Notes: CDS 110b,Winter 2014, Caltech

1 Discrete time linear quadratic control (LQR)

1.1 Optimal control, an introduction

Consider the discrete time dynamics

$$x_{t+1} = f(x_t, u_t, t)$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^p$ is the input with $u_t = g(y_t, t)$. The goal is to "control optimally". For the inverted pendulum, we might want to optimize and trade off between the following objectives:

1. minimize deviation from vertical
2. deviation from position
3. speed of hand movement

i.e. penalize the state *and* the control effort. We encode these objectives in a cost function $J(x, u), x = (x_0, ..., x_T), u = (u_0, ..., u_T), T \in \{1, 2, ...\}$. The optimal control problem can then be formulated in the following way

$$\begin{aligned} &\min_u J(x, u) \\ &\text{s.t. } x_{t+1} = f(x_t, u_t, t), u_t = g(y_t, t). \end{aligned} \tag{1}$$

We can not always solve for this as it is much too general, there are not enough assumptions for J or f. Generally speaking, finding the *global* minimum of a function $J(x, u)$, even without constraints on how x evolves, is an analytically and computationally intractable problem. In general, a desirable property of a function to be minimized is that it be *convex*: in this case, local minima (which can be found relatively easily via gradient descent methods, for example) are indeed global minima as well.

Given the interpretation of our cost function as a penalty on state and control effort, and our desire for it to be convex, a broad class of useful cost functions norms, or functions that "behave like" norms. We restrict our exposition to linear dynamics, because convexity is preserved under composition with affine functions: this is a key property of linear dynamics that will be needed for reasons that will become apparent.

1.2 Discrete time LQR formulation

We begin with a special case, and look at discrete time, finite horizon linear quadratic regulators (LQR) (ref stephen boyd, ee363, lecture 1). Consider the following dynamics:

$$\begin{aligned} & x_{t+1} = Ax_t + Bu_t \\ & x_0 = x(0) \\ & u_t = g(x_t, t), \text{(i.e., state feedback)} \end{aligned} \tag{2}$$

Given the cost function,

$$J_{LQR}(x, u) = \sum_{t=0}^{T-1} x_t^\top Q x_t + u_t^\top R u_t + x_T^\top Q_T x_T,$$

with matrices $Q \geq 0, Q_T \geq 0, R > 0$, corresponding to the state cost matrix, the terminal cost matrix and the control cost matrix respectively, the control problem can be formulated as the optimization problem:

$$\begin{aligned} & \min\ _{u_0, \ldots u_{T-1}} J_{LQR}(x, u) \\ & \text{s.t. } x_{t+1} = Ax_t + Bu_t \\ & x_0 = x(0) \\ & u_t = g(x_t, t), \text{state feedback} \end{aligned} \tag{3}$$

$J_{LQR}(x, u)$ is not generally a norm on (x, u). $Q \geq 0 \Rightarrow \exists x \neq 0$, such that $x^T Q x = 0$. We require $R > 0$, to ensure well posedness. A non strict inequality would imply that R has a null space, and this could lead to large control efforts in this subspace not being penalized in the cost.

Interpretation:

Let $y_t = Cx_t$ be the output that we want to minimize, and let $Q = C^T C \geq 0$. Let $R = \rho I, \rho > 0$. This satisfies the requirements of a convex cost function:

$$J(x,u) = \sum_t ||y_t||_2^2 + \rho \sum_t ||u_t||_2^2 + ||y_T||_2^2$$

Thus, ρ trades off explicitly between control effort and output deviation (this has connections to Pareto optimality.)

1.3 Least squares solution

We use the principle of superposition to write x_t in terms of x_0 and $u_0, ...u_{t-1}$.

$$\begin{array}{l} x_0 = x(0) \\ x_1 = Ax_0 + Bu_0 \\ x_2 = Ax_1 + Bu_1 = A^2 x0 + ABu_0 + Bu_1 \\ ... \\ x_t = A^t x_0 + \sum_{s=0}^{t-1} A^s B u_{t-1-s} \\ ... \end{array} \tag{4}$$

or, in matrix form,

$$\underbrace{\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}}_{\text{X}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ B & 0 & 0 & 0 \\ AB & \ddots & 0 & 0 \\ \vdots & & \ddots & 0 \\ A^{T-1}B & A^{T-2}B & \cdots & B \end{bmatrix}}_{\text{G}} \underbrace{\begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ x_{T-1} \end{bmatrix}}_{\text{U}} + \underbrace{\begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^T \end{bmatrix}}_{\text{F}} x_0 \tag{5}$$

giving

$$X = Fx_0 + GU.$$

We rewrite $J(x,u)$ in terms of F, G, X and U and note that it is unconstrained and strongly convex in optimization variable U.

$$\begin{array}{rcl} J(x,u) & = & x_0^\top Q x_0 + x_1^\top Q x_1 + ...x_{T-1}^\top Q x_{T-1} + x_T^\top Q x_T \\ & + & u_0^\top R u_0 + u_1^\top R u_1 + ...u_{T-1}^\top R u_{T-1} \end{array} \tag{6}$$

$$J(x,u) = X^T \underbrace{\begin{bmatrix} Q & & & \\ & \ddots & & \\ & & Q & \\ & & & Q_T \end{bmatrix}}_{\tilde{Q}} X + U^T \underbrace{\begin{bmatrix} R & & \\ & \ddots & \\ & & R \end{bmatrix}}_{\tilde{R}} U \tag{7}$$

Finally, we minimize the cost function

$$\min_U (Fx_0 + GU)^\top \tilde{Q}(Fx_0 + GU) + U^\top \tilde{R} U \tag{8}$$

which can be solved by differentiation,

$$\begin{aligned} J(U) &= x_0^\top F^\top \tilde{Q} F x_0 + 2U^\top G^\top \tilde{Q} F x_0 + U^\top (\tilde{R} + G^\top \tilde{Q} G) U \\ \nabla_U J(U) &= 2G^\top \tilde{Q} F x_0 + 2(\tilde{R} + G^\top \tilde{Q} G) U = 0 \\ U &= -(\tilde{R} + G^\top \tilde{Q} G)^{-1} G^\top \tilde{Q} F x_0 \end{aligned} \tag{9}$$

The problems with this method are that it is:

1. *Computationally costly.* Naively[a], the cost is $O(T^3 np^2)$, and we can do much better than this by using a recursive solution that will reduce cost to $O(Tn^3)$.

2. *Open loop control.* This is fine under assumptions of a perfect model, perfect state measurement, and perfect actuation. In reality, we require a feedback solution to mitigate these issues.

3. *Dependent on initial condition and can not be precomputed.*

We imposed a lot of structure on our problem, a quadratic cost and linear dynamics. This structure needs to be fully exploited to get scalable solutions.

a. In the homework, you will derive a way of solving this problem in $O(Tn^3)$.