geq n)$$

And as  $n \rightarrow \infty$ , we know  $T^A$  takes values in  $\{0, 1, 2, \dots\} \cup \{\infty\}$  so

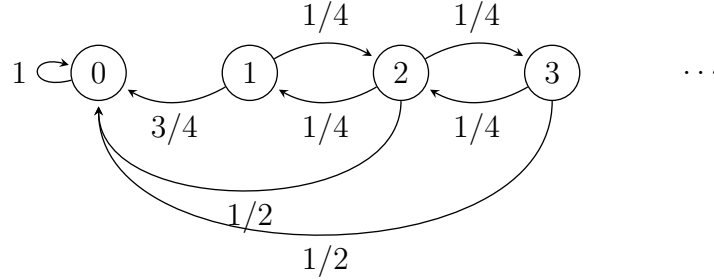
$$\sum_{n=1}^{\infty} \mathbb{1}_{\{T^A \geq n\}} = T^A$$

and taking the expectation of both sides, we are done.

**Example:** Consider the MC  $X$  with state space  $\mathcal{S} = \{0, 1, 2, \dots\}$ . Assume  $P_{00} = 1$  and for all  $x > 0$ ,

$$P_{xy} = \begin{cases} 1/2 & y = 0 \\ 1/4 & y = x + 1 \\ 1/4 & y = x - 1 \\ 0 & \text{otherwise} \end{cases}$$

modelled by



Let  $T = \inf\{n \geq 0 : X_n = 0\}$ . We are interested in the average duration  $d(x) = \mathbb{E}^x[T]$

By the Theorem,  $\{d(x) : x \in \mathcal{S}\}$  is the minimal nonnegative solution to

$$\begin{aligned} d(x) &= 1 + \frac{1}{2}d(0) + \frac{1}{4}d(x+1) + \frac{1}{4}d(x-1) \\ &= 1 + \frac{1}{4}d(x+1) + \frac{1}{4}d(x-1) \end{aligned}$$

for  $x \geq 1$  with  $d(0) = 0$ .

If we consider  $x$  an index, these equations are just a difference equation! (Excepting the “1” term which make it no longer homogeneous).

This allows us to follow the standard approach:

1. Find a particular sequence, ignoring the boundary condition.

Here,  $d_p(x) = 2$

2. Find the general solution.

Here,  $d_g(x) = d(x) - d_p(x)$  satisfies

$$d_g(x) = \frac{1}{4}d_g(x+1) + \frac{1}{4}d_g(x-1)$$

So  $r = \frac{1}{4}r^2 + \frac{1}{4} \implies r = 2 \pm \sqrt{3}$  and

$$d_g(x) = Ar_1^x + Br_2^x = A(2 + \sqrt{3})^x + B(2 - \sqrt{3})^x$$

for constants  $A, B$

3. Fit the boundary condition.

Here,  $d(x) = d_g(x) + d_p(x)$  with  $d(0) = 0$  so  $A + B = -2$  and

$$d(x) = A \left[ (2 + \sqrt{3})^x - (2 - \sqrt{3})^x \right] + 2 \left[ 1 - (2 - \sqrt{3})^x \right]$$

Since  $d(x)$  is nonnegative,  $A \geq 0$  and the minimality of  $d$  implies  $A = 0$  so

$$d(x) = 2[1 - (2 - \sqrt{3})^x]$$

**Example (Branching Process):** Consider a stochastic model for population growth. Let  $X_n$  denote the size of the population in the  $n$ -th generation.

Each individual independently reproduces a random number of offspring:

$$X_{n+1} = \sum_{i=1}^{X_n} Z_i^{(n)}$$

where  $\{Z_i^{(n)}\}$  are iid with distribution

$$\mathbb{P}(Z = k) = p_k, \quad k = 0, 1, 2, \dots, \quad \sum_{k=0}^{\infty} p_k = 1$$

Assume  $X_0 = 1$  and let  $T = \mathbb{1}\{n \geq 0 : X_n = 0\}$  be the time to extinction. We are interested in  $\mathbb{P}(T < \infty)$  (will the population ever go extinct?)

Define for each  $x \geq 0$ ,

$$H(x) = P^x(T_0 < \infty)$$

which is the probability of extinction given initial population  $x$ . We want  $\pi = H(1)$ .

Again by the theorem,  $\{H(x) : x \geq 0\}$  is the minimal nonnegative solution to

$$H(x) = \sum_{k=0}^{\infty} P_{xk} H(k)$$

for  $x \geq 1$  with  $H(0) = 1$ , where  $P_{xk} = \mathbb{P}(X_{n+1} = k \mid X_n = x)$  is the transition probability.

The key observation is that this system can be reduced to a single equation:

$$H(x) = [H(1)]^x = \pi^x$$



Therefore, we expect  $\pi$  to satisfy

$$\pi = \sum_{k=0}^{\infty} P_{1k} \pi^k = \sum_{k=0}^{\infty} P_k \pi^k = \phi(\pi) \quad (*)$$

And  $\forall \pi$  that satisfy the above,

$$H(x) = \pi^x = \left[ \sum_{i=0}^{\infty} P_i \pi^i \right]^x = \sum_{k=0}^{\infty} P_{xk} \pi^k = \sum_{k=0}^{\infty} P_{xk} H(k)$$

for  $x \geq 0$  where the coefficient of  $\pi^k$  is the sum of all possible ways to have  $k$  offspring in total. Therefore,  $\pi$  must be the guaranteed minimal solution.

To solve  $(*)$ , we need to study  $\phi$ .

$$\phi(0) = p_0$$

$$\phi(1) = 1$$

$$\phi'(0) = p_1$$

$$\phi'(1) = \sum_{k=0}^{\infty} k p_k = \mathbb{E}[Z]$$

so  $\phi$  is increasing and convex.

Excluding the trivial case  $p_0 = 0$  (the population will never go extinct), we assume  $p_0 > 0$ .

And with some calculus,

1. If  $p_0 > 0$  and  $\mathbb{E}[Z] \leq 1$ , then  $\pi = 1$
2. If  $p_0 > 0$  and  $\mathbb{E}[Z] > 1$ , then  $0 < \pi < 1$

To give a numerical example, assume  $p_k = p(1-p)^k$  for  $k \geq 0$  where  $p \in (0, 1)$  is a constant. This is equivalent to  $Z + 1$  be geometrically distributed with parameter  $p$  so

$$\begin{aligned} \mathbb{E}[Z] &= \frac{1}{p} - 1, \\ \phi(\pi) &= \sum_{k=0}^{\infty} p(1-p)^k \pi^k \\ &= \frac{p}{1 - (1-p)\pi} \end{aligned}$$

$\pi = \phi(\pi)$  has two roots,  $\{1, \frac{p}{1-p}\}$

Therefore, if  $p \geq 1/2$  (or equivalently  $\mathbb{E}[Z] \leq 1$ ), we have  $\pi_1 \leq \pi_2$  and  $\pi = 1$ . On the other hand, if  $p < 1/2$  (or  $\mathbb{E}[Z] > 1$ ),  $\pi_1 > \pi_2$  and  $\pi = \pi_2 < 1$

## 2.4 Oct 09

**Recall (Branching Process):** 1st step analysis gives us the *hitting probability* which guarantees a minimal nonnegative solution.

Take the case of a population where each member has a random number of offspring  $Z$  where  $\mathbb{P}(Z = 0) = p_0, \mathbb{P}(Z = 1) = p_1, \dots$

We can define  $\pi$  to be the probability of extinction. Clearly,

$$\begin{aligned}\pi &= \sum_{k=0}^{\infty} \mathbb{P}(Z = k) \pi^k \\ &= p_0 + p_1 \pi + p_2 \pi^2 + \dots \\ &= \phi(\pi)\end{aligned}$$

This is called a **moment generating function**.

We have  $\phi(0) = p_0$  and  $\phi(1) = 1$ . We can show that  $\phi$  is increasing and convex.

And then we can show that if  $\phi'(1) < 1$ , then  $\phi$  has a solution less than  $\pi = 1$  (a minimal solution). Otherwise,  $\phi(1)$  is the minimal solution.

Let's calculate.

$$\begin{aligned}\phi'(\pi) &= p_1 + 2p_2\pi + 3p_3\pi^2 + \dots \\ \phi'(1) &= p_1 + 2p_2 + 3p_3 + \dots \\ &= \mathbb{E}[Z]\end{aligned}$$

so  $\phi'(1) > 1 \iff \mathbb{E}[Z] > 1$ . Hence, the probability the population goes extinct depends on the expected number of offspring (not surprising!)

Now suppose  $Z + 1$  is geometric with parameter  $p$ . Then

$$\begin{aligned}\mathbb{P}(Z = 0) &= p \\ \mathbb{P}(Z = 1) &= pq \\ \mathbb{P}(Z = 2) &= pq^2 \\ &\vdots\end{aligned}$$

where  $q = (1 - p)$  and  $0 < p < 1$ .

So

$$\begin{aligned}\phi(\pi) &= p + pq\pi + pq^2\pi^2 + \dots \\ &= p(1 + q\pi + (q\pi)^2 + (q\pi)^3 + \dots) \\ &= \frac{p}{1 - q\pi}\end{aligned}$$

$\pi = \frac{p}{1 - q\pi}$  is a quadratic with two roots:  $\pi = \{1, \frac{p}{q}\}$ .

Therefore,

$$\pi_2 < 1 \iff o < q \iff p < \frac{1}{2} \iff \mathbb{E}[Z + 1] = \frac{1}{p} > Z \iff \mathbb{E}[Z] > 1$$

exactly as before.

## Martingales

**Martingale:** A stochastic process that on average stays constant

**Example (Random Walk):**

$$\begin{cases} S_0 = x \\ S_{n+1} = S_n + X_{n+1} \end{cases} \iff S_{n+1} = x + \sum_{k=1}^{n+1} X_{n+1}$$

where  $X_n$  are iid R.V. with (say)  $\mathbb{P}(X_n = \pm 1) = \frac{1}{2}$

But

$$\mathbb{E}[S_{n+1} \mid S_n, S_{n-1}, \dots, S_0] = \mathbb{E}[S_{n+1} \mid S_n] = S_n$$

so  $S_n$  is a martingale.

**Definition:** With

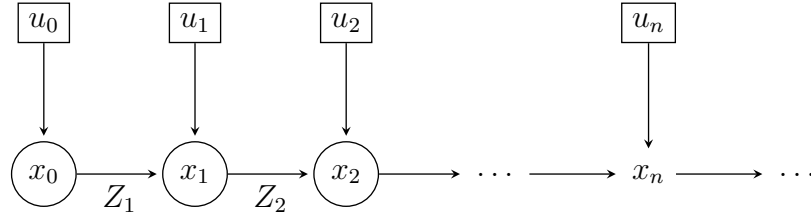
1.  $\mathcal{F}_n = (Z_0, Z_1, \dots, Z_n)$  is a *filtration process* (reference) containing the information up to time  $n$  for  $Z = \{Z_0, Z_1, \dots\}$ .

2.  $\{X_0, X_1, \dots\}$  is *adapted* to  $Z$ , i.e.  $X_n = X_n(Z_0, Z_1, \dots, Z_n)$

then  $\{X_n\}$  is a **martingale** (with respect to  $Z$ ) if

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$$

**Example:**



So

$$X_{n+1} = f(X_n, u_n, Z_{n+1})$$

and we call  $\{X_n\}$  a *Markov Controlled Process*

**Example:** Let  $S_n$  be the standard random walk ( $S_{n+1} = x_0 + \sum_{k=1}^{n+1} X_k$ ) such that  $X_n$  are iid with  $\mathbb{E}[X_n] = \mu$ .

Define  $Y_n = S_n - n\mu$ . We claim  $(Y_n)$  is a Martingale W.R.T  $(X_n)$ :

$$\begin{aligned} \mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[S_{n+1} - (n+1)\mu \mid \mathcal{F}_n] \\ &= \mathbb{E}[S_{n+1} \mid \mathcal{F}_n] - (n+1)\mu \\ &= \mathbb{E}[S_n + X_{n+1} \mid \mathcal{F}_n] - (n+1)\mu \\ &= \mathbb{E}[S_n \mid \mathcal{F}_n] + \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] - (n+1)\mu \\ &= \mathbb{E}[S_n \mid \mathcal{F}_n] + \mu - (n+1)\mu \\ &= S_n + \mu - (n+1)\mu \\ &= S_n - n\mu \\ &= Y_n \end{aligned}$$

## 2.5 Oct 11

**Example (Wald's Martingale):** Let  $X = \{X_0, X_1, \dots\}$  be a sequence of iid random variables and  $\theta \in \mathbb{R}$  a constant. Let  $\phi(\theta) = \mathbb{E}[e^{\theta X_1}]$  (the moment generating

function) and

$$S_n = x_0 + \sum_{i=1}^n X_i$$

for some  $x_0 \in \mathbb{R}$  and all  $n \geq 0$ . Then the process

$$M_n = \frac{1}{\phi^n(\theta)} e^{\theta S_n}$$

is a martingale with respect to the reference process  $\{X_0, X_1, \dots\}$ . It is not hard to verify this claim.

The adaptiveness is trivial and

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = \frac{1}{\phi^{n+1}(\theta)} \mathbb{E}[e^{\theta S_{n+1}} \mid \mathcal{F}_n] = \frac{1}{\phi^{n+1}(\theta)} e^{\theta S_n} \mathbb{E}[e^{\theta X_{n+1}} \mid \mathcal{F}_n]$$

Since  $X_{n+1}$  is independent of  $\mathcal{F}_n = (X_0, X_1, \dots, X_n)$  and has the same distribution as  $X_1$ , we have

$$\mathbb{E}[e^{\theta X_{n+1}} \mid \mathcal{F}_n] = \mathbb{E}[e^{\theta X_{n+1}}] = \mathbb{E}[e^{\theta X_1}] = \phi(\theta)$$

Therefore,  $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$  and  $\{M_n\}$  is a martingale.

**Example (Martingale Transform):** Let  $M = \{M_0, M_1, M_2, \dots\}$  be a martingale with respect to some reference process  $\{Z_0, Z_1, \dots\}$ . Let  $Y = \{Y_1, Y_2, \dots\}$  be a process such that  $Y_n$  is a function of  $\mathcal{F}_{n-1}$  for all  $n \geq 1$  (such processes are said to be *predictable*).

Define

$$X_n = \sum_{i=1}^n Y_i (M_i - M_{i-1}) = (Y \bullet M)_n$$

for all  $n \geq 0$ . Then  $\{X_n\}$  is a martingale with  $X_0 = 0$ .

The verification is not hard: The adaptiveness is trivial and

$$\begin{aligned} \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] &= X_n + \mathbb{E}[Y_{n+1}(M_{n+1} - M_n) \mid \mathcal{F}_n] \\ &= X_n + Y_{n+1} \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\ &= X_n \end{aligned}$$

Note that the predictability of  $Y$  is the only reason we can take  $Y_{n+1}$  outside the conditional expectation.

In general, this procedure generates many martingales from a given martingale. Moreover, this is indeed a preliminary version of stochastic integral if  $M_{n+1} - M_n$  is viewed as the differential  $dM$  of the martingale  $M$ .

## Optional Sampling

**Stopping time:** Let  $Z = \{Z_0, Z_1, Z_2, \dots\}$  be a reference process. A mapping  $\tau : \Omega \rightarrow \{0, 1, 2, \dots\} \cup \{\infty\}$  is said to be a *stopping time* if for every  $n$ , the indicator random variable  $\mathbb{1}_{\tau \leq n}$  is a function of  $\mathcal{F}_n = (Z_0, Z_1, \dots, Z_n)$ .

Heuristically, the decision to stop at time  $\tau$  is a stopping time iff the decision is based only on historical information.

**Example:** Let  $X = \{X_n\}$  be adapted WRT  $Z$  and  $a \in \mathbb{R}$  constant. Define

$$\tau = \inf\{n \geq 0 : X_n \geq a\}, \quad \sigma = \sup\{n \geq 0 : X_n \geq a\}$$

Both are random times but  $\tau$  is a stopping time while  $\sigma$  is not.

**Optional Sampling Theorem:** Let  $X = \{X_0, X_1, \dots\}$  be a martingale and  $\tau$  an arbitrary stopping time. Then

$$E[X_{\tau \wedge n}] = E[X_0]$$

for every  $n$ . If  $X$  is a supermartingale or submartingale, then the equality is replaced by  $\leq$  or  $\geq$  respectively.

*Proof:* We will only pose the martingale case as the others are similar.

Fix  $n \geq 0$ . Then

$$X_{\tau \wedge (n+1)} = X_\tau \mathbb{1}_{\tau \leq n} + X_{n+1} \mathbb{1}_{\tau > n}$$

The key observation is

$$\mathbb{1}_{\tau > n} = 1 - \mathbb{1}_{\tau \leq n}$$

is a function of  $\mathcal{F}_n$ . Therefore,

$$\mathbb{E}[X_{n+1} \mathbb{1}_{\tau > n} \mid \mathcal{F}_n] = \mathbb{1}_{\tau > n} \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{1}_{\tau > n} X_n$$

By the tower property, taking expectation on both sides we have

$$\mathbb{E}[X_{n+1} \mathbb{1}_{\tau > n}] = \mathbb{E}[X_n \mathbb{1}_{\tau > n}]$$

In particular,

$$\mathbb{E}[X_{\tau \wedge (n+1)}] = \mathbb{E}[X_\tau \mathbb{1}_{\tau \leq n}] + \mathbb{E}[X_{n+1} \mathbb{1}_{\tau > n}] = \mathbb{E}[X_{\tau \wedge n}]$$

Since this is true for all  $n$ , the claim follows readily as  $\tau \wedge 0 = 0$ .

**Remark:** Ideally, we would have  $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$  when  $X$  is a martingale in the Optional Sampling Theorem. However, one cannot hope this to be true in general. For example, let  $X$  be a symmetric simple random walk and  $\tau = \inf\{n \geq 0 : X_n = 1\}$ . Then  $\mathbb{E}[X_\tau] = 1$  while  $\mathbb{E}[X_0] = 0$ .

Therefore, we add conditions to ensure the  $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ :

1. The stopping time is bounded, i.e.  $\mathbb{P}(\tau \leq t) = 1$  for some  $t \in \mathbb{N}$
2. The process  $\{X_n\}$  is bounded for  $n = 1, \dots, \tau$

In this course, some examples will not satisfy these conditions but it can nonetheless be shown that  $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ . As we are not interested in rigor, *assume that the equality holds if  $X$  is a martingale and  $\tau$  is a stopping time.*

## 2.6 Oct 16

**Definition:**

- If  $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$ , we say the process is a martingale
- If  $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n$ , we say the process is a supermartingale
- If  $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n$ , we say the process is a submartingale

Further, using the Law of total expectation,  $\mathbb{E}[\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]] = \mathbb{E}[X_{n+1}]$  so

$$\begin{cases} \mathbb{E}[X_{n+1}] = \mathbb{E}[X_n] & \text{if martingale} \\ \mathbb{E}[X_{n+1}] \geq \mathbb{E}[X_n] & \text{if submartingale} \\ \mathbb{E}[X_{n+1}] \leq \mathbb{E}[X_n] & \text{if supermartingale} \end{cases}$$

**Recall (Optional Sampling Theorem):** If  $(X_0, X_1, \dots)$  is a martingale and  $\tau$  is a stopping time,

$$\mathbb{E}[X_{\tau \wedge n}] = \mathbb{E}[X_0] \quad \forall n$$

And under some conditions,

$$\mathbb{E}[X] = \lim_{n \rightarrow \infty} \mathbb{E}[X_{\tau \wedge n}] = \mathbb{E}[X_\tau]$$

Specifically, these equalities hold if

1. if  $\tau$  is bounded, i.e.  $\mathbb{P}(\tau \leq N) = 1$  for fixed  $N \in \mathbb{N}$ .
2. if  $(X_n)$  is bounded on  $n = 1, \dots, \tau$

**Example (Gambler's Ruin):** Let  $X = \{X_0, X_1, \dots\}$  be iid RVs with

$$\begin{aligned}\mathbb{P}(X_1 = 1) &= p \\ \mathbb{P}(X_1 = -1) &= q = 1 - p\end{aligned}$$

for  $p \in (0, 1)$ . One can think of  $X_n$  as the net winning of a gambler at the  $n$ -th round. The gambler starts with \$  $x$  and he will leave if he reaches \$ 0 or \$  $N$  for  $0 < x < N$ . The wealth process is

$$S_n = x + \sum_{i=1}^n X_i, \quad n \geq 0$$

From now on, we will let  $\{X_n\}$  be the reference process. Then  $S = \{S_0, S_1, \dots\}$  is adapted. Define the stopping time

$$\tau = \inf\{n \geq 0 : S_n \in \{0, N\}\}$$

We are interested in  $\mathbb{P}(S_\tau = 0)$ .

- CASE 1:  $p = q$ .

The wealth process is a martingale itself. By the Optional Sampling Theorem,

$$\mathbb{E}[S_\tau] = \mathbb{E}[S_0] = x$$

And  $0 \leq S \leq N$  until the game ends (Condition 2) but  $S_\tau \in \{0, N\}$  so

$$x = \mathbb{E}[S_\tau] = N\mathbb{P}(S_\tau = N)$$

It follows that

$$\mathbb{P}(S_\tau = 0) = 1 - \mathbb{P}(S_\tau = N) = 1 - \frac{x}{N} = \frac{N - x}{N}$$



- CASE 2:  $p \neq q$ .

Now the wealth process is not a martingale. Consider Wald's Martingale

$$M_n = \frac{1}{\phi^n(\theta)} e^{\theta S_n}$$

for fixed  $\theta$  and  $n \geq 0$  where

$$e^{\theta x} = \mathbb{E}[M_0] = \mathbb{E}[M_\tau] = \mathbb{E} \left[ \frac{1}{\phi^\tau(\theta)} e^{\theta S_\tau} \right]$$

However now  $\phi^\tau(\theta)$  cannot be taken outside the expectation. Since  $\theta$  is arbitrary, choose  $\theta$  such that  $\phi(\theta) = 1$ , i.e.  $\theta \in \{0, \log(p/q)\}$ . With  $\theta = 0$ , we have a constant martingale, which is useless.

Therefore, let  $\theta = \log(q/p)$  so

$$M_n = \left( \frac{q}{p} \right)^{S_n}$$

is a martingale and

$$\begin{aligned} \left( \frac{q}{p} \right)^x &= \mathbb{E}[M_0] = \mathbb{E}[M_\tau] = \mathbb{E} \left[ \left( \frac{q}{p} \right)^{S_\tau} \right] \\ &= \mathbb{P}(S_\tau = 0) + \left( \frac{q}{p} \right)^N \mathbb{P}(S_\tau = \left( \frac{q}{p} \right)^N) \end{aligned}$$

And  $\mathbb{P}(S_\tau = 0) + \mathbb{P}(S_\tau = N) = 1$  so

$$\mathbb{P}(S_\tau = 0) = \frac{(q/p)^x - (q/p)^N}{1 - (q/p)^N}$$

**Remark:** notice  $1 < M_n < (q/p)^N$  up to time  $\tau$  so we have again satisfied Condition 2.

We can define analogous OST for sub and supermartingales:

- Submartingale:  $\mathbb{E}[X_{\tau \wedge n}] \geq \mathbb{E}[X_0]$
- Supermartingale:  $\mathbb{E}[X_{\tau \wedge n}] \leq \mathbb{E}[X_0]$

## Stochastic Optimization

There are two main branches in Stochastic Optimization:

1. Optimal Stopping (e.g. Secretary Problem)
2. Optimal Control (e.g. utility maximization, Kelly criterion)

Let's solve some of them with Martingales.

**Optimal Stopping (Finite Horizon):** Let  $(X_0, X_1, \dots)$  be some process and  $N$  a fixed terminal time.

$$V(x) = \max_{0 \leq \tau \leq N} \mathbb{E} \left[ H(X_\tau) + \sum_{k=1}^{\tau-1} G(X_k) \mid X_0 = x \right]$$

where  $H(X_\tau)$  is a terminal cost and  $\sum_{k=1}^{\tau-1} G(X_k)$  is a running cost.

We can define this recursively

$$V_n(x) = \max_{0 \leq \tau \leq N} \mathbb{E} \left[ H(X_\tau) + \sum_{k=1}^{\tau-1} G(X_k) \mid S_n = x \right]$$

so our DPE is

$$\begin{cases} V_n(x) = \max(H(x), \mathbb{E}[G(x) + V_{n+1}(X_{n+1}) \mid X_n = x]) \\ V_N(x) = H(x) \end{cases}$$

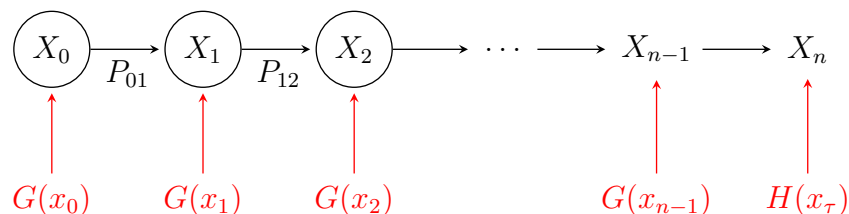
## 2.7 Oct 18

### Optimal Stopping (Finite Horizon)

Let  $(X_0, X_1, \dots)$  a Markov chain with  $P_{xy}$  transition matrix.

$$V(x) = \max(H(x), E^x[G(x) + V_{n+1}(X_{n+1})])$$

where  $N$  is a fixed terminal time,  $\tau$  is a stopping time, and  $E^x$  is the expectation given  $X_0 = x$ .



We can define value functions

$$V_n(x) = \max_{n \leq \tau \leq N} \mathbb{E} \left[ H(x_\tau) + \sum_{k=n}^{\tau-1} G(x_k) \middle| X_n = x \right]$$

for all  $0 \leq n \leq N$  and all  $x$  with  $V(x) = V_0(x)$ .

At time  $n$  with  $X_n = x$ , we can

1. Stop, with payoff  $H(x)$
2. Continue, with payoff

$$G(x) + \mathbb{E}[V_{n+1}(x_{n+1}) \mid X_n = x] = G(x) + \sum_y P_{xy} V_{n+1}(y)$$

where  $G(x)$  is the payoff at time  $n$  and  $V_{n+1}(y)$  is the optimal payoff for  $n+1$  given  $X_{n+1} = y$ .

Hence, our DPE is

$$\begin{cases} V_n(x) = \max \left\{ H(x), G(x) + \sum_y P_{xy} V_{n+1}(y) \right\} \\ V_N(x) = H(x) \end{cases}$$

At last, we are ready to justify DP.

**Theorem:** Suppose  $\{v_0(x), v_1(x), \dots, v_N(x)\}$  is a solution to the DPE. Then  $v_n(x) = V_n(x)$  for all  $n$ . Moreover, define  $\tau^* = \inf\{n \geq 0 : V_n(x) = H(x)\}$ . Then  $\tau^*$  is an optimal stopping time.

*Proof:* Define that is,  $M_n$  is a Markov Chain which behaves optimally for  $m \geq n$ .

We claim that  $\{M_n\}$  is a supermartingale.

$$\begin{aligned}
\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \mathbb{E} \left[ \sum_{k=0}^n G(x_k) + v_{n+1}(x_n) \mid \mathcal{F}_n \right] \\
&= \sum_{k=0}^n G(x_k) + \mathbb{E}[v_{n+1}(x_{n+1}) \mid \mathcal{F}_n] && (x_{0:n} \in \mathcal{F}_n) \\
&= \sum_{k=0}^n G(x_k) + \mathbb{E}[v_{n+1}(x_{n+1}) \mid x_n] && (\text{Markov property}) \\
&= \sum_{k=0}^{n-1} G(x_k) + G(x_n) + \mathbb{E}[v_{n+1}(x_{n+1}) \mid x_n] \\
&\leq \sum_{k=0}^{n-1} G(x_k) + v_n(x_n) && (\text{DPE}) \\
&= M_n
\end{aligned}$$

Then, by the Optional Sampling Theorem, for any  $0 \leq \tau \leq N$ ,  $\mathbb{E}[M_\tau] \leq \mathbb{E}[M_0]$ .

Therefore,

$$\mathbb{E} \left[ \sum_{k=0}^{\tau-1} G(x_k) + v_\tau(x_\tau) \right] \leq v_0(x_0)$$

But in the DPE, clearly  $v_n(x) \geq H(x)$  for all  $n$  so

$$\begin{aligned}
\mathbb{E} \left[ \sum_{k=0}^{\tau-1} G(x_k) + H(x_\tau) \right] &\leq \mathbb{E} \left[ \sum_{k=0}^{\tau-1} G(x_k) + v_\tau(x_\tau) \right] \\
&\leq v_0(x_0)
\end{aligned}$$

and

$$v_0(x_0) \geq V_0(x)$$

But if  $\tau^* = \inf\{n \geq 0 : v_n(x) = H(x)\}$ , then clearly  $v_{\tau^*}(x_{\tau^*}) = H(x_{\tau^*})$  so

$$\mathbb{E} \left[ \sum_{k=0}^{\tau-1} G(x_k) + H(x_\tau) \right] = \mathbb{E} \left[ \sum_{k=0}^{\tau-1} G(x_k) + v_\tau(x_\tau) \right]$$

Finally, for  $n < \tau^*$ ,  $v_n(x_n) \neq H(x_n)$  so  $v_n(x_n) = G(x_n) + \mathbb{E}[v_{n+1}(x_{n+1}) \mid x_n]$  so

$$\sum k = 0^{n-1} G(x_k) + v_{\tau^*}(x_{\tau^*}) = v_0(x_0)$$

and  $M_n$  is a martingale.