MARKOVIAN LEARNING

(OUTLINE OF PRESENTATION)

COIN TOSSING CHOICE

MOTIVATING EXAMPLE

EXAMPLE: Coin 1 has Prob.of Heads (win) 1/10

Coin 2 has Prob.of Heads (win) 6/10

I do not know the 1/10 and 6/10. I toss the coins many times and each time I win or loose. Each time I try to choose that which seems to have the better probability of success.

QUESTION :How to do that in a quick and sure way?

Main Math Problem

For n Coins, the problem is equivalent to finding the gobal maximum of a function f with domain {1,2,…n} and when I choose i=k, (i.e. coin k), I learn not f (k) but the outcome of an experiment which has probability of success f (k).

Clearly a hard problem, so do not expect fast convergence!

It can be handled with the Robbins –Monroe and Kiefer – Wolfowitz methodologies, but notice the special structure.

$$\max(\sum_{i=1}^{n} x_i d_i), \text{subject to: } \sum_{i=1}^{n} x_i = 1, x_i \geq 0, Prob(Coin(i) = success) = d_i)$$

Or

$$\max E[\sum_{i=1}^{n} x_i y_i], \text{subject to : } \sum_{i=1}^{n} x_i = 1, x_i \geq 0, Prob(y_i = 1) = d_i, \quad Prob(y = 1) = 1 - d_i$$

Notice that the d_i's are unknown probabilities!

Learning Methodology Based on Animal Learning

We choose a Coin at each instant of time (discrete), toss it and if the outcome is success(failure) we do not necessarily choose the same(other) coin next time ,but just increase(decrease) the chance of choosing it ,i.e.

Prob(of choosing Coin i at time t=k+1)=:

Prob(of choosing Coin i at time t=k) +θ(i,Success,k), if Coin i gave Success at time t=k (Reward!)

Prob(of choosing Coin i at time t=k) –θ(I,Failure,k), if Coin i gave Failure at time t=k (Penalty!)

The θ's are positive quantities which may depend on time k as well as on the Prob (of choosing Coin i at time t=k) (and perhaps on the previous probabilities used ,the previous Coins used and the history of success or failures associated with each Coin choice) and of course the resulting new probabilities are to be nonnegative and add up to one.

MATERIAL FROM S.LAKSHMIVARAHAN
"LEARNING ALGORITHMS THEORY AND APPLICATIONS"
SPRINGER,1981

- Ch.1; 1.1,1.2,1.3,1.4
- Ch.2; 2.1,2.2,2.3(up to page 34),2.5,2.6
- Ch.3; 3.1,3.2
- Ch.4; 4.1,4.2(up to page 112)

HOMEWORK PROBLEM

We have three Coins with Probability of success for each one:1/4,2/4,3/4.

Apply the Ergodic Algorithm for solving “it”, with several (at least three) admissible choices of Reward –Penalty functions and stepsizes. Use constant and time varying stepsizes. Compare and comment.