Lecture 18: Linear quadratic regulator

Jie Fu
jfu2@wpi.edu
RBE 502

16, 2018
April ~~11, 2017~~

Outline

- continuous-time LQR problem
- dynamic programming solution
- infinite horizon LQR
- discrete-time LQR and nonlinear LQR.

The slide is partially adapted from Lecture notes by Prof. Stephen Boyd [1].

[1] http://stanford.edu/class/ee363/lectures/clqr.pdf

Introduction

Consider the linear time invariant (LTI) system

$$\dot{x} = Ax + Bu$$

Control objective: bring the system from an arbitrary initial state to the origin.

rank $[B \ AB \ A^2B \cdots A^{k-1}B]$ full rank.

- A LTI system, if stabilizable, can always be stabilized by a linear feedback law $u = Kx$.
- If it is controllable, the close-loop poles can be placed ***anywhere*** on the S-plane. Thus the convergence can be made arbitrarily fast, which means ***large input amplitudes***.

E.g. $\dot{x} = x + u$, choosing $u = -kx$ and $k = 1.1, 1000, 10000$.

$= (1-k)x$

$1-k < 0$

$k > 1$

0.1

$x(t) = e^{-0.1t}x(0)$

Formulating an optimization problem

Considerations:

- The bounded input amplitudes.
- The speed of convergence.

Natually an optimization problem, let's see how it is formulated.
Criterion 1: Fast convergence.

$$\int_0^T x^T(t)Qx(t)dt$$

where Q is a ***nonnegative-definite symmetric*** matrix and T is the ***final time***.
- x^TQx : a measure of how much x deviates from the origin.
- Q: weighting matrix.
- the integral is the cumulative deviation of x from the zero during the interval $[0, T]$.

e.g. 1-d system. $Q=1$. $X(t) = x(t)$ $\int_0^T x^2(t)dt$

Formulating an optimization problem

Criterion 2: Bound the input amplitudes.

$$\int_0^T u^T(t)Ru(t)dt$$

where R is a nonnegative definite symmetric matrix.

- $u^T Ru$: a measure of how much u deviates from the zero.
- To reduce the amplitudes we have the total value as small as possible.

Criterion 3: Converge to the origin. the terminal cost, if $x(T) \neq 0$.

$$x(T)^T Q_f x(T)$$

Remark: Weighting matrices Q, R, Q_f may need to be selected via trials and errors.

Formulating an optimization problem

stable controller

problem: Choosing $u : [0, T] \to \mathbf{R}^m$ to minimize the total cost

$$J = \int_0^T \left(x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau) \right) \, d\tau + x(T)^T Q_f x(T)$$

- T is time horizon.
- $Q = Q^T \geq 0$, $Q_f = Q_f^T \geq 0$, and $R = R^T > 0$ are ***state cost, final state cost, and input cost*** matrices.

...an ***infinite-dimensional problem***: (the input $u : [0, T] \to \mathbf{R}^m$ is variable)

LQR Problem: linear quadratic regulator (regulating the controller).

infinite-horizon optimal control: $J = \int_0^\infty \left(x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau) \right) d\tau$

LQR solution via DP

we'll solve LQR problem using dynamic programming
for $0 \le t \le T$ we define the **value function** $V_t : \mathbf{R}^n \to \mathbf{R}$ by

state real value

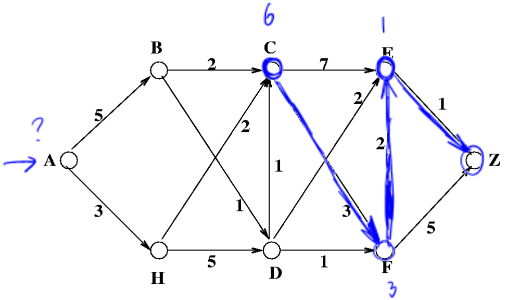
$$V_t(z) = \min_u \int_t^T x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)\, d\tau + x(T)^T Q_f x(T)$$

subject to $x(t) = z$, $\dot{x} = Ax + Bu$.

> $V_t(z) =$ *"minimal total cost for completing the task starting from state z at time t".*

- minimum is taken over all possible input signals $u : [t, T] \to \mathbf{R}^m$
- $V_t(z)$ is also called minimum LQR cost-to-go from z at time t.
- $V_T(z) = z^T Q_f z$ where $z = x(T)$ is the final state.

A simple dynamic programming example



$V(Z) = 0$

$V(E)$ = minimum cost to go from E

$= \min(\underbrace{\text{cost}(E, Z)}_{\text{step cost}} + \underbrace{V(Z)}_{\text{Next state cost to go}}) = \min(1 + 0) = 1$

$V(F) = \min(\text{cost}(F, E) + V(E), \text{cost}(F, Z) + V(Z)) = \min(2+1, 5+0) = 3$

$V(C) = \min(\text{cost}(C, E) + V(E), \text{cost}(C, F) + V(F)) = \min(7+1, 3+3) = 6$

LQR solution via DP

Let V_t be the optimal cost to go from time t.

fact

V_t is quadratic, *i.e.*, $V_t(z) = z^T P_t z$, where $P_t = P_t^T \geq 0$ (***symmetric, positive definite***).

What do we know about V_t?

① Terminal condition: $V_T(z) = z^T P_T z = z^T Q_f z$

$P_T = Q_f$

② For $t \in [0, T)$, have to compute P_t

suppose $V_t(z) = z^T P_t z$ is known,

want to known: $V_{(t+h)}(x(t+h))$

when $x(t) = z$.

Relation between $V_{(t+h)}(x(t+h))$ and $V_t(z)$

$$\underbrace{V_{t+h}(x(t+h))}_{\uparrow}$$

Next state minimal cost to go.

step cost: $\int_t^{t+h} (x(\tau)^T Q x(\tau) + u^T(\tau) R u(\tau))\, d\tau$

$$V_t(z) = \min_{w} \left(\text{step cost} + V_{t+h}(x(t+h)) \right) \quad (*)$$

suppose w is given

step cost: $\int_t^{t+h} (x(\tau)^T Q x(\tau) + u^T R u)\, d\tau = (x(t)^T Q x(t) + w^T R w) \cdot h$

$= (z^T Q z + w^T R w) h$

$$V_{t+h}(x(t+h)) = x(t+h)^T P_{t+h} x(t+h)$$

$$\dot{x} = Ax + Bu$$

$$x(t+h) = x(t) + (Ax(t) + Bu(t)) \cdot h$$

$$= z + (Az + Bw) h.$$

$$\boxed{P_{t+h} = P_t + \dot{P}_t \cdot h}$$

LQR solution via DP

We start with $x(t) = z$, let's take $u(t) = w \in \mathbf{R}^m$, a constant, over the time interval $[t, t+h]$, where $h > 0$ is small

The cost incurred over $[t, t+h]$ is and the state is

$$V_t(z) = \min_w \left((z^T Q z + w^T R w) h + (z + (Az + Bw)h)^T (P_t + \dot{P}_t h)(z + (Az + Bw)h) \right)$$

simplification: omit all h^2, h^3 terms.

$$= \min_w \Big((z^T Q z + w^T R w) h + z^T P_t z + ((Az + Bw)h)^T P_t z + z^T \dot{P}_t h z$$

$$+ \cancel{((Az + Bw)h)^T \dot{P}_t h (Az + Bw)h} + z^T P_t (Az + Bw)h$$

$$+ ((Az + Bw)h)^T \cancel{P_t (Az + Bw)h} + \cancel{((Az + Bw)h)^T} \dot{P}_t h z$$

$$+ z^T \dot{P}_t h \cancel{(Az + Bw)h} \Big)$$

$$= (z^T Q z) h + z^T \dot{P}_t h z + z^T P_t z + \min_w \left(w^T R w h + ((Az + Bw)h)^T P_t z + z^T P_t (Az + Bw) h \right)$$

Taking the derivative of the function:

$$2h w^T R + 2h^T z^T P_t B = 0$$

$$w^* = -\underbrace{R^{-1} B^T P_t}_{?}\,\underbrace{z}_{x(t)}.$$

LQR solution via DP

Recall $V_t = x(t)^T P_t x(t)$, and $P_{t+h} \approx P_t + h\dot{P}_t$.
min-cost-to-go from where we land is approximately

$$V_{t+h}(x(t+h)) \approx V_{t+h}(z + h(Az + Bw))$$

cost plus incurred cost at t is approximately

LQR solution via DP

To optimize

$$V_t(z) \approx \min_w \Big[z^T P_t z + \\ h \Big(z^T Q z + w^T R w + (Az + Bw)^T P_t z + z^T P_t (Az + Bw) + z^T \dot{P}_t z \Big) \Big]$$

minimize over w to get (approximately) optimal w:

HJB equation

$V_t(z) = \min_{w} ($ step cost + Next state cost $)$

$= -R^{-1}B^T P_t z$

now let's substitute w^* into equation.

omited h^2, h^3.

$$z^T P_t z = V_t(z) \approx z^T P_t z +$$
$$+ h \left(z^T Q z + {w^*}^T R w^* + (Az + Bw^*)^T P_t z + z^T P_t (Az + Bw^*) + z^T \dot{P}_t z \right)$$

after simplification,

$$-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t + Q$$

$-\frac{P(t+h) - P(t)}{h} = (\qquad)$

which is the ***Riccati differential equation*** for the LQR problem
we can solve it (numerically) using the ***final condition*** $P_T = Q_f$.

Summary of LQR for LTI system via DP

1. solve Riccati differential equation

$$-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t + Q, \qquad P_T = Q_f$$

(backward in time) $u = -R^{-1} B^T P_t \, x(t)$

2. optimal u is $u_{\mathrm{lqr}}(t) = K_t x(t)$, $K_t := -R^{-1} B^T P_t$

DP method readily extends to time-varying A, B, Q, R, and tracking problem (define $e(t) = x(t) - x_d(t)$ and drive the tracking error to 0.)

infinite horizon: $0 = A^T P + PA - PBR^{-1}B^T P + Q$

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}$$

Solve $P(t)$ backward in time

1. select a time step δt and total number of steps $N = \frac{T}{\delta t}$. By final condition $P_N = P(T) = Q_f$.
2. $-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t + Q$ in numerical computation via Eular approximation of $\dot{P}_t$.

Steady-state regulator

$$-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t + Q, \qquad P_T = Q_f$$

Usually P_t rapidly converges as t decreases below T.
by convergence, $\dot{P}_t = 0, P_t = P_{t^+} = P$ as the limit of P_t which satisfies a

quadratic matrix equation — algebraic Riccati equation

- P can be found by direct method or numerically integrating the Riccati differential equation. In matlab, $[X, L, G] = care(A, B, Q, R, S, E)$ computes the unique solution X of the continuous-time algebraic Riccati equation.
- for t not close to horizon T, LQR optimal input is approximately a linear, constant state feedback

$$u(t) = K_{\text{ss}} x(t), \qquad K_{\text{ss}} = -R^{-1} B^T P_{\text{ss}}$$

Variation: Infinite horizon

The infinite horizon cost function

$$J = \int_0^\infty x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)\, d\tau$$

The value function as

$$V(z) = \min_u \int_0^\infty x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)\, d\tau$$

subject to $x(0) = z$, $\dot{x} = Ax + Bu$

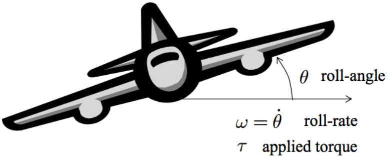
- we assume that (A, B) is controllable, so V is finite for all z
- V is quadratic and time-invariant: $V(z) = z^T P z$, where $P = P^T \geq 0$
- the optimal u is $u(t) = Kx(t)$, where $K = -R^{-1} B^T P$ (*i.e.*, a constant linear state feedback) $\dot{P} = 0$

For infinite horizon, in matlab, $\mathrm{K} = \mathrm{lqr}(\mathrm{A}, \mathrm{B}, \mathrm{Q}, \mathrm{R})$.

Remarks

- For finite-horizon $[0, T]$ optimal control, need to solve $P(t)$ backward in time and apply $u(t) = -R^{-1}B^T P(t)x$ which is a ***time varying*** feedback control.
- For infinite-horizon $[0, \infty)$ optimal control, P is a constant and u is a ***time-invariant*** control input.
- Require $R > 0$, $Q \geq 0$, and there exists $Q = H^T H$.

Example: Aircraft roll-dynamics [2]



$$\dot{\theta} = \omega$$
$$\dot{\omega} = -.875\omega - 20\tau$$
$$\dot{\tau} = -50\tau + 50u$$

$$x = \begin{bmatrix} \theta & \omega & \tau \end{bmatrix}'$$

with

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -.875 & -20 \\ 0 & 0 & -50 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 50 \end{bmatrix}.$$

Consider

$$J := \int_0^\infty \|x\|^2 + \rho \|u\|^2 \, dt,$$

where ρ is a positive constant, corresponding to $Q = \mathrm{diag}(1,1); R = \rho$.

[2] Hespanha, J. "Lecture notes on lqr/lqg controller design." notes, online note (2005).

Example

The control is $u := -Kx, \quad K := R^{-1}B^T P$ where P solves the algebraic Riccati equation

$$A^T P + PA - PBR^{-1}B^T P + Q = 0$$

Discrete-time LQR

$$x_{t+1} = Ax_t + Bu_t$$

and objective function

$$J = \sum_{\tau=0}^{N-1} (x_\tau^T Q x_\tau + u_\tau^T R u_\tau) + x_N^T Q_f x_N$$

For $t = 0, \ldots, N$ define the value function

$$V_t(z) = \min_{u_t, \ldots, u_{N-1}} \sum_{\tau=t}^{N-1} (x_\tau^T Q x_\tau + u_\tau^T R u_\tau) + x_N^T Q_f x_N$$

subject to $x(t) = z$ and the dynamics.

- $V_t(z)$ is the minimum LQR cost-to-go, starting from z at time t;
- $V_0(x_0)$ is the min cost (from state x_0 at time 0.)

Discrete-time LQR

- V_t is quadratic, *i.e.*, $V_t(z) = z^T P_t z$ where $P_t = P_t^T \geq 0$.
- $P_T = Q_f$ and can be computed backward in time and

$$P_t = Q + A^T P_{t+1} A - A^T P_{t+1} B (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$$

Discrete-time LQR

$$V_t = z^T Qz + \min_u (u^T Ru + V_{t+1}(z_{t+1}))$$

since $z_{t+1} = Az + Bu$,

$$V_t = z^T Qz + \min_u (u^T Ru + V_{t+1}(Az + Bu))$$

is the dynamic programming, Bellman or Hamilton-Jacobi equation.

Discrete-time LQR

$$V_t = z^T Qz + \min_u (u^T Ru + V_{t+1}(z_{t+1}))$$

since $z_{t+1} = Az + Bu$,

$$V_t = z^T Qz + \min_u (u^T Ru + V_{t+1}(Az + Bu))$$

is the dynamic programming, Bellman or Hamilton-Jacobi equation.
minimize u gives optimal input at time t

$$u_t^{\mathrm{lqr}} = \arg\min_u (u^T Ru + V_{t+1}(Az + Bu))$$

take derivative and since $V_{t+1}(z) = z^T P_{t+1} z$

$$\frac{\partial(u^T Ru + (Az + Bu)^T P_{t+1}(Az + Bu))}{\partial u} = 0$$

$$u^* = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} Az$$

put in V_t obtain the relation between P_t and P_{t+1}.

Conclusion

- continuous and discrete LQR for time invariant systems.
- extension to for time varying systems is straightforward.

Further readings ...

- the dual problem: linear quadratic estimator and Kalman filter.
- linear quadratic gaussian (LQG) control.