CMPUT 655: RL–1 Lecture 2 (Part 2)

9th September 2020

Contents

We will talk about

- Random variables
- Conditional probability
- Expectation of random variables
- Tiny introduction to the notation used in the RL book
- And discuss these a little bit in the context of MDPs

Some Resources on Probability

- Resources: Document by Rupam (shared on slack by Andy)
- Chapter 2 and Appendix of ML Notes (`https://marthawhite.github.io/mlcourse/notes.pdf`)
- For a more rigorous introduction, check Chapter 2 of the Bandit book (`https://tor-lattimore.com/downloads/book/book.pdf`)

Random Variables

$P(\{H\}) = 1/2$

$P(\{H, T\}) = 1$

R.V. value set of value

$X \quad x \quad \mathcal{X}$

$x \in \mathcal{X}$

$\{H, T\} \longrightarrow \mathbb{R}$

Example: Coin toss experiment

X : # of heads

$\{H\} \longrightarrow 1$

$\{T\} \longrightarrow 0$

Conditional Probability

$$P(A \mid B) \doteq \frac{P(A \cap B)}{P(B)}$$

Conditional Probability: Dice Example

$$S = \{1, 2, 3, 4, 5, 6\}$$

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

$$P(\{1, 2, 3\}) = 1/2$$

P(even # | divisible by 3)

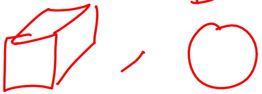
$$P(\{2, 4, \underline{6}\} \mid \{3, \underline{6}\}) = \frac{P(\{6\})}{P(\{3, 6\})} = \frac{1}{2}$$

Joint Probability and Marginalization

$$P(A, B) = P(A)\,P(B) \quad \text{[Independent]}$$

A

B



$$P(\{1,2\}, \{H\}) = P(\{1,2\}) \times P(\{H\}) = \frac{2}{6} \times \frac{1}{2}$$

$P(A,B)$	diff. comb. A, B
⋮	⋮

$$P(B=\{H\}) = P(\{1\}, \{H\}) + P(\{2\}, \{H\}) + P(\{3\}, \{H\}) + \cdots$$

$$\boxed{P(A) = \sum_b P(A, B=b)}$$

$$P(A) = \sum_b \sum_c P(A, B=b, C=c)$$

Chain Rule of Probability

$P(A, B)$

dice $\nearrow$ $\nwarrow$ coin

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

$$\Rightarrow \boxed{P(A,B) = P(A|B)\,P(B)}$$

$$p(x_1, x_2, x_3) = p(x_3|x_2, x_1) \cdot p(x_2|x_1) \cdot p(x_1)$$

$$p(x_1, x_2, x_3) = p(x_1|x_2, x_3)\,\underbrace{p(x_2, x_3)}$$

$$= \boxed{p(x_1|x_2, x_3)\,p(x_2|x_3)\,p(x_3)}$$

$$\underbrace{p(x_1, x_2|x_3)}\,p(x_3)$$

$$= \boxed{p(x_1|x_2, x_3)\,p(x_2|x_3)\,p(x_3)}$$

Expectation of Random Variables

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x \cdot p(X = x)$$

$$\mathbb{E}[X] = \int_{\mathcal{X}} x \, p(x) \, dx$$

Expectation of Random Variables: Dice Example

$X = \#$ on dice

$$\mathbb{E}[X] = \sum_{x \in \{1,2,3,4,5,6\}} x \cdot p(X = x)$$

$$= \frac{1+2+3+4+5+6}{6}$$

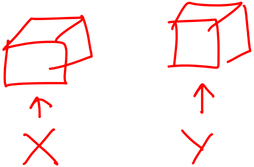
$$= 3.5$$

Linearity of Expectation and another Dice Example

$$\mathbb{E}[cX] = c\,\mathbb{E}[X]$$

constant

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$



$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$= 3.5 + 3.5$$

Conditional Expectation and Law of Total Expectation

$$\mathbb{E}[X \mid Y=y] = \sum_x x \, p(X=x \mid Y=y)$$

$$\mathbb{E}_x[X] = \mathbb{E}_y\left[\mathbb{E}_x[X \mid Y=y]\right]$$

$$\sum_x p(x) \times x$$

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

$$= \sum_y \sum_x x \, p(x \mid y) \, P(y)$$

$$= \sum_x x \sum_y \underbrace{p(x \mid y) \, p(y)}_{p(x,y)}$$

$$\sum_x x \sum_y p(x,y)$$

Example MDP

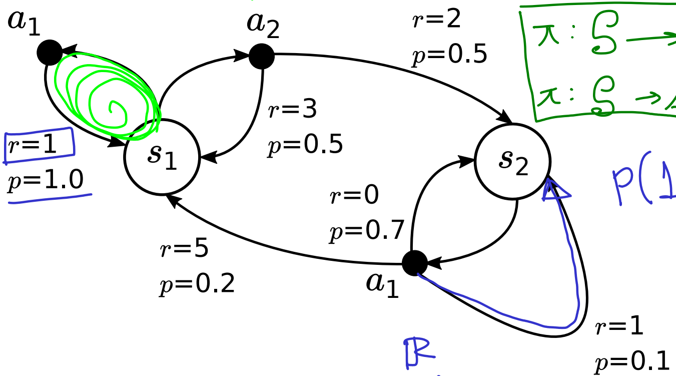
environment

$\pi(a_1 \mid s_1) = 1$

$\pi(a_2 \mid s_1) = 0$ → $s_1, a_1, 1, s_1, a_1, 1, s_1, \ldots$

$\pi : \mathcal{S} \to \mathcal{A}$ deterministic

$\pi : \mathcal{S} \to \Delta(\mathcal{A})$ (stochastic)



$p(1, s_2 \mid s_2, a_1) = 0.1$

$\mathbb{R}$

Figure: MDP $M = (\mathcal{S}, \mathcal{A}, \mathcal{R}, P, \gamma)$.

$\{s_1, s_2\}$ $\{a_1, a_2\}$

transition prob.

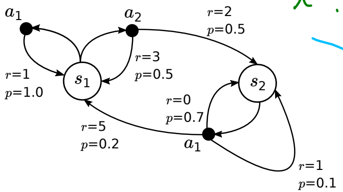
$\boxed{p(s', r \mid s, a)}$

Agent's Interaction and the Trajectory ($\sim$ stream of experience)

$h_1 = S_1 \; S_2 \; S_3 \; S_1 \Rightarrow$

vs

$h_2 = S_1 \; S_3 \; S_1 \Rightarrow$



π^*: deterministic & stationary

Figure: Trajectory $\tau_t = (S_0, A_0, R_1, S_1, A_1, R_2, S_2, \ldots, S_T)$.

$P(A_1 \mid S_0 A_0 R_1 S_1)$

Let us calculate the probability of seeing this trajectory:

$$P(\tau) = P(S_0 \; A_0 \; R_1 \; S_1 \; A_1 \; R_2 S_2 \ldots)$$

$$= P(S_0) \times \pi(A_0 \mid S_0) \times P(R_1 S_1 \mid S_0 A_0) \times \pi(A_1 \mid S_1)$$

$$\times \; P(R_2, S_2 \mid S_1 A_1) \times \ldots$$

$\rightarrow$ Markov property.

RL Book Notation

- Mathcal (fancy) Symbols:

  $\mathcal{S}$ $\mathcal{A}$ $\mathcal{R}$ → sets

- Capital Symbols:

  S_1 A_1 R_1 → random/placeholders

- Small Symbols:

  s_0 s_1

  $S_1 = s_1$

Different Transition Probabilities (Section 3.1 RL Book[1])

- $p(s', r | s, a) \doteq \Pr\left(S_{t+1} = s', R_{t+1} = r \,\middle|\, S_t = s, A_t = a\right)$

- $p(s' | s, a) =$

- (Dependent on the policy π)
  $p(s' | s) =$

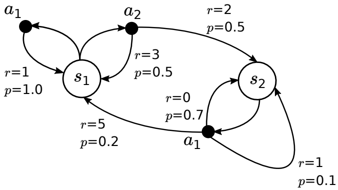


Figure: MDP

[1] http://incompleteideas.net/book/the-book-2nd.html

Different Reward Functions (Section 3.1 RL Book[2])

- The random variable R

- Reward function

$$r(s,a) \doteq \mathbb{E}\left[R_t \middle| S_{t-1}=s, A_{t-1}=a\right]$$
$$=$$

- $r(s,a,s') =$

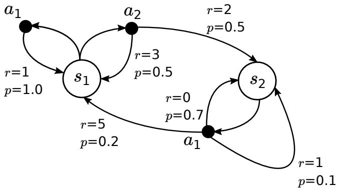


Figure: MDP

[2] http://incompleteideas.net/book/the-book-2nd.html

Summary