ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control
IDSC
Institut für Dynamische Systeme und Regelungstechnik

Final Exam **January 27th, 2011**

Dynamic Programming & Optimal Control (151-0563-01) Prof. R. D'Andrea

Solutions

Exam Duration: **150 minutes**

Number of Problems: **4**

Permitted aids: One A4 sheet of paper.

Use only the provided sheets for your solutions.

Problem 1 **25%**

Consider the dynamic system

$$x_{k+1} = x_k + x_{k-1} + u_k + w_k\,, \quad k = 0, 1,$$

where the input u_k is restricted to be 1 or -1, and the disturbance w_k takes values 1 and -1 with equal probability.

Given $x_0 = 0$ and $x_{-1} = 0$, find the optimal control policy that minimizes

$$\sum_{k=0}^{2} (x_k - 1)^2$$

using the dynamic programming algorithm. Furthermore, calculate the corresponding optimal cost.

Solution 1

Introduce a new state variable $y_k = x_{k-1}$. This allows the system to be rewritten as

$$\tilde{x}_{k+1} := \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + y_k + w_k + u_k \\ x_k \end{bmatrix},$$

with initial condition $\tilde{x}_0 = (0,0)$. Now, find all the states $\tilde{x}_k$ that can be reached from $\tilde{x}_0 = (0,0)$ with every possible u_k and w_k for $k = 1,2$.

k=1 and $\tilde{x}_0 = (0,0)$:

$$\begin{aligned} \{u_0 = 1 \quad \text{and} \quad w_0 = 1\} \quad &\Rightarrow \quad \tilde{x}_1 = (2,0) \\ \{u_0 = 1 \quad \text{and} \quad w_0 = -1\} \quad \text{or} \quad \{u_0 = -1 \quad \text{and} \quad w_0 = 1\} \quad &\Rightarrow \quad \tilde{x}_1 = (0,0) \\ \{u_0 = -1 \quad \text{and} \quad w_0 = -1\} \quad &\Rightarrow \quad \tilde{x}_1 = (-2,0) \end{aligned}$$

k=2 and $\tilde{x}_1 = (2,0)$:

$$\begin{aligned} \{u_1 = 1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (4,2) \\ \{u_1 = 1 \quad \text{and} \quad w_1 = -1\} \quad \text{or} \quad \{u_1 = -1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (2,2) \\ \{u_1 = -1 \quad \text{and} \quad w_1 = -1\} \quad &\Rightarrow \quad \tilde{x}_2 = (0,2) \end{aligned}$$

k=2 and $\tilde{x}_1 = (0,0)$:

$$\begin{aligned} \{u_1 = 1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (2,0) \\ \{u_1 = 1 \quad \text{and} \quad w_1 = -1\} \quad \text{or} \quad \{u_1 = -1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (0,0) \\ \{u_1 = -1 \quad \text{and} \quad w_1 = -1\} \quad &\Rightarrow \quad \tilde{x}_2 = (-2,0) \end{aligned}$$

k=2 and $\tilde{x}_1 = (-2,0)$:

$$\begin{aligned} \{u_1 = 1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (0,-2) \\ \{u_1 = 1 \quad \text{and} \quad w_1 = -1\} \quad \text{or} \quad \{u_1 = -1 \quad \text{and} \quad w_1 = 1\} \quad &\Rightarrow \quad \tilde{x}_2 = (-2,-2) \\ \{u_1 = -1 \quad \text{and} \quad w_1 = -1\} \quad &\Rightarrow \quad \tilde{x}_2 = (-4,-2) \end{aligned}$$

Calculate the terminal cost $J_2(\tilde{x}_2) = (x_2 - 1)^2$, for all possible states at $k = 2$.

$$\begin{aligned} J_2((4,2)) &= 9 & J_2((2,0)) &= 1 & J_2((0,-2)) &= 1 \\ J_2((2,2)) &= 1 & J_2((0,0)) &= 1 & J_2((-2,-2)) &= 9 \\ J_2((0,2)) &= 1 & J_2((-2,0)) &= 9 & J_2((-4,-2)) &= 25. \end{aligned}$$

Now, using the the dynamic programming algorithm,

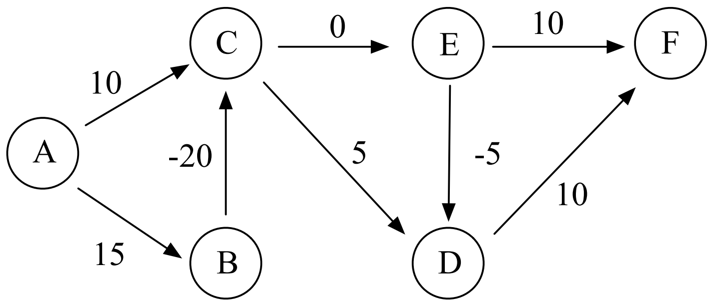
$$J_k(\tilde{x}_k) = J_k((x_k, x_{k-1})) = \min_{u_k \in \{-1,1\}} \underset{w_k}{E} \Big\{ (x_k - 1)^2 + J_{k+1}\big(\tilde{x}_{k+1}\big) \Big\},$$

for k = 1 and 0 gives

$$\begin{aligned} J_1((2,0)) &= 2 \quad \text{with} \quad \mu_1^*((2,0)) = -1, \\ J_1((0,0)) &= 2 \quad \text{with} \quad \mu_1^*((0,0)) = 1, \\ J_1((-2,0)) &= 14 \quad \text{with} \quad \mu_1^*((-2,0)) = 1, \quad \text{and} \\ J_0((0,0)) &= 3 \quad \text{with} \quad \mu_0^*((0,0)) = 1. \end{aligned}$$

Problem 2 25%

Consider the following transition graph.



a) Calculate the optimal cost to go and the shortest path from A to F using the dynamic programming algorithm.

b) Calculate the optimal cost to go and the shortest path from A to F using the label correcting algorithm. Use Breadth-first (First-in/First-out) search to determine at each iteration which node to remove from the `OPEN` bin.

Solve the problem by populating a table of the following form,

Iteration	Node exiting `OPEN`	`OPEN`	d_A	d_B	d_C	d_D	d_E	d_F
0	-	...						
...								

where the variable d_i denotes the length of the shortest path from node A to node i that has been found so far. State the resulting shortest path and the optimal cost-to-go.

c) How would the label correcting algorithm behave if a new arc from D to B is introduced with a weight of 10 ?

Solution 2

a) Dynamic Programming Algorithm:

$$\begin{aligned}
J(F) &= 0 \\
J(D) &= J(F) + 10 = 10 \quad \text{with} \quad \mu^*(D) = D \to F \\
J(E) &= \min\{10 + J(F), -5 + J(D)\} = 5 \quad \text{with} \quad \mu^*(E) = E \to D \\
J(C) &= \min\{0 + J(E), 5 + J(D)\} = 5 \quad \text{with} \quad \mu^*(E) = C \to E \\
J(B) &= -20 + J(C) = -15 \quad \text{with} \quad \mu^*(B) = B \to C \\
J(A) &= \min\{15 + J(B), 10 + J(C)\} = 0 \quad \text{with} \quad \mu^*(A) = A \to B
\end{aligned}$$

Therefore, the optimal cost to go from A to F is 0 and the shortest Path is $A \to B \to C \to E \to D \to F$

b) Label Correcting Algorithm:

Iteration	Node exiting `OPEN`	`OPEN`	d_A	d_B	d_C	d_D	d_E	d_F
0	–	A	0	∞	∞	∞	∞	∞
1	A	B,C	0	15	10	∞	∞	∞
2	B	C	0	15	-5	∞	∞	∞
3	C	D,E	0	15	-5	0	-5	∞
4	D	E,F	0	15	-5	0	-5	10
5	E	F,D	0	15	-5	-10	-5	5
6	F	D	0	15	-5	-10	-5	5
7	D	F	0	15	-5	-10	-5	0
8	F	–	0	15	-5	-10	-5	0

Therefore, optimal cost to go from A to F is 0 and the shortest Path is $A \to B \to C \to E \to D \to F$

c) The new arc will introduce a negative cycle ($D \to B \to C \to D$) and the label correcting algorithm will not terminate.

Problem 3 **25%**

Consider the dynamic system

$$\dot{x}(t) = x(t) + u(t)$$

with the free initial state $x(0) = x_0$ and the fixed terminal state $x(1) = 1$.

Use the Minimum Principle to find an input $u(t)$ and an initial condition x_0 that minimize the cost function

$$\int_0^1 u^2(t)dt.$$

Solution 3

Apply the Minimum Principle:

The system equation is

$$\dot{x}(t) = x(t) + u(t).$$

The Hamiltonian is given by

$$\begin{aligned} H(x(t), u(t), p(t)) &= g(x(t), u(t)) + p(t)f(x(t), u(t)) \\ &= u^2(t) + p(t)x(t) + p(t)u(t). \end{aligned}$$

The adjoint equation can be calculated as follows

$$\dot{p}(t) = -\frac{\partial H}{\partial x}(x^*(t), u^*(t), p(t)) = -p(t).$$

Solving this differential equation leads to

$$p(t) = Ae^{-t}, \quad A \text{ constant},$$

and using the extra boundary condition on the adjoint equation due to the free initial state we find

$$p(0) = 0 \Rightarrow A = 0.$$

The optimal input $u^*(t)$ is obtained by minimizing the Hamiltonian along the optimal trajectory

$$u^*(t) = \arg\min_u(t)(u^2(t) + p(t)x(t) + p(t)u(t)) \Rightarrow 2u^*(t) + p(t) = 0 \Rightarrow u^*(t) = -\frac{1}{2}p(t).$$

Since $p(t) \equiv 0$ the optimal input is

$$u^*(t) = -\frac{1}{2}p(t) \equiv 0.$$

Using the optimal input $u^*(t) \equiv 0$, we can solve the system equation to find the optimal trajectory

$$\dot{x}^*(t) = x^*(t) \Rightarrow x^*(t) = Be^t, \quad B \text{ constant}.$$

With the fixed terminal state $x(1) = 1$ we find

$$Be^1 = 1 \Rightarrow B = e^{-1},$$

and thus the initial condition $x(0) = x_0$ is

$$x(0) = e^{-1}.$$

Problem 4 25%

Fred and George want to play a pursuer-evader game. Their possible locations and the transitions between them are represented by the graph shown in Figure 1.

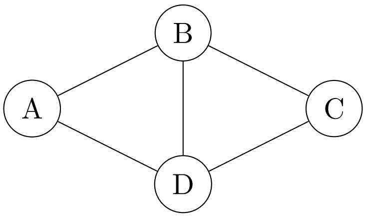


Figure 1

The game evolves in stages where at each stage both of them change their locations simultaneously. Fred is the pursuer and tries to catch George, who is the evader. Fred will have succeeded when he manages to be at the same location as George at the end of a stage. The possibility exists that, when they are adjacent, they both move toward each others' location. This does not result in George being caught. For example, if Fred is at location B and George is at location D and both move to the others' current location, the game will continue and the next stage will start with Fred being at location D and George being at location B.

If George is at location A or C he will move to one of the adjacent locations with equal probability. If he is at location B he will move to location C or D with equal probability and if he is at location D he will move to location A or B with equal probability. Fred also will never stay at his current location. If he is at location A or C he will to move to location B or D. If he is at location B he is only allowed to move to location C or D, and if he is at location D, he is only allowed to move to location A or B.

Fred's objective is to catch George in minimum expected time. The game ends when George is caught. There is no cost involved for moving between different locations.

Formulate the problem as a stochastic shortest path problem

$$x_{k+1} = w_k, \quad k = 0, 1, 2, \ldots,$$

with $x_k,\ w_k \in S,\ u_k \in U(x_k)$ and transition probabilities $p_{ij}(u) = P(w_k = j | x_k = i,\ u_k = u)$, where the objective is to minimize

$$\lim_{N \to \infty} E \left\{ \sum_{k=0}^{N-1} g(x_k, u_k) \right\}.$$

That is, define the set of states S, the control sets $U(x_k)$, the corresponding transition probabilities $p_{ij}(u)$, and the stage costs $g(x_k, u_k)$. Explain your steps. You do ***NOT*** have to solve the problem.

Solution 4

Set of states:

$$S = x_F \times x_G = \{A, B, C, D\} \times \{A, B, C, D\}$$

where x_F represents the location of Fred and x_G represents the location of George. The game ends if one of the terminal states $S_T = \{(A, A), (B, B), (C, C), (D, D)\}$ is reached.

Set of control sets:

$$U(x_F) = \{1, 2\}$$

Since Fred has two options at each location, a single control set is sufficient. If Fred is at location A or C, 1 means that he will take the upper arc to location B, and 2 means that he will take the lower arc to location D. If Fred is at location B, 1 means that he will take the arc to location C, 2 means that he will take the arc to location D. If Fred is at location D, 1 means that he will take the arc to location A, 2 means that he will take the arc to location B.

Transition possibilities:

$$\begin{aligned}
p_{(A,B),(B,C)}(1) = p_{(C,B),(B,C)}(1) = 0.5 \\
p_{(A,B),(B,D)}(1) = p_{(C,B),(B,D)}(1) = 0.5 \\
p_{(A,C),(B,B)}(1) = p_{(C,A),(B,B)}(1) = 0.5 \\
p_{(A,C),(B,D)}(1) = p_{(C,A),(B,D)}(1) = 0.5 \\
p_{(A,D),(B,A)}(1) = p_{(C,D),(B,A)}(1) = 0.5 \\
p_{(A,D),(B,B)}(1) = p_{(C,D),(B,B)}(1) = 0.5
\end{aligned}$$

$$\begin{aligned}
p_{(A,B),(D,C)}(2) = p_{(C,B),(D,C)}(2) = 0.5 \\
p_{(A,B),(D,D)}(2) = p_{(C,B),(D,D)}(2) = 0.5 \\
p_{(A,C),(D,B)}(2) = p_{(C,A),(D,B)}(2) = 0.5 \\
p_{(A,C),(D,D)}(2) = p_{(C,A),(D,D)}(2) = 0.5 \\
p_{(A,D),(D,A)}(2) = p_{(C,D),(D,A)}(2) = 0.5 \\
p_{(A,D),(D,B)}(2) = p_{(C,D),(D,B)}(2) = 0.5
\end{aligned}$$

$$\begin{aligned}
p_{(B,A),(C,B)}(1) = p_{(D,A),(A,B)}(1) = 0.5 \\
p_{(B,A),(C,D)}(1) = p_{(D,A),(A,D)}(1) = 0.5 \\
p_{(B,C),(C,B)}(1) = p_{(D,C),(A,B)}(1) = 0.5 \\
p_{(B,C),(C,D)}(1) = p_{(D,C),(A,D)}(1) = 0.5 \\
p_{(B,D),(C,A)}(1) = p_{(D,B),(A,C)}(1) = 0.5 \\
p_{(B,D),(C,B)}(1) = p_{(D,B),(A,D)}(1) = 0.5
\end{aligned}$$

$$
\begin{aligned}
p_{(B,A),(D,B)}(2) &= p_{(D,A),(B,B)}(2) = 0.5 \\
p_{(B,A),(D,D)}(2) &= p_{(D,A),(B,D)}(2) = 0.5 \\
p_{(B,C),(D,B)}(2) &= p_{(D,C),(B,B)}(2) = 0.5 \\
p_{(B,C),(D,D)}(2) &= p_{(D,C),(B,D)}(2) = 0.5 \\
p_{(B,D),(D,A)}(2) &= p_{(D,B),(B,C)}(2) = 0.5 \\
p_{(B,D),(D,B)}(2) &= p_{(D,B),(B,C)}(2) = 0.5
\end{aligned}
$$

The probabilities for staying at the terminal states are $p_{(A,A),(A,A)}(u) = p_{(B,B),(B,B)}(u) = p_{(C,C),(C,C)}(u) = p_{(D,D),(D,D)}(u) = 1,\ \forall u \in U(x_F)$. All other probabilities are zero.

Cost: (subscript k dropped for convenience)

$$
\begin{aligned}
g(x,u) &= c \quad x \notin S_T,\ u \in U(x_F) \\
g(x,u) &= 0 \quad x \in S_T,\ u \in U(x_F)
\end{aligned}
$$

Since the goal is to minimize the expected time until Fred catches George, c can be any constant positive value but, for simplicity, might be chosen to be 1. As soon as Fred has caught George, we assume zero cost to get a well-defined problem with finite overall cost.