

6.231 Dynamic Programming and Optimal Control - Homework 5

7.2 (The Strategic Quarterback)

(a) We formulate the problem as the following stochastic shortest path problem:

States (nodes). The states are characterized by three parameters (k, i, l) , where

- $k = 0, 1, 2, 3, 4$: Yardage status at the end of the k -th down. $k = 0$ represents the status in the beginning of a play.
- $i = 1, 2, \dots, 120$: Distance to the goal (where 120 is the length of the football field).
- $l = 0, 1, 2, \dots, (120 - i)$: Number of yards gained in the current 4-down cycle. (Note that you cannot have gained more than $120 - i$ yards if you're already i yards away from the goal. Note also that for $k = 0$, which indicates the very beginning of a play, only $l = 0$ is permissible.)

These states have no stage cost.

In addition, there are also two termination states:

- T: the state indicating that a touchdown has been scored, with stage cost 1.
- L: the state indicating that the play has been lost (i.e. the ball is turned over to the opponent), with stage cost 0.

State transition probabilities (arcs). Let $f_\lambda(i) = \frac{\lambda^i e^{-\lambda}}{i!}$ be the probability of drawing the integer i from a Poisson random variable with rate λ . Let $F_\lambda(i) = \sum_{j=0}^i f_\lambda(j)$ be the corresponding cumulative density function.

When in state (k, i, l) , with $k = 0, 1, 2, 3$, consider the following two policies:

- Running:
 - With probability $f_{\lambda_r}(j)$, we progress j yards, and move to state $(k+1, i-j, l+j)$. ($j = 0, 1, \dots, i-1$)
 - With probability $1 - F_{\lambda_r}(i-1)$, we progress i yards or more, and move to state T.

- Passing:
 - With probability p , the pass will be incomplete, and we remain at the same yardage, moving to state $(k + 1, i, l)$.
 - With probability q , the ball will be intercepted, and we move to state L.
 - With probability $(1 - p - q) \times f_{\lambda_p}(j)$, we progress j yards, and move to state $(k + 1, i - j, l + j)$. ($j = 0, 1, \dots, i - 1$)
 - With probability $(1 - p - q) \times [1 - F_{\lambda_r}(i - 1)]$, we progress i yards or more, and move to state T.

At the end of four downs, in state $(k = 4, i, l)$, the following state transitions occur:

- If $l \geq 10$, the play is refreshed, and we move to state $(k = 0, i, 0)$.
- Otherwise, we have not made 10 yards in the fourth down, and the ball is turned over to the other team, moving to state L.

Assumption 7.2.1 holds because the football field is of finite length (120 yards), and at the end of a play, if we have not entered the termination state T, we either enter the other termination state L, or the distance to the goal i decreases by at least 10. Therefore the game terminates in at most 12 cycles of four downs.

Bellman's equation. Let $J(k, i, l)$ be the optimal cost, which is the probability of scoring a touchdown starting from this state. According to the description above, we have

- For $k = 0, 1, 2, 3$:

$$\begin{aligned}
J(k, i, l) = \max\{ & \text{(run)} \sum_{j=0}^{i-1} f_{\lambda_r}(j) J(k + 1, i - j, l + j) + [1 - F_{\lambda_r}(i - 1)] J(T), \\
& \text{(pass)} pJ(k + 1, i, l) + qJ(L) \\
& + (1 - p - q) \left[\sum_{j=0}^{i-1} f_{\lambda_p}(j) J(k + 1, i - j, l + j) + [1 - F_{\lambda_p}(i - 1)] J(T) \right] \}
\end{aligned}$$

- For $k = 4$:

$$J(k, i, l) = \begin{cases} J(0, i, 0) & \text{if } l \geq 10, \\ 0 & \text{otherwise.} \end{cases}$$

- $J(T) = 1$.
- $J(L) = 0$.

(b) I implemented the above value iteration process in Python, with the code attached. The value and policy converged in 60 iterations, taking less than a minute to run.

The resulting policies for each state can be summarized as follows:

- In the beginning of a play ($k = 0$), pass if the distance to the goal is 20 yards or more; run otherwise.
- After the first down ($k = 1$),
 - If 2 yards or less were gained in the first down, pass if distance to goal is 8 yards or more; run otherwise.
 - If 4 yards were gained in the first down, pass if distance to goal is 17 yards or more; run otherwise.
 - If 4 or more yards were gained in the first down, always run.
- After the second down ($k = 2$),
 - Pass if 5 yards or less have been gained in the previous downs and the goal is more than 5 yards away.
 - Otherwise, run.
- After the third down ($k = 3$): The policy is attached. Note that for $l \geq 7$, the policy is no longer a simple threshold policy between pass and run. This is possibly because the expected value of “scoring in the current play” vs “trying to make 10 yards so as to get a new play” are comparable.

MIT OpenCourseWare
<http://ocw.mit.edu>

6.231 Dynamic Programming and Stochastic Control
Fall 2015

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.