16/10/2014
MPRO - Stochastic Optimization

Exercises inspired by "Dynamic Programming and Optimal Control" by Dimitri Bertsekas.

Exercise 1. *Consider the dynamic system given by*

$$x_{t+1} = x_t + u_t, \qquad x_0 = 5 \ ,$$

and the cost function (to be minimized)

$$\sum_{t=0}^{3} u_t^2 + x_3^2 \ .$$

The constraint set is given by

$$\Big\{ u \in \mathbb{Z} \quad | \quad 0 \leq x_t + u \leq 5 \Big\} \ .$$

1. *Find the optimal value and controls by use of Dynamic Programming.*
2. *Now we consider, for this question only, an additional constraint on the final state :* $x_3 = 5$*, and solve the problem by use of Dynamic Programming.*
3. *We now consider a stochastic version of our problem. The dynamic is given by*

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \boldsymbol{U}_t + \boldsymbol{W}_t, \qquad x_0 = 5 \ ,$$

and the cost function by

$$\mathbb{E}\Big[\sum_{t=0}^{3} \boldsymbol{U}_t^2 + \boldsymbol{X}_3^2 \Big] \ ,$$

where $\boldsymbol{W}_t$ *is* $+1$ *or* -1 *with probability* $1/2$ *(except if* $x_t + u_t$ *is* 0 *or* 5 *in which case* $\boldsymbol{W}_t$ *is equal to* 0*). Solve the problem by Dynamic Programming.*

Exercise 2. *Suppose we have a machine that is either running or broken down. If it is running at the beginning of a given week it guarantee a profit of* \$100*. At the beginning of a week we can either perform maintenance (for a cost of* \$25*) or do nothing. If we perform maintenance there is* 50% *probability that the machine broke at the end of the week, if we didnot the probability is* 80%*. If the machine is broken down we can either repair it for* \$50 *that will put it in a "running" state, or do nothing. Find the optimal repair and maintenance policy that maximizes total expected profit assuming a running machine at the start of first week.*

Exercise 3. *Consider the following dynamical system*

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \boldsymbol{U}_t + \boldsymbol{W}_t, \qquad \boldsymbol{X}_0 = 0 ,$$

with the cost function (to be minimized)

$$\mathbb{E}\Big[\sum_{t=1}^{T} \boldsymbol{X}_t^2\Big] .$$

The control $\boldsymbol{U}_t$ *is assumed to be non-anticipative, that is measurable with respect to the past noises* $\big(\boldsymbol{W}_0, \cdots, \boldsymbol{W}_t\big)$ *that are observable.*

1. *Assume that the noises* $\boldsymbol{W}_t$ *are independent and square integrable of known law. Write the Dynamic Programming equation and find the optimal policy and value of the problem.*
2. *Now, assume that the noise process is an AR-1. More precisely assume that*

   $$\boldsymbol{W}_{t+1} = \alpha \boldsymbol{W}_t + \boldsymbol{\varepsilon}_t ,$$

   where $\big(\varepsilon_t\big)$ *is an i.i.d process. Consider the extended state* $\boldsymbol{Y}_t = (\boldsymbol{X}_t, \boldsymbol{W}_t)$*, write a DP equation and find he optimal policy and value of the control problem.*
3. *How could this be generalized to an AR-n ?*