Course Notes On Dynamic Optimization (Fall 2023)
Lecture 5: Continuous-Time Discrete-Event Problems

Instructor: Daniel Russo

Email: djr2174@gsb.columbia.edu

Graduate Instructor: David Cheikhi
Email: d.cheikhi@columbia.edu

These notes are based of scribed notes from a previous edition of the class. I have done some follow up light editing, but there may be typos or errors.

Topics:

- Continuous time discounted problems with discrete events.
- For exponential inter-event times, a reduction to discounted discrete time problems via uniformization
- Near-reduction for general semi-Markov problems.

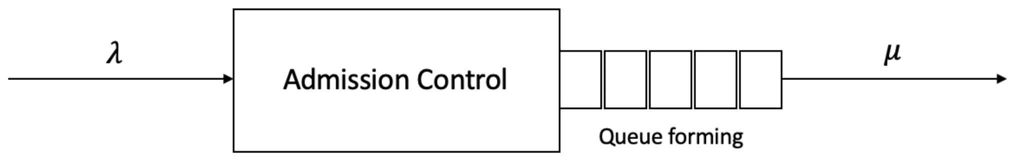
1 Discounted Continuous Time Problems

Some of the techniques we have seen so far also apply to continuous time problems, where decisions are still made a discrete but possible random time instances. The results go through for general finite state problems. For countable state problems, we need some regularity conditions ensuring that the time between decision epochs does not become infinitesimally small.

1.1 Prototypical Example: Queuing Admission Control Problem

We begin with a motivating example. Consider a queuing admission control problem with the following characteristics:

1. Single server queue.
2. Poisson arrivals rate with parameter λ.
3. Service times follow an exponential with parameter μ.
4. Control decision (binary): admit or reject.



While this problem is formulated in continuous time, decisions are made at the discrete (random) times when there is a change in the system state. There are two types of costs:

1. cost c per customer in queue per time instant.
2. cost D per job dropped (cost for turning people away).

Problem Formulation

1. Decisions are made at discrete (random) times $t_1 < t_2 < ... < t_k < ....$
   - It will also be useful to introduce the following notation: $\tau_k = t_k - t_{k-1}$ for $k = 1, 2, ...$
2. At time t_k, the decision maker observes the state $i(t_k)$ and selects a (feasible) control $u(t_k) \in U(i_{t_k})$ applied throughout $[t_k, t_{k+1})$.
3. $\tau_{k+1} | i(t_k) = i, u(t_k) = u \sim \exp(\nu_i(u))$, where $\nu_i(u) \leqslant \nu \forall i, u$.
   - We can consider ν to be a uniform control on rate at which events occur, for example we can consider $\nu = \max_{i,u} \nu_i(u)$.
4. The next state is j with probability $P_{ij}(U)$, independent of τ_k.

Objective

Minimize th expected cost

$$\limsup_{T\to\infty} \mathbb{E}^{\pi}\Big[\int_{t=0}^{T} e^{-\beta t} g(i(t), u(t)) dt\Big] = \limsup_{N\to\infty} \mathbb{E}^{\pi}\Big[\int_{t=0}^{t_N} e^{-\beta t} g(i(t), u(t)) dt\Big]$$
$$= \limsup_{N\to\infty} \sum_{k=0}^{N} \mathbb{E}^{\pi}\Big[\int_{t=t_k}^{t_{k+1}} e^{-\beta t} g(i_k, u_k) dt\Big]$$

where $i_k = i(t)$ and $u_k = u(t)$ for $t \in [t_k, t_{k+1})$. Here π is an admissible policy, i.e. $u(t_k) \in U(i(t_k))$. The expectation is taken with over the state transitions as well as the random transition times.

1.2 Warm-up Case: $\nu_i(u) = \nu \ \forall i, u$

In the simple case when the time between decision epochs is distributed i.i.d. the objective becomes:

$$\limsup_{N\to\infty} \sum_{k=0}^{N} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} e^{-\beta t} dt\Big] \mathbb{E}[g(i_k, u_k)] = \limsup_{N\to\infty} \frac{1}{\beta+\nu} \sum_{k=0}^{N} \alpha^k \mathbb{E}[g(i_k, u_k)]$$

where $\alpha = \mathbb{E}[e^{-\beta\tau}] = \int_0^\infty e^{-\beta t}\nu e^{-\nu t}dt = \frac{\nu}{\beta+\nu}$. This follows from the calculation

$$\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} e^{-\beta t}dt\Big] = \mathbb{E}\Big[e^{-\beta t_k}\int_0^{\tau_{k+1}} e^{-\beta t}dt\Big] = \mathbb{E}[e^{-(\beta\tau_1+\ldots+\beta\tau_k)}]\frac{1}{\beta}(1-\mathbb{E}[e^{-\beta\tau}]) = \prod_{i=1}^{k}\mathbb{E}[e^{-\beta\tau_i}] = \alpha^k,$$

where we use the fact that the random variables $\{\tau_k\}$ are i.i.d with rate ν. Note how this objective mimics the infinite horizon discounted objective which we are now familiar with.

2 Non-Uniform Transition Rates: $\nu_i(U) \leqslant \nu$

The main idea for the non-uniform case is to reduce to the uniform case by introducing fictitious events (where the event occurs but nothing happens, can be viewed as a transition to self-state). This can be viewed as an analogue of thinning, a commonly used technique in the study of Poisson processes. Like before, we assume the inter-event times to be i.i.d $\tau_1, \ldots, \tau_k \sim \exp(\nu)$, but to account for non-uniform transition rates we introduce some "fictitious" events where the system state does not change (that is the system transitions to the same state).

New Transition Probabilities under Uniform Version

$$\widetilde{P_{ij}}(u) = \begin{cases} \frac{\nu_i(u)}{\nu}P_{ij}(u) & \text{if } i \neq j \\ \frac{\nu_i(u)}{\nu}P_{ii}(u) + \frac{\nu-\nu_i(u)}{\nu} & \text{if } i = j \end{cases}$$

The interpretation is that we have added a fictitious events that occur at rate $\frac{\nu-\nu_i(u)}{\nu}$ A fraction $\nu_i(u)/\nu$ of events of events are "genuine" and drawn according to the probabilities $p_{ij}(u)$.

Remark: by adding fictitious events, we add new periods at the control can be re-chosen. But since the system state does not actually change the original control would still be optimal. This last fact uses crucially that the time between events is exponentially distributed (and hence memory-less).

Bellman Operator

$$TJ(i) = \min_{u\in U(i)} \widetilde{g}(i,u) + \alpha\sum_j \widetilde{P_{ij}}(u)J(j)$$

where $\alpha = \frac{\nu}{\beta+\nu}$, $\widetilde{g}(i,u) = \frac{g(i,u)}{\beta+\nu}$. Again, we are essentially back in the infinite horizon discounted regime with modified cost and transition probabilities. Analyzing both the above cases heavily relied on the memoryless property given that the inter-event times were exponentially distributed, essentially $\{\tau_k\}$ were not dependent on the history. We relax this below.

3 General Semi-Markov Problems

We now consider the more general class of Semi-Markov problems, in which the time between events may not be exponentially distributed. We will see that again, most of the core theory we have developed for discrete time discounted dynamic programming applies to these problems. Now, conditioned on the current sate i and control u, the next state and the time for a transition to occur are drawn jointly from some distribution Qiu:

$$P(\tau_k \leqslant \tau, i_{k+1} = j | i_k = i, u_k = u, \text{History up to } k) = Q_{iu}(\tau, j)$$

In addition, we assume that, conditioned on i_k and u_k, the next outcome (τ_k, i_{k+1}) is independent of the history $\{(t_\ell, u_\ell, i_\ell)\}_{\ell<k}$ prior to time k.

For such problems, we define a new Bellman operator as:

$$TJ(i) = \min_{u \in U(i)} \mathbb{E}\Big[\int_{t_0}^{t_1} e^{-\beta t} g(i(t), u(t)) dt + e^{-\beta t_1} J(i(t_1)) \Big| i(t_0) = i, u(t_0) = u \Big]$$

where $t_0 = 0$ and the expectation integrates over the event time t_1 and the next state $i(t_1)$. Applying the tower property of conditional expectations, we can rewrite this as: Now:

$$TJ(i) = \min_u \widetilde{g}(i,u) + \sum_j P_{ij}(u) \alpha_{ij}(u) J(j)$$

where $\alpha_{ij}(u) = \mathbb{E}[e^{-\beta t_1} | i(t_1) = j, i(t_0) = i, u(t_0) = u]$ acts as a discount factor, down-weighting how much future events matter. The function $\widetilde{g}(i,u) = \mathbb{E}[\int_{t_0}^{t_1} e^{-\beta t} g(i(t), u(t)) dt]$ captures the expected discounted flow costs incurred by applying control u in state i until the next event occurs

While this Bellman operator differs from the discounted case, it is still a monotone contraction operator obeying a variant of the constant-shift property. As a result, our main theoretical results extend to Semi-Markov problems. Precisely, suppose the effective discount factors are uniformly bounded above as $\alpha_{ij}(u) \leqslant \bar{\alpha} < 1 \forall i, j, u$. This is always true if the state space and action space is finite, assuming that $\mathbb{P}(\tau_k = 0) = 0$. Then T satisfies the following properties.

- Monotonicity: $J \preceq J' \implies TJ \preceq TJ'$.
- Strictly sub-constant shift: For $c \geqslant 0$, $(TJ + ce) \preceq TJ + \bar{\alpha} ce$ and $(TJ - ce) \succeq TJ - \bar{\alpha} ce$.
- Contraction: $\|TJ - TJ'\|_\infty \leqslant \bar{\alpha} \|TJ - TJ'\|_\infty$.

The contraction property follows from the other two, which can be seen by repeating the proof we gave in the discounted case