Optimal control and the linear quadratic regulator

Shankar Sastry, Forrest Laine, Claire Tomlin

February 3, 2021

These notes represent an introduction to the theory of optimal control and the linear quadratic regulator (LQR).

There exist two main approaches to optimal control:

1. via the Calculus of Variations (making use of the Maximum Principle);
2. via Dynamic Programming (making use of the Principle of Optimality).

Both approaches involve converting an optimization over a function space to a pointwise optimization. The methods are based on the following simple observations:

1. For the calculus of variations, the optimal curve should be such that neighboring curves do not lead to smaller costs. Thus the 'derivative' of the cost function about the optimal curve should be zero: one takes small variations about the candidate optimal solution and attempts to make the change in the cost zero.
2. For dynamic programming, the optimal curve remains optimal at intermediate points in time.

In these notes, both approaches are discussed for optimal control; the results are then specialized to the case of linear dynamics and quadratic cost.

The notes in Sections 1 and 2 are from Shankar Sastry's notes for the course EECS 290A [1], which provide an excellent summary of the two approaches. The notes in Section 3 are from Forrest Laine's notes for the course EE 291E (from Spring 2018).

1 Optimal Control based on the Calculus of Variations

There are numerous books on optimal control. Commonly used books which we will draw from are Athans and Falb [2], Berkovitz [4], Bryson and Ho [5], Pontryagin et al [6], Young [7], Kirk [8], Lewis [9] and Fleming and Rishel[10]. The history of optimal control is quite well rooted in antiquity, with allusion being made to Dido, the first Queen of Carthage, who when asked to take as much land as could be covered by an ox-hide, cut the ox-hide into a tiny strip and proceeded to enclose the entire area of what came to be know as Carthage in a circle of the appropriate radius[1]. The calculus of variations is really the ancient precursor to optimal control. Iso perimetric problems of the kind that gave Dido her kingdom were treated in detail by Tonelli and later by Euler. Both Euler and Lagrange laid the foundations of mechanics in a variational setting culminating in the Euler Lagrange equations. Newton used variational methods to determine the shape of a body that minimizes drag, and Bernoulli formulated his brachistochrone problem in the seventeenth century, which attracted the attention of Newton and L'Hôpital. This intellectual heritage was revived and generalized by Bellman [3] in the context of dynamic programming and by Pontryagin and his school in the so-called Pontryagin principle for optimal control ([6]).

Consider a nonlinear possibly time varying dynamical system described by

$$\dot{x} = f(x, u, t) \tag{1}$$

with state $x(t) \in \mathbb{R}^n$ and the control input $u \in \mathbb{R}^{n_i}$. Consider the problem of minimizing the performance index

$$J = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t)dt \tag{2}$$

where t_0 is the initial time, t_f the final time (free), $L(x, u, t)$ is the running cost, and $\phi(x(t_f), t_f)$ is the cost at the terminal time. The initial time t_0 is assumed to be fixed and t_f variable. Problems involving a cost only on the final and initial state are referred to as Mayer problems, those involving only the integral or running cost are called Lagrange problems and costs of the form of equation (2) are referred to as Bolza problems. We will also have a constraint on the final state given by

$$\psi(x(t_f), t_f) = 0 \tag{3}$$

where $\psi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^p$ is a smooth map. To derive necessary conditions for the optimum, we will perform the calculus of variations on the cost function of (2) subject to the constraints of equations (1), (3). To this end, define the modified cost function, using the Lagrange multipliers $\lambda \in \mathbb{R}^p, p(t) \in \mathbb{R}^n$,

$$\tilde{J} = \phi(x(t_f), t_f) + \lambda^T \psi(x(t_f), t_f) + \int_{t_0}^{t_f} \left[L(x, u, t) + p^T (f(x, u, t) - \dot{x}) \right] dt \tag{4}$$

Defining the *Hamiltonian* $H(x, u, t)$ using what is referred to as a *Legendre transformation*

$$H(x, p, u, t) = L(x, u, t) + p^T f(x, u, t) \tag{5}$$

[1]The optimal control problem here is to enclose the maximum area using a closed curve of given length.

The variation of (4) is given by assuming independent variations in $\delta u(), \delta x(), \delta p(), \delta\lambda$, and δt_f:

$$\begin{aligned} \delta\tilde{J} = &\ (D_1\phi + D_1\psi^T\lambda)\delta x|_{t_f} + (D_2\phi + D_2\psi^T\lambda)\delta t|_{t_f} + \psi^T\delta\lambda \\ + &\ (H - p^T\dot{x})\delta t|_{t_f} \\ + &\ \int_{t_0}^{t_f}\left[D_1H\delta x + D_3H\delta u - p^T\delta\dot{x} + (D_2H^T - \dot{x})^T\delta p\right]dt \end{aligned} \tag{6}$$

The notation D_iH stands for the derivative of H with respect to the i th argument. Thus, for example,

$$D_3H(x,p,u,t) = \frac{\partial H}{\partial u} \quad D_1H(x,p,u,t) = \frac{\partial H}{\partial x}$$

Integrating by parts for $\int p^T\delta\dot{x}dt$ yields

$$\begin{aligned} \delta\tilde{J} = &\ (D_1\phi + D_1\psi^T\lambda - p^T)\delta x(t_f) + (D_2\phi + D_2\psi^T\lambda + H)\delta t_f + \psi^T\delta\lambda \\ + &\ \int_{t_0}^{t_f}\left[(D_1H + \dot{p}^T)\delta x + D_3H\delta u + (D_2^TH - \dot{x})^T\delta p\right]dt \end{aligned} \tag{7}$$

An extremum of $\tilde{J}$ is achieved when $\delta\tilde{J} = 0$ for all independent variations $\delta\lambda, \delta x, \delta u, \delta p$. These conditions are recorded in the following

Table of necessary conditions for optimality:

Table 1

Description	Equation	Variation
Final State constraint	$\psi(x(t_f), t_f) = 0$	$\delta\lambda$
State Equation	$\dot{x} = \frac{\partial H}{\partial p}^T$	δp
Costate equation	$\dot{p} = -\frac{\partial H}{\partial x}^T$	δx
Input stationarity	$\frac{\partial H}{\partial u} = 0$	δu
Boundary conditions	$D_1\phi - p^T = -D_1\psi^T\lambda\|_{t_f}$	$\delta x(t_f)$
	$H + D_2\phi = -D_2\psi^T\lambda\|_{t_f}$	δt_f

The conditions of Table (1) and the boundary conditions $x(t_0) = x_0$ and the constraint on the final state $\psi(x(t_f), t_f) = 0$ constitute the necessary conditions for optimality. The end point constraint equation is referred to as the transversality condition:

$$\begin{aligned} D_1\phi - p^T = &\ -D_1\psi^T\lambda \\ H + D_2\phi = &\ -D_2\psi^T\lambda \end{aligned} \tag{8}$$

The optimality conditions may be written explicitly as

$$\begin{aligned} \dot{x} = & \ \frac{\partial H}{\partial p}^T(x, u^*, p) \\ \dot{p} = & \ -\frac{\partial H}{\partial x}^T(x, u^*, p) \end{aligned} \tag{9}$$

with the stationarity condition reading

$$\frac{\partial H}{\partial u}(x, u^*, p) = 0$$

and the endpoint constraint $\psi(x(t_f), t_f) = 0$. *The key point to the derivation of the necessary conditions of optimality is that the Legendre transformation of the Lagrangian to be minimized into a Hamiltonian converts a functional minimization problem into a static optimization problem on the function* $H(x, u, p, t)$.

The question of when these equations also constitute sufficient conditions for (local) optimality is an important one and needs to be ascertained by taking the second variation of $\tilde{J}$. This is an involved procedure but the input stationarity condition in Table (1) hints at the **sufficient condition for local minimality** of a given trajectory $x^*(\cdot), u^*(\cdot), p^*(\cdot)$ being a local minimum as being that the Hessian of the Hamiltonian,

$$D_2^2 H(x^*, u^*, p^*, t) \tag{10}$$

being positive definite along the optimal trajectory. A sufficient condition for this is to ask simply that the $n_i \times n_i$ Hessian matrix

$$D_2^2 H(x, u, p, t) \tag{11}$$

be positive definite. As far as conditions for **global minimality** are concerned, again the stationarity condition hints at a sufficient condition for global minimality being that

$$u^*(t) = \underset{\{\text{ min over u }\}}{\text{argmin}} \quad H(x^*(t), u, p^*(t), t) \tag{12}$$

Sufficient conditions for this are, for example, the convexity of the Hamiltonian $H(x, u, p, t)$ in u.

Finally, there are instances in which the Hamiltonian $H(x, u, p, t)$ is not a function of u at some values of x, p, t. These cases are referred to as *singular extremals* and need to be treated with care, since the value of u is left unspecified as far as the optimization is concerned.

1.1 Fixed Endpoint problems

In the instance that the final time t_f is fixed, the equations take on a simpler form, since there is no variation in δt_f. Then, the boundary condition of equation (8) becomes

$$p^T(t_f) = D_1\phi + D_1\psi^T\lambda|_{t_f} \tag{13}$$

Further, if there is no final state constraint the boundary condition simplifies even further to

$$p(t_f) = D_1\phi^T|_{t_f} \tag{14}$$

1.2 Time Invariant Systems

In the instance that $f(x,u,t)$ and the running cost $L(x,u,t)$ are not explicitly functions of time, there is no final state constraint and the final time t_f is fixed, the formulas of Table (1) can be rewritten as

State Equation	$\dot{x} = \frac{\partial H}{\partial p}^T = f(x,u^*)$
Costate Equation	$\dot{p} = -\frac{\partial H}{\partial x}^T = -D_1 f^T p + D_1 L^T$
Stationarity Condition	$0 = \frac{\partial H}{\partial u} = D_2 L^T + D_2 f^T p$
Transversality Conditions	$D_1\phi - p^T = -D_1\psi^T\lambda$
	$H(t_f) = 0$

In addition, it may be verified that

$$\frac{dH^*}{dt} = \frac{\partial H^*}{\partial x}(x,p)\dot{x} + \frac{\partial H^*}{\partial p}\dot{p} = 0 \tag{15}$$

thereby establishing that $H^*(t) \equiv 0$.

1.3 Connections with Classical Mechanics

Hamilton's principle of least action states (under certain conditions [2]) that a conservative system moves so as to minimize the time integral of its "action", defined to be the difference between the kinetic and potential energy. To make this more explicit we define $q \in \mathbb{R}^n$ to be the vector of generalized coordinates of the system and denote by $U(q)$ the potential energy of the system and $T(q,\dot{q})$ the kinetic energy of the system. Then Hamilton's principle of least action asks to solve an optimal control problem for the system

$$\dot{q} = u$$

[2]For example, there is no dissipation or no nonholonomic constraints. Holonomic or integrable constraints are dealt with by adding appropriate Lagrange multipliers. If nonholonomic constraints are dealt with in the same manner, we get equations of motion, dubbed vakonomic by Arnold [11] which do not correspond to experimentally observed motions. On the other hand, if there are only holonomic constraints, the equations of motion that we derive from Hamilton's principle of least action is equivalent to Newton's laws.

with Lagrangian

$$L(q,u) = T(q,u) - U(q)$$

The equations (9) in this context have $H(q,u,p) = L(q,u) + p^T u$. $u^* = u^*(p,q)$ is chosen so as to minimize the Hamiltonian H. A necessary condition for stationarity is that $u^*(p,q)$ satisfies

$$0 = \frac{\partial H}{\partial u} = \frac{\partial L}{\partial u} + p^T \tag{16}$$

The form of the equations (9) in this context is that of the familiar *Hamilton Jacobi equations*. The costate p has the interpretation of momentum.

$$\begin{array}{rll} \dot{q} = & \frac{\partial H^*}{\partial p}(p,q) & = u^*(p,q) \\ \dot{p} = & -\frac{\partial H^*}{\partial q}(p,q) & \end{array} \tag{17}$$

Combining the second of these equations with (16) yields the familiar *Euler Lagrange equations*

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0 \tag{18}$$

2 Optimal Control based on Dynamic Programming

To begin this discussion, we will embed the optimization problem which we are solving in a larger class of problems, more specifically we will consider the original cost function of equation (2) from an initial time $t \in [t_0, t_f]$ by considering the cost function on the interval $[t, t_f]$:

$$J(x(t),t) = \phi(x(t_f),t_f) + \int_t^{t_f} L(x(\tau),u(\tau),\tau)d\tau$$

Bellman's principle of optimality says that if we have found the optimal trajectory on the interval from $[t_0, t_f]$ by solving the optimal control problem on that interval, the resulting trajectory is also optimal on all subintervals of this interval of the form $[t, t_f]$ with $t > t_0$, provided that the initial condition at time t was obtained from running the system forward along the optimal trajectory from time t_0. The optimal value of $J(x(t),t)$ is referred to as the "cost-to go". To be able to state the following key theorem of optimal control we will need to define the "optimal Hamiltonian" to be

$$H^*(x,p,t) := H(x,u^*,p,t)$$

Theorem 1 The Hamilton Jacobi Bellman equation

Consider, the time varying optimal control problem of (2) with fixed endpoint t_f and time varying dynamics. If the optimal value function, i.e. $J^*(x(t_0), t_0)$ is a smooth function of x, t, then $J^*(x, t)$ satisfies the **Hamilton Jacobi Bellman** *partial differential equation*

$$\frac{\partial J^*}{\partial t}(x,t) = -H^*(x, \frac{\partial J^*}{\partial x}(x,t), t) \tag{19}$$

with boundary conditions given by $J^*(x, t_f) = \phi(x, t_f)$ for all $x \in \{x : \psi(x, t_f) = 0\}$.

Proof: The proof uses the principle of optimality. This principle says that if we have found the optimal trajectory on the interval from $[t, t_f]$ by solving the optimal control problem on that interval, the resulting trajectory is also optimal on all subintervals of this interval of the form $[t_1, t_f]$ with $t_1 > t$, provided that the initial condition at time t_1 was obtained from running the system forward along the optimal trajectory from time t. Thus, from using $t_1 = t + \Delta t$, it follows that

$$J^*(x,t) = \begin{matrix} \min \\ u(\tau) \\ t \le \tau \le t + \Delta t \end{matrix} \left[\int_t^{t+\Delta t} L(x,u,\tau) d\tau + J^*(x + \Delta x, t + \Delta t) \right] \tag{20}$$

Taking infinitesimals and letting $\Delta t \to 0$ yields that

$$-\frac{\partial J^*}{\partial t} = \begin{matrix} \min \\ u(t) \end{matrix} \left(L + (\frac{\partial J^*}{\partial x}) f \right) \tag{21}$$

with the boundary condition being that the terminal cost is

$$J^*(x, t_f) = \phi(x, t_f)$$

on the surface $\psi(x) = 0$. Using the definition of the Hamiltonian in equation (5), it follows from equation (21) that the Hamilton Jacobi equation of equation (19) holds.

□

Remarks:

1. The preceding theorem was stated as a necessary condition for extremal solutions of the optimal control problem. As far as minimal and global solutions of the optimal control problem, the Hamilton Jacobi Bellman equations read as in equation (21). In this sense, the form of the Hamilton Jacobi Bellman equation in (21) is more general.

2. The **Eulerian conditions** of Table (1) are easily obtained from the Hamilton Jacobi Bellman equation by proving that $p^T(t) := \frac{\partial J^*}{\partial x}(x,t)$ satisfies the costate equations of that Table. Indeed, consider the equation (21). Since $u(t)$ is unconstrained, it follows that it should satisfy

$$\frac{\partial L}{\partial u}(x^*, u^*) + \frac{\partial f}{\partial u}^T p = 0 \tag{22}$$

Now differentiating the definition of $p(t)$ above with respect to t yields

$$\frac{dp}{dt}^T = \frac{\partial^2 J^*}{\partial t \partial x}(x^*, t) + \frac{\partial^2 J^*}{\partial x^2} f(x^*, u^*, t) \tag{23}$$

Differentiating the Hamilton Jacobi equation (21) with respect to x and using the relation (22) for a stationary solution yields

$$-\frac{\partial^2 J^*}{\partial t \partial x}(x^*, t) = \frac{\partial L}{\partial x} + \frac{\partial^2 J^*}{\partial x^2} f + p^T \frac{\partial f}{\partial x} \tag{24}$$

Using equation (24) in equation (23) yields

$$-\dot{p} = \frac{\partial f}{\partial x}^T p + \frac{\partial L}{\partial x}^T \tag{25}$$

establishing that p is indeed the co-state of Table 1. The boundary conditions on $p(t)$ follow from the boundary conditions on the Hamilton Jacobi Bellman equation.

2.1 Constrained Input Problems

In the instance that there are no constraints on the input, the extremal solutions of the optimal control problem are found by simply extremizing the Hamiltonian and deriving the stationarity condition. Thus, if the specification is that $u(t) \in U \subset \mathbb{R}^{n_i}$ then, the optimality condition is that

$$H(x^*, u^*, p^*, t) \leq H(x^*, u, p^*, t) \quad \forall u \in U \tag{26}$$

If the Hamiltonian is convex in u and U is a convex set, there are no specific problems with this condition. In fact, when there is a single input and the set U is a single closed interval, there are several interesting examples of Hamiltonians for which the optimal inputs switch between the endpoints of the interval, resulting in what is referred to as **bang bang control**. However, problems can arise when U is either not convex or compact. In these cases, a concept of a **relaxed control** taking values in the convex hull of U needs to be introduced. As far as an implementation of a control $u(t) \in \text{conv}U$, but not in U, a probabilistic scheme involving switching between values of U whose convex combination u is needs to be devised.

2.2 Free end time problems

In the instance that the final time t_f is free, the transversality conditions are that

$$\begin{aligned} p^T(t_f) = & \quad D_1\phi + D_1\psi^T\lambda \\ H(t_f) = & \quad -(D_2\phi + D_2\psi^T\lambda) \end{aligned} \tag{27}$$

2.2.1 Minimum time problems

A special class of minimum time problems of especial interest is minimum time problems, where t_f is to be minimized subject to the constraints. This is accounted for by setting the Lagrangian to be 1, and the terminal state cost $\phi \equiv 0$, so that the Hamiltonian is $H(x,u,p,t) = 1 + p^T f(x,u,t)$. Note that by differentiating $H(x,u,p,t)$ with respect to time, we get

$$\frac{dH^*}{dt} = D_1 H^* \dot{x} + D_2 H^* \dot{u} + D_3 H^* \dot{p} + \frac{\partial H^*}{\partial t} \tag{28}$$

Continuing the calculation using the Hamilton Jacobi equation,

$$\frac{dH^*}{dt} = (\frac{\partial H^*}{\partial x} + \dot{p}) f(x, u^*, t) + \frac{\partial H^*}{\partial t} = \frac{\partial H^*}{\partial t} \tag{29}$$

In particular, if H^* is not an explicit function of t, it follows that $H^*(x,u,p,t) \equiv H$. Thus, for minimum time problems for which $f(x,u,t)$ and $\psi(x,t)$ are not explicitly functions of t, it follows that $0 = H(t_f) \equiv H(t)$.

3 Linear Quadratic Regulator (LQR)

In this section, we present the Linear Quadratic Regulator (LQR) as a practical example of a continuous time optimal control problem with an explicit differential equation describing the solution. We will then present the discrete-time version of the LQR problem as a means of introducing discrete-time optimal control problems and their relation to the continuous version. We will also present the discrete-time LQR in the framework of convex optimization and mention some methods for computing solutions.

The material here will borrow from Stephen Boyd's lectures in Stanford's EE363 class, which gives an excellent presentation of this topic [14].

3.1 Continuous-time Linear Quadratic Regulator

The Linear Quadratic Regulator is a classical problem first formulated by Rudolf Kalman in the 1960's [15]. The problem involves finding the optimal control policies for a system with linear dynamics and a quadratic running-cost. In particular, we consider the time-invariant dynamical system described by

$$\dot{x} = Ax + Bu \tag{30}$$

with state $x(t) \in \mathbb{R}^n$ and control input $u(t) \in \mathbb{R}^m$.

We wish to minimize the performance index

$$J = \frac{1}{2}\left(x(T)^\intercal Q_T x(T) + \int_{t_0}^{T} x(\tau)^\intercal Q x(\tau) + u(\tau)^\intercal R u(\tau) d\tau \right) \tag{31}$$

Here $Q_T = Q_T^\intercal \succeq 0$, $Q = Q^\intercal \succeq 0$, and $R = R^\intercal \succ 0$ define the quadratic cost terms of this objective.

We can write the necessary conditions of optimally given by the calculus of variations in Table 1 for this problem. We have:

$$\begin{aligned} H(x,u,p) &= \frac{1}{2}x^\intercal Q x + \frac{1}{2}u^\intercal R u + p^\intercal (Ax + Bu) && (32) \\ \dot{x} &= Ax + Bu && (33) \\ x(0) &= x_0 && (34) \\ \dot{p} &= -A^\intercal p - Qx && (35) \\ 0 &= Ru + B^\intercal p && (36) \\ p(T) &= Q_T x(T) && (37) \end{aligned}$$

Note that $D_2 H(x,u,p) = R \succ 0$, so the necessary conditions for optimality are also sufficient for a global minimum. Furthermore, because from Eq. (36) we have that $u(t) = -R^{-1}B^\intercal p(t)$, we can combine the above equations into the following differential equation:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^\intercal \\ -Q & -A^\intercal \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}. \tag{38}$$

Subject to boundary point conditions $x(0) = x_0$ and $p(T) = Q_T x(T)$.

Claim: There exists a positive-semi-definite matrix V given by the following matrix differential equation, known as the Riccati differential equation:

$$-\dot{V} = A^\intercal V + VA - VBR^{-1}B^\intercal V + Q, \quad V(T) = Q_T, \tag{39}$$

such that $p(t) = V(t)x(t)$.

Proof of claim: Show that p = Vx satisfies the relation $\dot{p} = -A^\intercal p - Qx$

$$
\begin{aligned}
\dot{p} &= \dot{V}x + V\dot{x} && (40)\\
&= -(A^\intercal V + VA - VBR^{-1}B^\intercal V + Q)x + V(Ax - BR^{-1}B^\intercal p) && (41)\\
&= -Qx - A^\intercal Vx + VBR^{-1}B^\intercal Vx - VBR^{-1}B^\intercal Vx && (42)\\
&= -Qx - A^\intercal p && (43)
\end{aligned}
$$

Hence, by first integrating the value function V backwards in time starting from the initial (final) condition $V(T) = Q_T$, we can substitute $p(t) = V(t)x(t)$ and solve forward in time the system $\dot{x} = Ax + Bu^*$, where $u^*(t) = -R^{-1}B^\intercal V(t)x(t)$, starting from the initial condition $x(0) = x_0$. This provides a form for the optimal trajectory $x(t)$, as well as the optimal control sequence $u^*(t)$ as a linear function of the state.

3.2 Discrete-Time LQR

We have derived forms of the value function (and accompanying linear feedback policies) for the continuous-time LQR problem in terms of matrix differential equations. In order to compute these quantities as presented, however, we must numerically integrate the Riccati differential equation. Alternatively, we can discretize our dynamics using Euler Integration, leading to a discrete-time LQR problem. The solution to this discrete-time version will approach the solution to the continuous-time version in the limit as we decrease the interval of time-discretization to 0.

To see this, first choose a step size $h > 0$ such that $Nh = T$ for some $N \in \mathbb{N}$. The discrete dynamics equations according to forward-Euler integration become:

$$
\begin{aligned}
x_{k+1} = x((k+1)h) \approx x(kh) + h\dot{x}(kh) = (I + hA)x_k + hBu_k && (44)\\
&& (45)
\end{aligned}
$$

We similarly approximate the continuous-time performance index by the following:

$$
\begin{aligned}
J &\approx \frac{1}{2}\left(x(Nh)^\intercal Q_T x(Nh) + h\sum_{k=0}^{N-1} x(kh)^\intercal Q x(kh) + u(kh)^\intercal R u(kh)\right) && (46)\\
&= \frac{1}{2}\left(x_N^\intercal Q_T x_N + \sum_{k=0}^{N-1} x_k^\intercal \bar{Q} x_k + u_k^\intercal \bar{R} u_k\right). && (47)
\end{aligned}
$$

Here $\bar{R} = hR$ and $\bar{Q} = hQ$.

The dynamic programming solution to the discrete-time LQR problem gives the relation $u_k = K_k x_k$ where

$$K_k = -(\bar{R} + \bar{B}^\intercal V_{k+1} \bar{B})^{-1} \bar{B}^\intercal V_{k+1} \bar{A} \tag{48}$$

$$V_k = \bar{Q} + \bar{A}^\intercal V_{k+1} \bar{A} - \bar{A}^\intercal V_{k+1} \bar{B} (\bar{R} + \bar{B}^\intercal V_{k+1} \bar{B})^{-1} \bar{B}^\intercal V_{k+1} \bar{A} \tag{49}$$

Here V_k is a discrete-time verison of the value function, and we have defined $\bar{A} = (I + hA)$ and $\bar{B} = hB$. Substituting and removing all terms with h^2 or higher-order terms, we get

$$V_k \approx hQ + V_{k+1} + hA^\intercal V_{k+1} + hV_{k+1}A - hV_{k+1}BR^{-1}B^\intercal V_{k+1} \tag{50}$$

Rearranging, we have

$$-\frac{1}{h}(V_{k+1} - V_k) = Q + A^\intercal V_{k+1} + V_{k+1}A - V_{k+1}BR^{-1}B^\intercal V_{k+1} \tag{51}$$

And in the limit $h \to 0$, we recover the continuous-time Riccati differential equation

$$\dot{V} = -(Q + A^\intercal V + VA - VBR^{-1}B^\intercal V). \tag{52}$$

Therefore, we have that $V_k \to V(kh)$ as $h \to 0$.

3.3 Discrete-Time LQR as a Convex Optimization Problem

We have now seen two methods for practically computing the optimal control policy and state trajectory for linear systems subject to a quadratic cost. One solution is to numerically integrate the discrete-time Riccati differential equation backward in time and then forward integrate the state trajectory using the control policy defined in terms of the value function. The other is to discretize the problem up-front and use the discrete-time solution as an approximation to the continuous-time solution.

One of the advantages of discretizing the problem up-front is that in discrete-time, finding optimal trajectories becomes a finite-dimensional optimization problem, and we can leverage powerful tools from optimization theory to find these solutions. To see this, we will cast our discrete-time LQR problem as a convex optimization problem, which will open the door to us viewing optimal control problems as optimization problems and studying useful solution techniques.

In presenting this method, we will use a slightly more general form of the Linear-Quadratic problem. We will consider a system with time-varying affine dynamics of the form $x_{t+1} =$

$A_t x_t + B_t u_t + c_t$. We consider minimizing the performance objective given by time-varying cost

$$J = \frac{1}{2}\left(x_N^\intercal Q_T x_N + 2q_N^\intercal x_N + \sum_{k=0}^{N-1} x_k^\intercal Q x_k + 2u_k^\intercal S_k x_k + u_k^\intercal R u_k + 2q_k^\intercal x_k + 2r_k^\intercal u_k\right) \tag{53}$$

If we stack the variables x_k, u_k into a single trajectory vector,

$$z = \left[x_0, u_0, ..., x_{N-1}, u_{N-1}, x_N\right], \tag{54}$$

we can express our performance index as simply

$$J = \frac{1}{2} z^\intercal H z + h^\intercal z \tag{55}$$

Here the terms H and h are given by

$$H = \begin{bmatrix} Q_0 & S_0^\intercal & & & & \\ S_0 & R_0 & & & & \\ & & Q_1 & S_1^\intercal & & \\ & & S_1 & R_1 & & \\ & & & & \ddots & \\ & & & & & Q_N \end{bmatrix}, h = \begin{bmatrix} q_0 \\ r_0 \\ q_1 \\ r_1 \\ \vdots \\ q_N \end{bmatrix} \tag{56}$$

Similarly, we can express the initial condition constraint as well as the dynamics constraints in the compact form $Gz + g = 0$ where

$$G = \begin{bmatrix} I & & & & & & \\ -A_0 & -B_0 & I & & & & \\ & & -A_1 & -B_1 & I & & \\ & & & & -A_2 & -B_2 & I \\ & & & & & & \ddots \end{bmatrix}, g = \begin{bmatrix} -x_{init} \\ -c_0 \\ -c_1 \\ -c_2 \\ \vdots \end{bmatrix} \tag{57}$$

Combining the performance objective and constraint terms, we have the following optimization problem:

$$\min_z \quad \frac{1}{2} z^\intercal H z + h^\intercal z \tag{58}$$

$$\text{s.t.} \quad Gz + g = 0 \tag{59}$$

The objective is a quadratic function of z and the constraints are linear. Hence, this is a standard Equality-Constrained Quadratic Program. The Karush-Kuhn-Tucker conditions of optimality [16] for this problem are given by

$$Hz + h + G^{\mathsf{T}}\lambda = 0 \tag{60}$$

$$Gz + g = 0, \tag{61}$$

Here λ is a vector of Lagrange multipliers of appropriate dimension associated with the constraint $Gz = g$. As it turns out, the vector $\lambda = \left[p_0, ..., p_N\right]$ where $p_k = V_k x_k$ are the discrete-time co-state vectors. These conditions define the following system of equations:

$$\begin{bmatrix} H & G^{\mathsf{T}} \\ G & 0 \end{bmatrix} \begin{bmatrix} z \\ \lambda \end{bmatrix} = \begin{bmatrix} -h \\ -g \end{bmatrix}. \tag{62}$$

It turns out that system above has a unique sparsity structure, and can be permuted such that the matrix on the left-hand side is banded. This allows for very efficient computation of the solution vectors z and λ.

Not all types of convex problems have closed-form solutions like this one. In fact, most do not. However, there exist many efficient solvers for convex problems which are guaranteed to find globally optimal solutions.

References

[1] S. S. Sastry, *Lectures in Optimal Control and Dynamic Games*, Notes for the course EECS290A, Advanced Topics in Control Theory, University of California, Berkeley, 1996.

[2] M. Athans and P. Falb, *Optimal Control*, Mc Graw Hill, 1966.

[3] R. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

[4] L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, 1974.

[5] A. E. Bryson and Y-C. Ho, *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, 1969.

[6] L. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley-New York, 1962.

[7] L. C. Young, *Optimal Control Theory*, Cambridge University Press, 2nd edition, 1980.

[8] D. Kirk, *Optimal Control Theory: An Introduction*, Prentice Hall, 1970.

[9] F. Lewis, *Optimal Control*, Wiley-New York, 1986.

[10] W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control*, Springer Verlag, 1975.

[11] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer Verlag, 2nd edition, 1989.

[12] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1947.

[13] T. Basar and G. Olsder, *Dynamic Noncooperative Games*, Academic Press, 2nd Edition, 1995.

[14] Stephen Boyd, *EE363: Linear Dynamical Systems*, Stanford course website (2021): https://web.stanford.edu/class/ee363/.

[15] Rudolf Emil Kalman and others. *Contributions to the theory of optimal control*, Bol. Soc. Mat. Mexicana, 5(2), pp. 102–119, 1960.

[16] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.