Course Notes On Dynamic Optimization (Fall 2023)
Lecture 2B: Inventory Control

Instructor: Daniel Russo

Email: djr2174@gsb.columbia.edu

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

September, 2023

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.

See also Bertsekas, *Dynamic Programming and Optimal Control* Vol. 1 Section 3.1

Here, we consider the special case of linear system when the cost is quadratic. We have,

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k && k = 0, 1, \ldots, N-1 \\ g(x_k, u_k) &= x_k^\top Q x_k + u^\top R u && k = 0, 1, \ldots, N-1 \end{aligned}$$

Where $x_k \in \mathbb{R}^n, u_k \in \mathbb{R}^m$, $Q \in \mathbb{S}^{n\times n}, Q \succeq 0, R \in \mathbb{S}^{m\times m}, R \succ 0$, $\mathbb{E}(w_k) = 0$ $k = 0, 1, \ldots, N-1$ and w_k's have finite second moments. We assume w_k to be iid random vectors (In fact, we only need independence. Identical distributions are assumed to simplify notation. Similarly, A, B, R, Q can all depend on k).

Main results: Nearly all dynamic programming problems are intractable when the state space consists of continuous vectors of moderate dimensions. The very special structures in linear quadratic control allow us to avoid the curse of dimensionality. We summarize the results below:

1. The optimal cost-to-go function is quadratic:

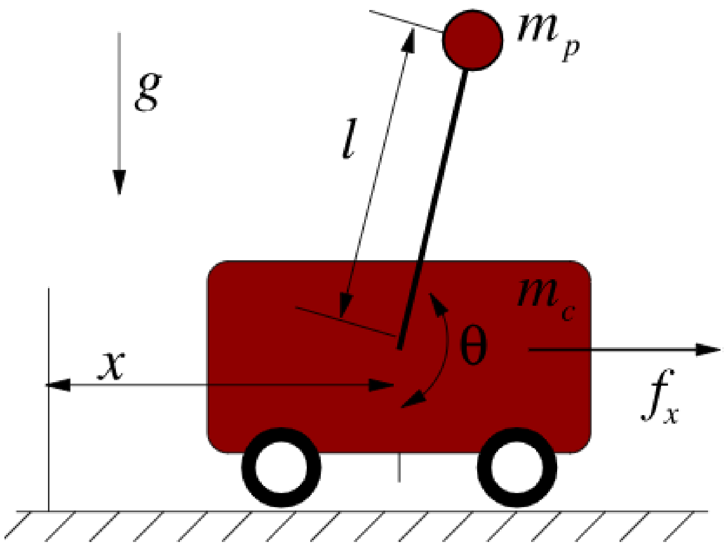
   $$J_i^*(x) = x^\top K_i x + c_i$$

   where $K_i \succ 0$ is a symmetric positive definite matrix. The $K_i's$ can be computed by recursive linear algebra as $K_i = A^\top(K_{i+1} - K_{i+1}B(B^\top K_{i+1}B + R)^{-1}B^\top K_{i+1})A + Q$.

2. The optimal policy $\pi^* = \left(\mu_0^*, \ldots, \mu_{N-1}^*\right)$ takes the form $\mu_i^*(x) = L_i x$, where $L_i = -(R + B^\top K_i B)^{-1} B^\top K_i A$.

3. The optimal policy has no dependence on the distribution of the disturbances. The optimal policy is the same as it would be in a system with no noise at all, a property known as *certainty equivalence*.

1 A glimpse of how LQR is used



Consider the cart-pole problem depicted in this figure. The goal is to keep a pendulum balanced upright by moving the cart horizontally.

Let z be the horizontal position of the cart (depicted by x in the figure), θ be the angle of the pendulum and take $x = [z, \theta, \dot{z}, \dot{\theta}]^\top$ to be the state variable. The variable u is the energy exerted along the horizontal access. It is possible to derive a nonlinear differential equation

$$x'(\tau) = f(x(\tau), u(\tau)) \qquad \tau \in [0, \infty)$$

that describes the laws of motion.

To tackle this problem we linearize the dynamics around unstable equilibrium at $\theta = 0$ (i.e. vertical cart pole). More generally, let us posit that there is an equilibrium point (x^*, u^*) such that $f(x^*, u^*) = 0$. Picking the Jacobian matrices $A = \frac{\partial}{\partial x} f(x^*, u^*)$ and $B = \frac{\partial}{\partial u} f(x^*, u^*)$, we approximate the nonlinear dynamics near the equilibrium point through the linear dynamics

$$x'(\tau) = \tilde{A}(x(\tau) - x^*) + B(u(\tau) - u^*).$$

To fit the discrete time formultaion in these notes, we define the variables

$$(x_0, u_0) = (x(0) - x^*, u(0) - u^*), \quad (x_1, u_1) = (x(\epsilon) - x^*, u(\epsilon) - u^(0)) \ldots$$

which, upto $o(\epsilon^2)$ error, obey the discrete time linear system

$$x_{k+1} = \underbrace{[I + \tilde{A}]}_{:=A} x_k + B u_k.$$

2 Derivation of the main results.

The derivation involves lots of algebra and is tedious. Refer to the textbook for a complete derivation.

Our aim is to prove that the given policy is optimal by the principle of mathematical induction. We prove this in 3 steps.

Step 1: $h(x, u)$ is convex quadratic in (x, u).

Why? This is because

- g is convex quadratic
- $J_N^*(x)$ is convex quadratic
- $(x, u, w) \to f(x, u, w)$ is an affine function, and composition of a convex quadratic function with an affine function is convex quadratic as well.

To see this more explicitly, note that the function $x \to J_N^*(x)$ is a convex function. $(x, u, w) \to f(x, u, w)$ is an affine function. Thus, $\forall w\ (x, u) \to J_N^*(f(x, u, w))$ is a convex function, and taking expectation of this w.r.t w preserves convexity. Since, $g(x, u)$ is also convex, $h(x, u)$ is just a sum of two convex functions, and hence is convex in turn.

To see that it is also quadratic, we write its full expansion. We have,

$$\begin{aligned} h(x,u) &= u^\top Ru + x^\top Qx + \mathbb{E}\left((Ax+Bu+w)^\top K_N\,(Ax+Bu+w)\right) + c_N \\ &= u^\top(R + B^\top K_N B)u + x^\top(Q + A^\top K_N A)x + 2x^\top A^\top K_N Bu + \mathbb{E}\left(w^\top Qw\right) + c_N \end{aligned}$$

which, clearly, is quadratic in (x, u).

Step 2: The minimizer $x \mapsto \arg\min_{u\in\mathbb{R}^m} h(x, u)$ is a linear function of state.

To see this, we apply the first order conditions for minimality. At the point of minimality, the first derivative $\nabla_u h(x, u)$ should vanish. We have,

$$\begin{aligned} \nabla_u h(x,u) &= 2\left(R + B^\top K_N B\right)u + 2B^\top K_N Ax \\ \nabla_u h(x,u) = 0 &\implies u = -(R + B^\top K_N B)^{-1}B^\top K_N Ax \\ \therefore \mu^*_{N-1}(x) &= L^*_{N-1}x \qquad \text{where } L^*_{N-1} = -(R + B^\top K_N B)^{-1}B^\top K_N A \end{aligned}$$

Note here, that indeed, L^*_{N-1} doesn't depend on the distribution of w_k's.

Step 3: Induction step

We first note that $J^*_{N-1}(x) = \min_u h(x, u) = h(x, L^*_{N-1}x)$ is a composition of a convex quadratic h with a linear function $x \to (x, L^*_{N-1}x)$, thus $J^*_{N-1}(x)$ is also convex quadratic. We can write

$$J^*_{N-1}(x) = \min_u h(x,u) = h(x, L^*_{N-1}x) = x^\top K_{N-1}x + c_{N-1}$$

$$\begin{aligned} \text{Where}\quad K_{N-1} &= {L^*_{N-1}}^\top(R + B^\top K_N B)L^*_{N-1} + Q + A^\top K_N A + 2A^\top K_N B L^*_{N-1} \\ &= A^\top K_N B(R + B^\top K_N B)^{-1}B^\top K_N A - 2A^\top K_N B(R + B^\top K_N B)^{-1}B^\top K_N A + Q + A^\top K_N A \\ &= A^\top(K_N - K_N B(B^\top K_N B + R)^{-1}B^\top K_N)A + Q \\ c_{N-1} &= c_N + \mathbb{E}\left(w^\top Qw\right) \end{aligned}$$

The above recurrence relation is called the Riccati equation. We note here, that K_{N-1} is symmetric whenever K_N, Q and R are symmetric. It can also be verified that K_{N-1} is positive semidefinite whenever K_{N-1}, R and Q are positive semidefinite. These are all the ingredients required to take the induction further and prove that $\mu^*_{N-2}(x)$ is also linear in x with the same arguments as above. Thus, with the principle of mathematical induction, we have managed to prove that the optimal policy $\pi^* = (\mu^*_0, \ldots, \mu^*_{N-1})$ takes the form $\mu^*_k(x) = L_k x$ as required.