Lecture 4: Linear Quadratic Regulators

Want to do control for a broader class of systems:

1. Want LTV since trajectory following is important.
2. Hard to choose closed loop roots. Too negative, then large control inputs. Too close to zero, then slow convergence.

We will use an optimization-based approach, but first

1. When is the problem well-posed.
2. How to solve.

A. Positive Semidefinite Matrices

Def:

a) A square matrix is called <u>symmetric</u> if it is equal to its transpose.

b) A symmetric matrix A is called <u>positive definite</u> if $x^T Ax > 0$ for any $x \neq 0$ denoted $A > 0$.

c) A symmetric matrix is called <u>positive semidefinite</u> if $x^T Ax \geq 0$, denoted $A \geq 0$.

Examples:

a) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then $x^T Ax = x_1^2 + x_2^2 > 0 \ \forall x \neq 0$, so $A > 0$.

b) Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then $x^T Ax = (x_1 + x_2)^2 \geq 0 \ \forall x$, so $A \geq 0$.

Remark: We use quadratic polynomials (which have bounded minimum value) to show that a matrix is positive definite or positive semidefinite, but if we show that a polynomial can be written as a positive definite or positive semidefinite matrix does it have a bounded minimum value.

Theorem. *A quadratic function*

$$f(x) = x^T Dx + C^T x + C_0$$

where $D \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^n$, $C_0 \in \mathbb{R}$, and $x \in \mathbb{R}^n$ has one bounded minimum value iff $D \geq 0$. Furthermore, the minimum value is achieved for a unique value of x iff $D > 0$.

But how to check if D is positive semidefinite or positive definite?

Theorem. *A matrix is positive semidefinite iff all of its eigenvalues are real and non-negative. Furthermore, A matrix is positive definite iff all of its eigenvalues are real and positive.*

LQR Formulation and Solution

$$\begin{aligned} &\min_{u:[0,T]\to\mathbb{R}^m} \int_0^T \left(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t) \right) dt + x^T(T)Q(T)x(T) \\ &\text{subject to: } \ \dot{x}(t) = A(t)x(t) + B(t)u(t) \end{aligned}$$

Remarks:

1) T is the time horizon.

2) To ensure problem is well formulated we will require $Q(t) = Q^T(t) \geq 0 \; \forall t \in [0, T]$ and $R(t) = R^T(t) > 0 \; \forall t \in [0, T]$.

3) R is called the input cost. $Q(t)$ is called the state cost $\forall t \in [0, T)$. $Q(T)$ is called the final cost.

4) This is an infinite dimensional problem since u is defined for all times.

5) This optimization tries to drive an LTV system to zero, can apply to tracking problems by applying a transformation that places the trajectory at the origin.

6) Note that no initial condition is specified.

Theorem. *The optimal input $u^* : [0, T] \to \mathbb{R}^m$ to LQR is a state feedback controller:*

$$u^*(t) = -K(t)x(t) \; \forall t \in [0, T]$$

where

$$K(t) = R^{-1}(t)B^T(t)P(t) \; \forall t \in [0, T]$$

where $P(t)$ is the solution to the Ricatti Differential Equation:

$$-\dot{P}(t) = A^T(t)P(t) + P(t)A(t) - P(t)B(t)R^{-1}(t)B^T(t)P(t) + Q(t)$$

with final condition P(T)=Q(T)

Remarks:

1) To apply this we can just solve for $P(t)$ starting from $T \to 0$, then compute $K(t)$.

2) In practice have to solve a nonlinear differential equation (use Euler).

3) Solution does not depend on the initial condition so it works everywhere.

Proof. We solve this problem using Bellman's optimality principal be defining the value function $v : \mathbb{R}^n \times [0, T] \to \mathbb{R}$ by

$$\begin{aligned} v(z, t) = \min_{u:[t,T]\to\mathbb{R}^m} & \int_t^T \Big(x^T(\tau)Qx(\tau) + u^T(\tau)Ru(\tau)\Big)d\tau + x^T(T)Q_T x(T) \\ \text{subject to: } & x(t) = z \\ & \dot{x} = Ax + Bu \end{aligned}$$

(Note that we have dropped the time dependence for convenience but the proof works with the time dependence included.) $v(z, t)$ gives the minimum LQR cost to go from state z and time t. If we select $t = T$ then $v(z, T) = z^T Q_T z$.

Lemma. *For LQR $v(x, t) = x^T(t)P(t)x(t)$ where $P(t) \geq 0$*

Proof: Anderson & Moore

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision"

-Bellman, 1957 "Dynamic Programming"

Lets see how to apply this idea: Suppose we start with $x(t) = z$ and lets assume $u(t) = w \in \mathbb{R}^m$ a constant over the time interval $[t, t + h]$ for a small h. The incurred cost over $[t, t + h]$ is

$$\int_t^{t+h} \Big(x^T(\tau)Qx(\tau) + w^T Rw\Big)d\tau \approx h\Big(z^T Qz + w^T Rw\Big)$$

and we end up at $x(t+h) \approx z + h(Az + Bw)$ The min cost-to-go from where we land is:

$$\begin{aligned} v(z+h(Az+Bw),t+h) &= (z+h(Az+Bw))^T P(t+h)(z+h(Az+Bw)) \\ &\approx (z+h(Az+Bw))^T (P(t)+h\dot{P}(t))(z+h(Az+Bw)) \\ &\approx z^T P(t)z + h\Big((Az+Bw)^T P(t)z + z^t P(t)(Az+Bw) + z^t \dot{P}(t)z\Big) \end{aligned}$$

According to Bellman's Principal of optimality:

$$\begin{aligned} v(z,t) &= \int_t^{t+h} (x^T(\tau)Qx(\tau) + w^T Rw)d\tau + v(z+h(Az+Bw),t+h) \\ &\approx z^T P(t)z + h\Big(z^T Qz + w^T Rw + (Az+Bw)^T P(t)z + z^T P(t)(Az+Bw) + z^T \dot{P}(t)z\Big) \end{aligned}$$

If we minimize over w (take $\frac{d}{dw}v(z,t)$ and set equal to 0 and solve):

$$\begin{aligned} 2hw^T R + 2hz^T P(t)B &= 0 \\ \text{so: } w^*(t) &= -R^{-1}B^T P(t)z \\ \text{therefore: } u^*(t) &= -K(t)x(t) \end{aligned}$$

Now to get an equation for $P(t)$, recall: $v(z,t) = z^T P(t)z$ but it is also:

$$z^T P(t)z = v(z,t) = z^T P(t)z + h\Big(z^T Qz + w^{*T}(t)Rw^*(t) + (Az+Bw^*(t))^T P(t)z + z^T P(t)(Az+Bw^*(t)) + z^T \dot{P}(t)z\Big)$$

After simplification:

$$-\dot{P}(t) = A^T P(t) + P(t)A - P(t)BR^{-1}B^T P(t) + Q$$

with $P(T) = Q_T$ ∎