Module 11: Introduction to Optimal Control
Lecture Note 1

1 Introduction to optimal control

In the past lectures, although we have designed controllers based on some criteria, but we have never considered optimality of the controller with respect to some index. In this context, Linear Quadratic Regular is a very popular design technique.

The optimal control theory relies on design techniques that maximize or minimize a given performance index which is a measure of the effectiveness of the controller.

Euler-Lagrange equation is a very popular equation in the context of minimization or maximization of a functional.

A functional is a mapping or transformation that depends on one or more functions and the values of the functionals are numbers. Examples of functionals are performance indices which will be introduced later.

In the following section we would discuss the Euler-Lagrange equation for discrete time systems.

1.1 Discrete Euler-Lagrange Equation

A large class of optimal digital controller design aims to minimize or maximize a performance index of the following form.

$$J = \sum_{k=0}^{N-1} F(k, \boldsymbol{x}(k), \boldsymbol{x}(k+1), \boldsymbol{u}(k))$$

where $F(k, \boldsymbol{x}(k), \boldsymbol{x}(k+1), \boldsymbol{u}(k))$ is a differentiable scalar function and $\boldsymbol{x}(k) \in R^n$, $\boldsymbol{u}(k) \in R^m$.

The minimization or maximization of J is subject to the following constraint.

$$\boldsymbol{x}(k+1) = f(k, \boldsymbol{x}(k), \boldsymbol{u}(k))$$

The above can be the state equation of the system, as well as other equality or inequality constraints.

Design techniques for optimal control theory mostly rely on the calculus of variation, according to which, the problem of minimizing one function while it is subject to equality constraints is solved by adjoining the constraint to the function to be minimized.

Let $\boldsymbol{\lambda}(k+1) \in R^{n\times 1}$ be defined as the Lagrange multiplier. Adjoining J with the constraint equation,

$$J_a = \sum_{k=0}^{N-1} F(k, \boldsymbol{x}(k), \boldsymbol{x}(k+1), \boldsymbol{u}(k)) + < \boldsymbol{\lambda}(k+1), [\boldsymbol{x}(k+1) - f(k, \boldsymbol{x}(k), \boldsymbol{u}(k))] >$$

where $< . >$ denotes the inner product.

Calculus of variation says that the minimization of J with constraint is equivalent to the minimization of J_a without any constraint.

Let $\boldsymbol{x}^*(k)$, $\boldsymbol{x}^*(k+1)$, $\boldsymbol{u}^*(k))$ and $\boldsymbol{\lambda}^*(k+1)$ represent the vectors corresponding to optimal trajectories. Thus one can write

$$\begin{aligned}
\boldsymbol{x}(k) &= \boldsymbol{x}^*(k) + \epsilon\boldsymbol{\eta}(k)\\
\boldsymbol{x}(k+1) &= \boldsymbol{x}^*(k+1) + \epsilon\boldsymbol{\eta}(k+1)\\
\boldsymbol{u}(k) &= \boldsymbol{u}^*(k) + \delta\boldsymbol{\mu}(k)\\
\boldsymbol{\lambda}(k+1) &= \boldsymbol{\lambda}^*(k+1) + \gamma\boldsymbol{\nu}(k+1)
\end{aligned}$$

where $\boldsymbol{\eta}(k)$, $\boldsymbol{\mu}(k)$, $\boldsymbol{\nu}(k)$ are arbitrary vectors and ϵ, δ, γ are small constants.

Substituting the above four equations in the expression of J_a,

$$\begin{aligned}
J_a = & \sum_{k=0}^{N-1} F(k, \boldsymbol{x}^*(k) + \epsilon\boldsymbol{\eta}(k), \boldsymbol{x}^*(k+1) + \epsilon\boldsymbol{\eta}(k+1), \boldsymbol{u}^*(k) + \delta\boldsymbol{\mu}(k)) + \\
& < \boldsymbol{\lambda}^*(k+1) + \gamma\boldsymbol{\nu}(k+1), [\boldsymbol{x}^*(k+1) + \epsilon\boldsymbol{\eta}(k+1) - f(k, \boldsymbol{x}^*(k) + \epsilon\boldsymbol{\eta}(k), \boldsymbol{u}^*(k) + \delta\boldsymbol{\mu}(k))] >
\end{aligned}$$

To simplify the notation, let us denote J_a as

$$J_a = \sum_{k=0}^{N-1} F_a(k, \boldsymbol{x}(k), \boldsymbol{x}(k+1), \boldsymbol{u}(k)), \boldsymbol{\lambda}(k+1))$$

Expanding F_a using Taylor series around $\boldsymbol{x}^*(k),\ \boldsymbol{x}^*(k+1),\ \boldsymbol{u}^*(k))$ and $\boldsymbol{\lambda}^*(k+1)$, we get

$$\begin{aligned} F_a(k,\boldsymbol{x}(k),\boldsymbol{x}(k+1),\boldsymbol{u}(k),\boldsymbol{\lambda}(k+1)) = F_a(k,\boldsymbol{x}^*(k),\boldsymbol{x}^*(k+1),\boldsymbol{u}^*(k),\boldsymbol{\lambda}^*(k+1)) + \\ \left\langle \epsilon\boldsymbol{\eta}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k)} \right\rangle + \left\langle \epsilon\boldsymbol{\eta}(k+1), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k+1)} \right\rangle + \left\langle \delta\boldsymbol{\mu}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{u}^*(k)} \right\rangle + \\ \left\langle \gamma\boldsymbol{\nu}(k+1), \frac{\partial F_a^*(k)}{\partial \boldsymbol{\lambda}^*(k+1)} \right\rangle + \text{ higher order terms} \end{aligned}$$

where

$$F_a^*(k) = F_a(k,\boldsymbol{x}^*(k),\boldsymbol{x}^*(k+1),\boldsymbol{u}^*(k),\boldsymbol{\lambda}^*(k+1))$$

The necessary condition for J_a to be minimum is

$$\begin{aligned} \left.\frac{\partial J_a}{\partial \epsilon}\right|_{\epsilon=\delta=\gamma=0} &= 0 \\ \left.\frac{\partial J_a}{\partial \delta}\right|_{\epsilon=\delta=\gamma=0} &= 0 \\ \left.\frac{\partial J_a}{\partial \gamma}\right|_{\epsilon=\delta=\gamma=0} &= 0 \end{aligned}$$

Substituting F_a into the expression of J_a and applying the necessary conditions,

$$\sum_{k=0}^{N-1}\left[\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k)} \right\rangle + \left\langle \boldsymbol{\eta}(k+1), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k+1)} \right\rangle\right] = 0 \tag{1}$$

$$\sum_{k=0}^{N-1}\left\langle \boldsymbol{\mu}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{u}^*(k)} \right\rangle = 0 \tag{2}$$

$$\sum_{k=0}^{N-1}\left\langle \boldsymbol{\nu}(k+1), \frac{\partial F_a^*(k)}{\partial \boldsymbol{\lambda}^*(k+1)} \right\rangle = 0 \tag{3}$$

Equation (1) can be rewritten as

$$\begin{aligned} \sum_{k=0}^{N-1}\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k)} \right\rangle &= -\sum_{k=1}^{N}\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle \\ &= -\sum_{k=0}^{N-1}\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle + \\ &\quad \left.\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle\right|_{k=0} - \left.\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle\right|_{k=N} \end{aligned}$$

where

$$F_a^*(k-1) = F_a(k-1, \boldsymbol{x}^*(k-1), \boldsymbol{x}^*(k), \boldsymbol{u}^*(k-1), \boldsymbol{\lambda}^*(k))$$

Rearranging terms in the last equation, we get

$$\sum_{k=0}^{N-1} \left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k)} + \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle + \left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle \Bigg|_{k=0}^{k=N} = 0 \tag{4}$$

According to the fundamental lemma of calculus of variation, equation (4) is satisfied for any $\boldsymbol{\eta}(k)$ only when the two components of the equation are individually zero. Thus,

$$\frac{\partial F_a^*(k)}{\partial \boldsymbol{x}^*(k)} + \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} = 0 \tag{5}$$

$$\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle \Bigg|_{k=0}^{k=N} = 0 \tag{6}$$

Equation (5) is known as the discrete Euler-Lagrange equation and equation (6) is called the transversality condition which is nothing but the boundary condition needed to solve equation (5).

Discrete Euler-Lagrange equation is the necessary condition that must be satisfied for J_a to be an extremal.

With reference to the additional conditions (2) and (3), for arbitrary $\boldsymbol{\mu}(k)$ and $\boldsymbol{\nu}(k+1)$,

$$\frac{\partial F_a^*(k)}{\partial u_j^*(k)} = 0, \quad j = 1, 2, \cdots m \tag{7}$$

$$\frac{\partial F_a^*(k)}{\partial \lambda_i^*(k+1)} = 0, \quad i = 1, 2, \cdots n \tag{8}$$

Equation (8) leads to

$$\boldsymbol{x}^*(k+1) = f(k, \boldsymbol{x}^*(k), \boldsymbol{u}^*(k))$$

which means that the state equation should satisfy the optimal trajectory. Equation (7) gives the optimal control $\boldsymbol{u}^*(k)$ in terms of $\boldsymbol{\lambda}^*(k+1)$

In a variety of the design problems, the initial state $\boldsymbol{x}(0)$ is given. Thus $\boldsymbol{\eta}(0) = 0$ since $\boldsymbol{x}(0)$ is fixed. Hence the tranversality condition reduces to

$$\left\langle \boldsymbol{\eta}(k), \frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)} \right\rangle \Bigg|_{k=N} = 0$$

Again, a number of optimal control problems are classified according to the final conditions.

If $\boldsymbol{x}(N)$ is given and fixed, the problem is known as fixed-endpoint design. On the other hand, if $\boldsymbol{x}(N)$ is free, the problem is called a free endpoint design.

For fixed endpoint ($\boldsymbol{x}(N)$ = fixed, $\boldsymbol{\eta}(N) = 0$) problems, no transversality condition is required to solve.

For free endpoint the transversality condition is given as follows.

$$\left.\frac{\partial F_a^*(k-1)}{\partial \boldsymbol{x}^*(k)}\right|_{k=N} = 0$$

For more details, one can consult **Digital Control Systems** by B. C. Kuo.