Optimal control

Wikipedia: Optimal control

★ General method

Minimize the continuous-time cost functional

$$J = \Phi[x(t_0), t_0, x(t_f), t_f] + \int_{t_0}^{t_f} \mathcal{L}[x(t), u(t), t]dt$$

subject to the first-order dynamic constraints (the state equation)

$$\dot{x}(t) = a[x(t), u(t), t]$$

the algebraic path constraints

$$b[x(t), u(t), t] \leq 0$$

and the boundary conditions

$$\Phi[x(t_0), t_0, x(t_f), t_f] = 0$$

where $x(t)$ is the state, $u(t)$ is the control, t is the independent variable (generally speaking, time), t_0 is the initial time, and t_f is the terminal time. The term Φ and $\mathcal{L}$ are called the endpoint cost and Lagrangian, respectively.

★ Linear quadratic control

A special case of the general nonlinear optimal control problem given in the previous section is the linear quadratic (LQ) optimal control problem. The LQ problem is stated as follows. Minimize the quadratic continuous-time cost functional

$$J = \frac{1}{2}x^T(t_f)S_f x(t_f) + \frac{1}{2}\int_{t_0}^{t_f}[x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]dt$$

Subject to the linear first-order dynamic constraints

$$(x) = A(t)x(t) + B(t)u(t)$$

and the initial condition

$$x(t_0) = x_0$$

A particular form of the LQ problem that arises in many control system problems is that of the linear quadratic regulator (LQR) where all of the matrices (i.e., A, B, Q, and R) are constant, the initial time is arbitrarily set to zero, and the terminal time is taken in the limit $t_f \leftarrow \infty$ (this last assumption is what is known as infinite horizon). The LQR problem is stated as follows. Minimize the infinite horizon quadratic continuous-time cost functional

$$J = \frac{1}{2}\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)]dt$$

Subject to the linear time-invariant first-order dynamic constraints

$$\dot{x}(t) = Ax(t) + Bu(t)$$

and the initial condition

$$x(t_0) = x_0$$

In the finite-horizon case the matrices are restricted in that Q and R are positive semi-definite and positive definite, respectively. In the infinite-horizon case, however, the matrices Q and R re not only positive-semidefinite and positive-definite, respectively, but are also constant. These additional restrictions on Q and R in the infinite-horizon case are enforced to ensure that the cost functional remains positive. Furthermore, in order to ensure that the cost function is bounded, the additional restriction is imposed that the pair (A, B) is controllable. Note that the LQ or LQR cost functional can be thought of physically as attempting to minimize the control energy (measured as a quadratic form).

    The infinite horizon problem (i.e., LQR) may seem overly restrictive and essentially useless because it assumes that the operator is driving the system to zero-state and hence driving the output of the system to zero. This is indeed correct. However the problem of driving the output to a desired nonzero level can be solved after the zero output one is. In fact, it can be proved that this secondary LQR problem can be solved in a very straightforward manner. It has been shown in classical optimal control theory that the LQ (or LQR) optimal control has the feedback form

$$u(t) = -K(t)x(t)$$

where $K(t)$ is a properly dimensioned matrix, given as

$$K(t) = R^{-1}B^T S(t)$$

and $S(t)$ is the solution of the differential Riccati equation. The differential Riccati equation is given as

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

For the finite horizon LQ problem, the Riccati equation is integrated backward in time using the terminal boundary condition

$$S(t_f) = S_f$$

For the infinite horizon LQR problem, the differential Riccati equation is replaced with the algebraic Riccati equation (ARE) given as

$$0 = -SA - A^T S + SBR^{-1}B^T S - Q$$

Understanding that the ARE arises from infinite horizon problem, the matrices A, B, Q and R are all constant. It is noted that there are in general multiple solutions to the algebraic Riccati equation and the positive definite (or positive semi-definite) solution is the one that is used to compute the feedback gain. The LQ (LQR) problem was elegantly solved by Rudolf Kalman.

* Numerical methods for optimal control
* Discrete-time optimal control

EX: Finite time

Consider the problem of a mine owner who must decide at what rate to extract ore from his mine. He owns rights to the ore from date 0 and T. At date 0 there is x_0 ore in the ground, and the instantaneous stock of ore $x(t)$ declines at the rate the mine owner extracts it $u(t)$. The mine owner extracts ore at cost $u(t)^2/x(t)$ and sells ore at a constant price p. He does not value the ore remaining in the ground at time T (there is no "scrap value"). He chooses the rate of extraction in time $u(t)$ to maximize profits over the period of ownership with no time discounting.

1. Discrete-time version
The manager maximizes profit Π:

$$\Pi = \sum_{t=0}^{T-1} [pu_t - \frac{u_t^2}{x_t}]$$

subject to the law of evolution for the state variable x_t

$$x_{t+1} - x_t = -u_t$$

Form the Hamiltonian and differentiate:

$$H = pu_t - \frac{u_t^2}{x_t} - \lambda_{t+1} u_t$$

$$\frac{\partial H}{\partial u_t} = p - \lambda_{t+1} - 2\frac{u_t}{x_t} = 0$$

$$\lambda_{t+1} - \lambda_t = -\frac{\partial H}{\partial x_t} = -(\frac{u_t}{x_t})^2$$

As the mine owner does not value the ore remaining at time T,

$$\lambda_T = 0$$

Using the above equations, it is easy to solve for the x_t and λ_t series

$$\lambda_t = \lambda_{t+1} + \frac{(p - \lambda_{t+1})^2}{4}$$

$$x_{t+1} = x_t \frac{2 - p + \lambda_{t+1}}{2}$$

and using the initial and turn-T conditions, the x_t series can be solved explicitly, giving u_t.

Figure 1: Discrete-time version

2. Continuous-time version
The manager maximizes profit Π:

$$\Pi = \int_0^T [pu(t) - \frac{u(t)^2}{x(t)}]dt$$

subject to the law of evolution for the state variable $x(t)$

$$\dot{x}(t) = -u(t)$$

Form the Hamiltonian and differentiate:

$$H = pu(t) - \frac{u(t)^2}{x(t)} - \lambda(t)u(t)$$

$$\frac{\partial H}{\partial u} = p - \lambda(t) - 2\frac{u(t)}{x(t)} = 0$$

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x} = -(\frac{u(t)}{x(t)})^2$$

As the mine owner does not value the ore remaining at time T,

$$\lambda(T) = 0$$

Using the above equations, it is easy to solve for the differential equations governing $u(t)$ and $\lambda(t)$

$$\dot{\lambda}(t) = -\frac{(p - \lambda(t))^2}{4}$$

$$u(t) = x(t)\frac{p - \lambda(t)}{2}$$

and using the initial and turn-T conditions, the functions can be solved to yield

$$x(t) = \frac{(4 - pt + pT)^2}{(4 + pT)^2}x_0$$

Figure 2: Continuous-time version