Latex Notes

Asif Hamid

October 29, 2022

Contents

Chapter 1

Introduction

1.1 Introduction to reinforcement learning

1.1.1 Definitions

Reinforcement learning is:

- **agent-oriented learning:** learning by interacting with an environment
- **trial and error** only given delayed evaluative feedback
- **science of the mind** one which is neither natural science nor applied technology

Framework:

1. agent percieves the **state** of the environment
2. based on the state, it chooses an **action**
3. the action gives the agent a **reward**
4. a **policy** aims to maximize the agent's **long term expected reward**

1.2 Bandit

1.2.1 Definition

One-armed bandit Simplest RL problem

- pull the lever
- get some reward
- choose the best lever!

k-armed bandit extends to k arms

- at every time step t, choose an action A_t from k possibilties
- recieve a reward R_t dependent only on the action taken (i.i.d)

- $q_a = \mathbb{E}[R_t | A_t = a], \forall a \in 1, \ldots k$

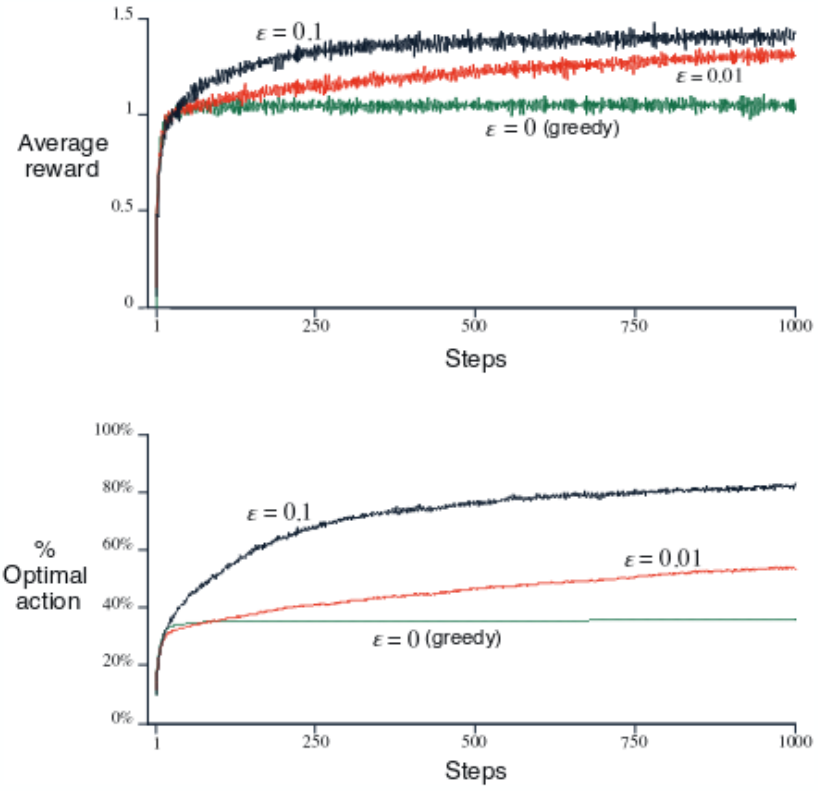

Figure 1.1: ϵ-greedy methods on 10-arm bandit

1.3 Notes for latex class

Observe the two dollar sign for the command epsilon in the caption. Equation ;

$$\begin{aligned} A &= \frac{\pi r^2}{2} \\ &= \frac{1}{2}\pi r^2 \end{aligned} \tag{1.1}$$

Note here the use of split env. Also the & operator which is used for alignment. If you can not understand this we will discuss this on monday. Another equation

example;

$$f(x) = \frac{P(x)}{Q(x)} \qquad and \qquad f(x) = \tfrac{P(x)}{Q(x)} \tag{1.2}$$

Note how I have used the command hspace in (1.2) as we discussed in class on *friday* to insert a gap for and.

The following equation is for the summation and integration, you can follow the same and adapt for different equations; This is to use integration;

$$y = \int_a^b x^2 \, dx$$

$$z = \oint_V f(s) \, ds$$

Observe the use of as we discussed in the class. Also, try using **Split env** for the above two equations.

This one for summation;

$$\sum_{n=1}^{\infty} 2^{-n} = 1 \tag{1.3}$$

In (1.3), analyse the use of underscores and curly braces. Try to understand this or we will discuss it in next class.

We insert the limits with the equation env but we can also insert them in the text as Limit $\lim_{x \to \infty} f(x)$ inside text.

Chapter 2

Deep neural network

2.1 Motivation

Machine learning

this is the first example.
text text this is the next section.

$$y = 3x$$

$$x = 3x \tag{2.1}$$

In equation 2.1 we discussed a linear system of the form $x = 2x \nleq y$.

Observe the changes in the table code i have made. centering for center alignment, use of hspace and use of **Table env** for caption and labeling. **Tabular env** is for inserting table only. I forgot to mention this in the class.

In section 2.1 we discussed the motivation for the report.

1. Detergent
2. Rice.
3. Dal.
4. fruits.

The items have following prices;

Table 2.1: List of prices

First	second	third	fourth	fifth
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•

- Price of item-2 is Rs 70/kg.
- 60
- 20

The next section discusses further prices.
In Figure 1.1 we have shown the results.
This is an example of latex.