ECE 566 Project Proposal: Markov Decision Process and its Application in Finance

Ranvir Rana, Sakshi Agarwal

1 Introduction

Planning under uncertainty is fundamental to solving many important real-world problems, including applications in robotics, network routing, scheduling, and financial decision making. Markov Decision Processes provide a formal framework for modeling these tasks and for deriving optimal solutions. Markov Decision Process provide a framework for Decision making in various situations where the outcome is partially random and partially under the control of the decision maker. MDPs are applicable when the effects of actions follow Markov Property: i.e. the effect of an action taken at a state depend only on that state and not on the prior history.

2 Applications

MDPs are used in various areas like Computer game-playing programs, scheduling algorithm for data centers, financial engineering applications, robotic navigation and control, etc. There are various variants of MDPs that are used in Reinforcement learning which is a growing field owing to its use in automated driving vehicles. We will focus on the application of MDP to finance especially the Consumption-Investment Problem, which deals with a person who wants to increase his wealth at each time step. He wishes to allocate his wealth between consumption and investment with an objective to maximize his consumption utility.

3 Consumption Investment Problem

A Consumption-Investment problem is defined as follows: Suppose at time t , the current wealth x_t generates output $h(x_t)$ (h is a production function) and a part of it a_t, is consumed, and the rest $i_t = h(x_t) - a_t$ is invested. This investment will lead to another wealth x_{t+1} at time t+1, which is given by the equation:

$$x_{t+1} = \xi_t(h(x_t) - a_t)$$

where $\{\xi_t\}$ is a sequence of i.i.d. random variables. ξ_t can be thought of as the return per invested dollar.
The aim is to maximize the investor's consumption utility over all policies $\pi \in \Pi$:

$$v(\pi, x) = E_x^{\pi}[\sum_{t=0}^{\infty} \alpha^t U(a_t)]$$

where U is the utility function and α is the discount factor.

4 Problem Formulation

The problem can be formulated as a continuous state Markov Decision Process as follows:
$S = x_t$ are the continuous states, $A = a_t$ are the continuous actions
$P(s, a, s') = P(\xi_t = \frac{x_{t+1}}{h(x_t) - a_t})$ is the transition probability.
$R(s, a, s') = U(a_t)$ is the reward function

We optimize the value function written above using discretization techniques to solve continuous MDPs.

5 Analyzing different business scenarios

The production function $h(x_t)$ can take different shapes in different real life scenarios. We will run the MDP for different production functions and observe the consumption utility w.r.t time for different production functions.