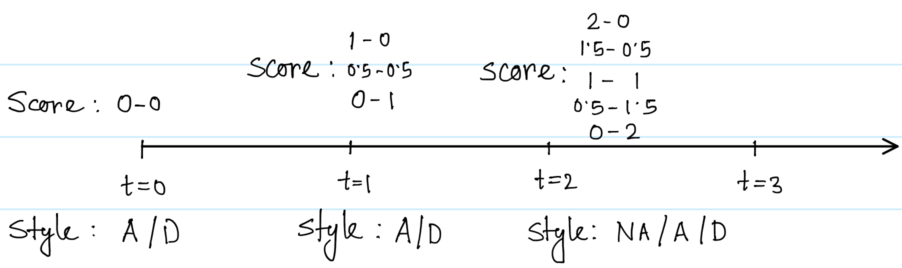
Lecture 1

Thursday, 1 December 2022

Consider a 2 player board game, where each game can have a winner else the game is a draw. There are two style of playing the game, attacking and defensive. If you play in attacking mode you win the game with probability p_w and lose the game with probability $1 - p_w$. If you play in defensive mode the result will be a draw with probability p_d and lose the game with probability $1 - p_d$. Winner gets 1 point, draw gives 0.5 whereas a lose adds nothing to the score. Suppose you are playing this a match of 2 games with someone. If the match is a tie after 2 games the match enters in sudden death mode, i.e., next whoever wins first, wins the match.

Find the best strategy to play this match.

- What is strategy?
  - For each game you have to decide your style of play.
- Should your decisions be influenced by current situation?
  - Yes. Same as closed loop control vs open loop control.



We will see how to solve such sequential decision making problems where the environment is random.

'Solve' in the sense that given a criterion to <u>optimize</u> in <u>Expectation</u>, we want to find the 'policy' that does it.

alternative word for strategy.

What kind of object 'policy' is?

What kind of object 'policy' is ?

– policy is a mapping from 'situation' to 'decision'

In this course we will assume some good property on the randomness of the environment. —"Markov."

Examples of real life problem which we can model using MDP:

1. Investment Portfolio management.
2. Periodic Maintainence activity.
3. Route Planning
4. Inventory Management.
5. Dynamic Pricing
6. Queueing.

* <u>Example in inventory management</u>:

Consider a store and a particular product, 'p' it sells.
Consider N periods of taking ordering decisions.
Let C be the capacity of the store corresp. to product 'p'.

x_k : Stock available at the beginning of k^{th} period.
u_k : stock ordered " " " " " "
assume that order is delievered instantaneously.
w_k : demand during k^{th} period ; w_k, $k = 0, 1, \ldots, N-1$ are independent random variable.

$$x_{k+1} = \underbrace{x_k}_{\text{state variable}} + \underbrace{u_k}_{\text{decision variable}} - \underbrace{w_k}_{\text{randomness in the system}}$$

a. Negative stock ⇒ backlogged demand → shortage cost.
Positive stock → excess inv...

a. Negative stock ⇒ backlogged demand → shortage cost.
Positive stock ⇒ excess inventory → holding cost.
Let $r(x_k)$ captures above two costs.

b. Purchasing cost : cu_k where c is unit price

c. Terminal cost : $R(x_N)$

- We want to minimize $R(x_N) + \sum_{k=0}^{N-1} (r(x_k) + cu_k)$

but this is random variable so we want

$$\min_{\substack{u_k,\ k=0,1,\ldots,N-1 \\ 0 \leq u_k \leq C - x_k}} \ \mathbb{E}_{w_k,\ k=0,\ldots,N-1} \left[R(x_N) + \sum_{k=0}^{N-1} (r(x_k) + cu_k) \right]$$

Review of Probability Theory :

Function: A mapping $f : X \to Y$ is called a function if $x = y$ then $f(x) = f(y)$.

X is called domain of f.

$f(X) = \{ f(x) : x \in X \}$ is called range of f.

Sequence : A real-valued sequence is a function from $\mathbb{N}$ to $\mathbb{R}$.

Example - $f(n) = \frac{1}{n}$ is a sequence. We denote as $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$

Sample Space : Sample space is the collection of all possible outcomes of a random experiment. Sample space is denoted by Ω.

Sample point : Elements of sample space are called sample points. $\omega \in \Omega$ is a sample point.

Example : Tossing a coin once is a random experiment. Head and tail are only possible outcomes.

Head and tail are only possible outcomes.

Therefore, $\Omega = \{H, T\}$ → Notation for elements in this abstract space.

H & T are two sample points here.

Another Example: Consider the random experiment of tossing a coin thrice. The sample space is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

So far we have not talk about probability. We assign probability to subsets of Ω (Events) in such a way that

1. $\mathbb{P}(\Omega) = 1$
2. $\mathbb{P}(A) \geq 0$, $A \subset \Omega$
3. $\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)$, $A_i \in \Omega$ $\forall i \in \mathbb{N}$ and A_i's are mutually disjoint.

– We call these 'Axioms of probability'

We denote the set of all events by $\mathcal{F}$, and we will call $(\Omega, \mathcal{F}, \mathbb{P})$ as probability space.

Back to examples: If we specify that for the coin tossed, probability of head is $p \in [0,1]$ that's enough to assign probability to each sample point in both the sample spaces. Verify this assigned probability obeys all three axioms.

Some Properties:

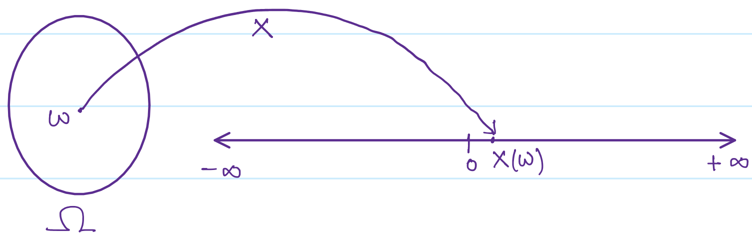
1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. If $A \subset B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$
3. $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$ if A_i's are mutually disjoint.
4. $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$

Random Variable: A random variable is a function from the sample space to the real line.

$$X : \Omega \to \mathbb{R}$$

Example: Lets define a random variable $X : \{H, T\} \to \mathbb{R}$ such that

$$X(H) = 1, \quad X(T) = 0$$



Probability Law: Given a random variable $X : \Omega \to \mathbb{R}$, the probability measure gives rise to a probability law P_X on $X(\Omega)$, i.e.,

$$P_X = \mathbb{P} \circ X^{-1}$$

$$\cdot \quad P_X(B) = \mathbb{P}(X^{-1}(B)) \qquad \text{for every } B \subset X(\Omega) \subset \mathbb{R}$$

Expectation: Expectation of a random variable $X : \Omega \to \mathcal{X} \subset \mathbb{R}$ is defined as

$$\mathbb{E}X = \sum_{x \in \mathcal{X}} x \, P_X(\{x\})$$

Some Properties

1. If $g : \mathcal{X} \to \mathbb{R}$ then $g(X)$ is a random variable and

$$\mathbb{E}\, g(X) = \sum_{y \in g(\mathcal{X})} y \, P_{g(X)}(y)$$

$$= \sum_{x \in \mathcal{X}} g(x) \cdot P_X(x)$$

2. Expectation is linear. Let $a, b \in \mathbb{R}$, X, Y are random variables on Ω, then

variables on Ω, then

$$\mathbb{E}[aX + bY] = a\mathbb{E}X + b\mathbb{E}Y$$

3. Monotonicity: If $\mathbb{P}\{X \geq Y\} = 1$ and $\mathbb{E}Y$ is well defined with $\mathbb{E}Y > -\infty$, then $\mathbb{E}X \geq \mathbb{E}Y$.

Conditional Probability: Let $A, B \subseteq \Omega$ then define conditional probability of event A given event B where $\mathbb{P}(B) \neq 0$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Verify that $\mathbb{P}(\cdot|B)$ obeys the axioms of probability.

Conditional Expectation: Let X, Y are two random variables defined on same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X : \Omega \to \mathcal{X} \subset \mathbb{R}$, $Y : \Omega \to \mathcal{Y} \subset \mathbb{R}$. We define a random variable $\mathbb{E}[X|Y]$ called conditional expectation of X given Y such that

$$\mathbb{P}\left(\mathbb{E}[X|Y] = \sum_{x \in \mathcal{X}} x \cdot P(X = x | Y = y)\right) = P_Y(y)$$

$$= P(\{\omega \mid Y(\omega) = y\})$$

Example: Let X and Y take values on $\{1, 2\}$ and $\{1, -1\}$ respectively with joint prob. mass function P_{XY} s.t.

$$P_{XY}(1,-1) = \frac{1}{5}, \quad P_{XY}(1,1) = \frac{1}{3}$$

$$P_{XY}(2,-1) = \frac{1}{5} \quad P_{XY}(2,1) = \frac{4}{15}$$

Find $\mathbb{E}[X|Y]$ and $\mathbb{E}[Y|X]$.

Ans. $\mathbb{E}[X|Y] = \begin{cases} \frac{3}{2} & \text{w.p. } \frac{2}{5} \\ \frac{13}{9} & \text{w.p. } \frac{3}{5} \end{cases}$ $\quad \mathbb{E}[Y|X] = \begin{cases} \frac{3}{2} & \text{w.p. } \frac{8}{15} \\ \frac{10}{7} & \text{w.p. } \frac{7}{15} \end{cases}$