Optimal Control Theory

A Preprint

Hao Jin
Academy for Advanced Interdisciplinary Studies
Peking University
jin.hao@pku.edu.cn

August 28, 2019

1 The Maximum Principle: Continuous Time

1.1 Problem Setting

$$\max_{u(t)\in\Omega(t)} J = \int_0^T F(x,u,t)dt + S(x(T),T)$$
$$s.t. \quad \dot{x} = f(x,u,t),\ x(0) = x_0.$$

1.2 Necessary Conditions

The necessary conditions of $u^*(t)$ being optimal control and $x^*(t)$ being optimal path are

$$\begin{aligned}
&\dot{x}^* = f(x^*,u^*,t),\ x^*(0) = x_0,\\
&\dot{\lambda} = -H_x(x^*,u^*,\lambda,t),\ \lambda(T) = S_x(x^*(T),T),\\
&H(x^*,u^*,\lambda,t) \geq H(x^*,u,\lambda,t),\ \forall u \in \Omega(t),\ t \in [0,T],
\end{aligned}$$

where

$$H(x,u,\lambda,t) = F(x,u,t) + \lambda f(x,u,t).$$

Once the terminal time T is undecided, the terminal time $T = T^*$ has to satisfy

$$H(x^*(T^*),u^*(T^*),\lambda(T^*),T^*) + S_T(x^*(T^*),T^*) = 0$$

1.3 Sufficiency Conditions

Theorem 1. *Let $u^*(t)$, and the corresponding $x^*(t)$ and $\lambda(t)$ satisfy the necessary conditions for all $t \in [0,T]$. Then, u^* is an optimal control if $H^0[x,\lambda_t,t]$ is concave in x for each t and $S(x,T)$ is concave in x where*

$$H^0(x(t),\lambda(t),t) = \max_{u(t)\in\Omega(t)} H(x(t),u(t),\lambda(t),t)$$

2 The Maximum Principle: Mixed Inequality

2.1 Problem Setting

$$
\begin{aligned}
\max_u J = & \int_0^T F(x,u,t)dt + S(x(T),T) \\
s.t. \quad & \dot{x} = f(x,u,t),\ x(0) = x_0. \\
& g(x,u,t) \geq 0, \\
& a(x(T),T) \geq 0, b(x(T),T) = 0 \ \ (or\ x(T) \in Y(T))
\end{aligned}
$$

2.2 Necessary Conditions

The necessary conditions of $u^*(t)$ being optimal control and $x^*(t)$ being optimal path are separated as below:

- Basic dynamics of the optimal path x^*

$$
\begin{aligned}
&\dot{x}^* = f(x^*,u^*,t),\ x^*(0) = x_0, \\
&a(x^*(T),T) \geq 0, b(x^*(T),T) = 0.
\end{aligned}
$$

- Dynamics of the shadow price λ and the terminal conditions

$$
\begin{aligned}
&\dot{\lambda} = -L_x(x^*,u^*,\lambda,\mu,t), \\
&\lambda(T) = S_x(x^*(T),T) + \alpha a_x(x^*(T),T) + \beta b_x(x^*(T),T),\ where\ \alpha \geq 0\ and\ \alpha a(x^*(T),T) = 0, \\
&or\ (\lambda(T) - S_x(x^*(T),T))(y - x^*(T)) \geq 0, \forall y \in Y(T),
\end{aligned}
$$

where

$$L(x,u,\lambda,\mu,t) = H(x,u,\lambda,t) + \mu g(x,u.t).$$

- The Hamilton maximizing condition

$$H(x^*,u^*,\lambda,t) \geq H(x^*,u,\lambda,t), \forall u\ satisfying\ g(x^*,u,t) \geq 0, \forall t \in [0,T].$$

- Lagrange multiplier has to satisfy

$$
\begin{aligned}
&\frac{\partial L}{\partial u}|_{u=u^*} = (\frac{\partial H}{\partial u} + \mu\frac{\partial g}{\partial u})|_{u=u^*} = 0, \\
&\mu \geq 0, and\ \mu g(x^*,u^*,t) = 0.
\end{aligned}
$$

Once the terminal time T is undecided, the terminal time $T = T^*$ has to satisfy

$$H(x^*(T^*),u^*(T^*),\lambda(T^*),T^*) + S_T(x^*(T^*),T^*) + \alpha a_T(x^*(T^*),T^*) + \beta b_T(x^*(T^*),T^*) = 0.$$

The Lagrange multiplier $\lambda(t)$ is interpreted as the marginal value of an increment in the state variable x at time t, while other Lagrange multipliers $\alpha,\ \beta,\ \mu(t)$ is interpreted the influence ratio on the objective function J with respect to the ϵ-relaxation of the constraints.

2.3 Sufficiency Conditions

Theorem 2. *Let $(x^*,u^*,\lambda,\mu,\alpha,\beta)$ satisfy the necessary conditions for all $t \in [0,T]$. Then, u^* is an optimal control if $H^0[x,\lambda(t),t]$ is concave in x for each t and $S(x,T)$ is concave in x, g is quasiconcave in (x,u), a is quasiconcave in x, and b is linear in x, then (x^*,u^*) is optimal.*

3 The Maximum Principle: Pure State and Mixed Inequality Constraints

Pure state constraints are constraints that require certain inequality constraints only on state variables, which include constraints that require certain state variables to remain non-negative. Such pure state constraints are different and more complicated. That is because that when such constraints become tight, they do not provide any direct information to the decision maker on how to choose values for the control variables. Specifically, such pure state constraints generally result in the discontinuous marginal valuation of the state variable.

3.1 Problem Setting

$$
\begin{aligned}
\max_u\ & J = \int_0^T F(x,u,t)dt + S(x(T),T) \\
s.t.\ \ & \dot{x} = f(x,u,t),\ x(0) = x_0. \\
& g(x,u,t) \geq 0, \\
& a(x(T),T) \geq 0, b(x(T),T) = 0\ \ (or\ x(T) \in Y(T)), \\
& h(x,t) \geq 0.
\end{aligned}
$$

3.2 Necessary Conditions: Direct Method

Direct method directly adds the pure state constraint into the Lagrange function with a multiplier. The necessary conditions of $u^*(t)$ being optimal control and $x^*(t)$ being optimal path are separated as below:

- Basic dynamics of the optimal path x^* satisfying mixed inequality and pure state constraints

$$
\begin{aligned}
& \dot{x}^* = f(x^*,u^*,t),\ x^*(0) = x_0, \\
& g(x^*,u^*,t) \geq 0, h(x^*,t) \geq 0, \\
& a(x^*(T),T) \geq 0, b(x^*(T),T) = 0.
\end{aligned}
$$

- Dynamics of the shadow price λ^d and the terminal conditions

$$
\begin{aligned}
& \dot{\lambda}^d = -L_x^d(x^*,u^*,\lambda^d,\mu,\eta^d,t), \\
& \lambda^d(T^-) = S_x(x^*(T),T) + \alpha a_x(x^*(T),T) + \beta b_x(x^*(T),T) + \gamma^d h_x(x^*(T),T), \\
& \quad where\ \alpha \geq 0, \alpha a(x^*(T),T) = 0, \gamma^d \geq 0, \gamma^d h(x^*(T),T) = 0
\end{aligned}
$$

  where

$$
\begin{aligned}
& H^d(x,u,\lambda^d,t) = F(x,u,t) + \lambda^d f(x,u,t), \\
& L^d(x,u,\lambda^d,\mu,\eta^d,t) = H^d(x,u,\lambda^d,t) + \mu g(x,u.t) + \eta^d h(x,t).
\end{aligned}
$$

- The Hamiltonian maximizing conditions

$$
\begin{aligned}
& H^d(x^*(t),u^*(t),\lambda^d(t),t) \geq H^d(x^*(t),u,\lambda^d(t),t), \\
& \quad at\ each\ t \in [0,T]\ for\ all\ u\ satisfying\ g(x^*(t),u,t) \geq 0.
\end{aligned}
$$

- Jumping conditions at any time τ where λ^d is discontinuous

$$
\begin{aligned}
\lambda^d(\tau^-) =& \lambda^d(\tau^+) + \zeta^d(\tau)h_x(x^*(\tau),\tau), \\
H^d(x^*(\tau),u^*(\tau^-),\lambda^d(\tau^-).\tau) =& H^d(x^*(\tau),u^*(\tau^+),\lambda^d(\tau^+).\tau) - \zeta^d(\tau)h_t(x^*(\tau),\tau)
\end{aligned}
$$

- Lagrange multipliers have to satisfy

$$
\begin{aligned}
& \frac{\partial L}{\partial u}|_{u=u^*} = 0, \frac{dH^d}{dt} = \frac{dL^d}{dt} = \frac{\partial L^d}{\partial t} \\
& \mu \geq 0, and\ \mu g(x^*,u^*,t) = 0, \\
& \eta^d \geq 0, and\ \eta^d h(x^*,t) = 0, \\
& \zeta^d(\tau) \geq 0, and\ \zeta^d(\tau)h(x^*(\tau),\tau) = 0.
\end{aligned}
$$

Once the terminal time T is undecided, the terminal time $T = T^*$ has to satisfy

$$
H^d(x^*(T^*),u^*(T^*),\lambda^d(T^*),T^*) + S_T(x^*(T^*),T^*) + \alpha a_T(x^*(T^*),T^*) + \beta b_T(x^*(T^*),T^*) + \gamma^d h_T(x^*(T^*),T^*) = 0.
$$

3.3 Sufficiency Conditions: Direct Method

Theorem 3. *Let $(x^*,u^*,\lambda^d,\mu,\alpha,\beta,\gamma^d,\eta^d)$ and the jump parameters $\zeta^d(\tau)$ at each τ, where λ^d is discontinuous, satisfy the necessary conditions for all $t \in [0,T]$. If $H^{0d}[x,\lambda^d(t),t]$ is concave in x for each t and $S(x,T)$ is concave in x, g is quasiconcave in (x,u), h and a are quasiconcave in x, and b is linear in x, then (x^*,u^*) is optimal.*

3.4 Necessary Conditions: Indirect Method

Assumption 1. *h(x,t) is a constraint of first order if the first time derivative of h is the first time a term in control u appears in the expression by putting f(x,u,t) for $\dot{x}$ after each differentiation.*

Indirect method makes an assumption that all $h(x,t)$ in problem setting satisfies the above assumption. In this way, the first derivative of $h(x,t)$ is added to the Lagrange function instead of $h(x,t)$ in direct method. The necessary conditions of $u^*(t)$ being optimal control and $x^*(t)$ being optimal path are separated as below:

- Basic dynamics of the optimal path x^* satisfying mixed inequality and pure state constraints

$$\begin{aligned}&\dot{x}^* = f(x^*, u^*, t),\ x^*(0) = x_0,\\ &g(x^*, u^*, t) \geq 0, h(x^*, t) \geq 0,\\ &a(x^*(T), T) \geq 0, b(x^*(T), T) = 0.\end{aligned}$$

- Dynamics of the shadow price λ and the terminal conditions

$$\begin{aligned}&\dot{\lambda} = -L_x(x^*, u^*, \lambda, \mu, \eta, t),\\ &\lambda(T^-) = S_x(x^*(T), T) + \alpha a_x(x^*(T), T) + \beta b_x(x^*(T), T) + \gamma h_x(x^*(T), T),\\ &\quad where\ \alpha \geq 0, \alpha a(x^*(T), T) = 0, \gamma \geq 0, \gamma h(x^*(T), T) = 0\end{aligned}$$

  where
$$\begin{aligned}&H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t),\\ &L(x, u, \lambda, \mu, \eta, t) = H(x, u, \lambda, t) + \mu g(x, u.t) + \eta h^1(x, u, t),\\ &h^1(x, u, t) = \frac{\partial h(x, t)}{\partial t}\end{aligned}$$

- The Hamiltonian maximizing conditions

$$H^d(x^*(t), u^*(t), \lambda(t), t) \geq H^d(x^*(t), u, \lambda(t), t),$$
$$at\ each\ t \in [0, T]\ for\ all\ u\ satisfying\ g(x^*(t), u, t) \geq 0,\ and\ h^1(x, u, t) \geq 0\ whenever h(x^*(t), t) = 0.$$

- Jumping conditions at any time τ where λ^d is discontinuous

$$\begin{aligned}\lambda(\tau^-) &= \lambda(\tau^+) + \zeta(\tau) h_x(x^*(\tau), \tau),\\ H(x^*(\tau), u^*(\tau^-), \lambda(\tau^-).\tau) &= H(x^*(\tau), u^*(\tau^+), \lambda(\tau^+).\tau) - \zeta(\tau) h_t(x^*(\tau), \tau)\end{aligned}$$

- Lagrange multipliers have to satisfy

$$\begin{aligned}&\frac{\partial L}{\partial u}|_{u=u^*} = 0, \frac{dH^d}{dt} = \frac{dL^d}{dt} = \frac{\partial L^d}{\partial t}\\ &\mu \geq 0, and\ \mu g(x^*, u^*, t) = 0,\\ &\eta \geq 0, and\ \eta h(x^*(t), t) = 0,\\ &\zeta(\tau) \geq 0, and\ \zeta(\tau) h(x^*(\tau), \tau) = 0.\end{aligned}$$

3.5 Relations between Direct Method and Indirect Method

Assume $[\tau_1, \tau_2]$ to be the boundary interval of the optimal path x^* where $h(x^*(t), t) = 0,\ \forall t \in [\tau_1, \tau_2]$. Any span of $[\tau_1, \tau_2]$ is not a boundary interval. Then in the interior of this boundary interval, the direct method has relationship with indirect method as follows:

$$\eta^d(t) = -\ \dot{\eta}(t) \geq 0,\ t \in [\tau_1, \tau_2],$$
$$\lambda^d(t) = \lambda(t) + \eta(t) h_x(x^*(t), t),\ t \in [\tau_1, \tau_2].$$

The jump parameter at an entry time τ_1, an exit time τ_2, or a contact time τ has to satisfy:

$$\zeta^d(\tau_1) = \zeta(\tau_1) - \eta(\tau_1^+),\ \zeta^d(\tau_2) = \eta(\tau_2^-),\ \zeta^d(\tau) = \zeta(\tau).$$

The terminal conditions of the shadow price in these two methods have to satisfy:

$$\gamma^d = \gamma + \eta(T^-).$$

4 The Maximum Principle: Discrete Time

4.1 Preliminary on Nonlinear Programming Problems

- Nonlinear Programming Problems Setting: $\texttt{max}_x\ h(x),\ \texttt{s.t.}\ f(x) = a,\ g(x) \geq b$
- Introduction of Lagrange Multipliers into Lagrange function: $L(x, \lambda, \mu) = h(x) + \lambda[f(x) - a] + \mu[g(x) - b]$
- Constraint Qualification: The following full-rank condition holds at x, if the constraint qualification holds at x.

$$\texttt{rank}\begin{bmatrix} \partial g/\partial x & \texttt{diag}(g) \\ \partial f/\partial x & 0 \end{bmatrix} = \texttt{min}(s + r, s + n),$$

  where $\partial g/\partial x$ and $\partial f/\partial x$ are $s \times n$ and $r \times n$ gradient matrices. In other words, the gradients of the equality constraints and of the active inequality constraints at the candidate point under consideration are linearly independent.

- **Necessary Conditions**: If h, f, and g are differentiable, x^* maximizes the nonlinear programming problem, and the constraint qualification holds at x^*, then there exist multipliers λ and μ such that (x^*, λ, μ) satisfy the $Kuhn - Tucker$ conditions as below:

$$\begin{aligned} L_x(x^*, \lambda, \mu) = h_x(x^*) + \lambda f_x(x^*) + \mu g_x(x^*) =& 0 \\ f(x^*) =& a \\ g(x^*) \geq& b \\ \mu \geq 0, \mu(g(x^*) - b) =& 0 \end{aligned}$$

- **Sufficient Conditions**: If h, f, and g are differentiable, f if affine, g is concave, and (x^*, λ, μ) satisfy the necessary conditions above, then x^* is a solution to the maximization problem.

4.2 Discrete Time with Inequality on Control

- Problem Setting

$$\begin{aligned} &\max_u J = \sum_{k=0}^{T-1} F(x^k, u^k, k)dt + S(x(T), T) \\ &s.t.\ \ \Delta x^k = x^{k+1} - x^k = f(x^k, u^k, k),\ k = 0, \ldots, T-1,\ x^0 = x_0. \\ &\qquad g(u^k, k) \geq b^k,\ k = 0, \ldots, T-1 \end{aligned}$$

- Necessary Conditions: If for every k, $H^k = H(x^k, u^k, k) = F(x^k, u^k, k) + \lambda^{k+1} f(x^k, u^k, k)$ and $g(u^k, k)$ are concave in u^k, and the constraint qualification holds, then the necessary conditions for u^{k*}, $k = 0, 1, \ldots, T-1$, to be an optimal control with the corresponding state x^{k*} are

$$\begin{aligned} &\Delta x^{k*} = f(x^{k*}, u^{k*}, k), x^0\ given, \\ &\Delta \lambda^k = -\frac{\partial H^k}{\partial x^k}(x^{k*}, u^{k*}, \lambda^{k+1}, k), \lambda^T = \frac{\partial S(x^{T*}, T)}{\partial x^{T*}}, \\ &H^k(x^{k*}, u^{k*}, \lambda^{k+1}, k) \geq H^k(x^{k*}, u^k, \lambda^{k+1}, k), \\ &\quad for\ all\ u^k\ such\ that\ g(u^k, k) \geq b^k, k = 0, 1, \ldots, T-1. \end{aligned}$$

4.3 A General Discrete Maximum Principle

Necessary conditions mentioned above are the same with a general discrete problem under some specific assumptions:

$$\begin{aligned} &\max\ J = \sum_{k=0}^{T-1} F(x^k, u^k, k) \\ &\texttt{s.t.}\Delta x^k = f(x^k, u^k, k), x^0\ given \\ &u^k \in \Omega_k, k = 0, 1, \ldots T-1, \end{aligned}$$

where

- $F(x^k, u^k, k)$ and $f(x^k, u^k, k)$ are continuously differentiable in x^k for every u^k and k.
- The sets $\{-F(x, \Omega^k, k), f(x, \Omega^k, k)\}$ are *b-directionally convex* for every x and k, where $b = (-1, 0, \dots, 0)$. That is ,given v and w in Ω^k and $0 \leq \lambda \leq 1$, there exists $u(\lambda) \in \Omega^k$ such that

  $$F(x, u(\lambda), k) \geq \lambda F(x, v, k) + (1 - \lambda)F(x, w, k)$$

  and

  $$f(x, u(\lambda), k) \geq \lambda f(x, v, k) + (1 - \lambda)f(x, w, k)$$

  for every x and k.It should be noted that convexity implies b-directional convexity, but not the converse.
- Ω^k satisfies the Kuhn-Tucker constraint qualification.

5 Stochastic Optimal Control

5.1 Problem Setting

$$\begin{aligned}
&\max_U J = \mathbb{E}[\int_0^T F(X_t, U_t, t)dt + S(X_T, T)], \\
&s.t. \;\; dX_t = f(X_t, U_t, t)dt + G(X_t, U_t, t)dZ_t, \; X_0 = x_0, \\
&\qquad where \; Z_t, t \in [0, T] \; is \; a \; standard \; Wiener \; process.
\end{aligned}$$

5.2 Methods of Hamilton-Jacobi-Bellman Equations

Let $V(x, t)$ be the value function, be the expected value of the objective function from t to T, when an optimal policy is followed from t to T, given $X_t = x$.

$$V(x, t) = \max_U \; \mathbb{E}[F(x, U, t)dt + V(x + dX_t, t + dt)].$$

By Taylor's expansion, we have

$$\begin{aligned}
V(x + dX_t, t + dt) = V(x, t) + V_x dX_t + V_t dt + \frac{1}{2}V_{xx}(dX_t)^2 \\
+ \frac{1}{2}V_{tt}(dt)^2 + \frac{1}{2}V_{xt}dX_t dt + \texttt{higher-order terms}.
\end{aligned}$$

We can formally write some differential terms as

$$\begin{aligned}
(dX_t)^2 = f^2(dt)^2 + G^2(dZ_t)^2 + 2fGdZ_t dt, \\
dX_t dt = f(dt)^2 + GdZ_t dt.
\end{aligned}$$

By the rules of stochastic calculus: $(dZ_t)^2 = dt, \; dZ_t dt = 0, \; dt^2 = 0$, we can rewrite the $V(x, t)$ as

$$V = \max_U [Fdt + V + V_x f dt + V_t dt + \frac{1}{2}V_{xx}G^2 dt + o(dt)],$$

which is equivalent to a partial differential equation of $V(x, t)$

$$0 = \max_U [F + V_x f + V_t + \frac{1}{2}V_{xx}G^2],$$

and the boundary condition $V(x, T) = S(x, T)$. Through solving $V(x, t)$, it is trivial to obtain corresponding optimal control through maximization at each t.

6 Differential Games

The theory of differential games is actually an extension of optimal control theory. In the setting of differential games, there may be more than one decision maker, each having one's own objective function that each is trying to maximize.

6.1 Two-Person Zero-Sum Differential Games

- There are two players 1 and 2, whose controls are respectively u and v.
- The dynamics of the environment state is $\dot{x} = f(x, u, v, t), x(0) = x_0$.
- We assume $u(t) \in U, v(t) \in V, \forall t \in [0, T]$, where U and V are convex sets.
- Consider the objective function $J(u, v) = S[x(T)] + \int_0^T F(x, u, v, t)dt$, which player 1 wants to maximize and player 2 wants to minimize. That's the reason why the game is zero-sum.
- The solution (u^*, v^*) is known as the minimax solution if $J(u^*, v) \geq J(u^*, v^*) \geq J(u, v^*), \forall u \in U, \forall v \in V$.
- The necessary for (u^*, v^*) are given by an extension of the maximum principle as below:

$$H = F + \lambda f$$
$$\dot{\lambda} = -H_x, \lambda(T) = S_x(x(T))$$
$$H(x^*, u^*, v, \lambda, t) \geq H(x^*, u^*, v^*, \lambda, t) \geq H(x^*, u, v^*, \lambda, t), \forall u \in U, \forall v \in V.$$

- The saddle point conditions of H can be reduced as below if u, v are unconstrained:

$$H_u = 0 \; and \; H_v = 0$$
$$H_{uu} \leq 0 \; and \; H_{vv} \geq 0$$

6.2 Nash Differential Games

- **Problem Setting**

$$\dot{x} = f(x, u^1, u^2, \ldots, u^N, t),$$
$$J^i = s^i(x(T)) + \int_0^T F^i(x, u^1, u^2, \ldots, u^N, t)dt$$
$$\texttt{Agent i with control } u^i \texttt{ tends to maximize } J^i$$

- **Nash Solutions** $\{u^{1*}, u^{2*}, \ldots, u^{N*}\}$ has the property that

$$J^i(u^{1*}, u^{2*}, \ldots, u^{N*}) = \max_{u^i \in U_i} J^i(u^{1*}, u^{2*}, \ldots, u^{(i-1)*}, u^i, u^{(i+1)*}, \ldots, u^{N*})$$

  for $i = 1, 2, \ldots, N$.

6.3 Open-Loop Nash Equations

- The open-loop Nash solution is defined when the set of trajectories are given as functions of time. In this way, the maximum principle is given as follows:

$$H^i(x, u^1, u^2, \ldots, u^N, \lambda^i) = F^i + \lambda^i f,$$
$$\dot{\lambda}^i = -H_x^i, \lambda^i(T) = S_x^i(x(T)),$$
$$H^i(u^{1*}, u^{2*}, \ldots, u^{N*}) \geq H^i(u^{1*}, u^{2*}, \ldots, u^{(i-1)*}, u^i, u^{(i+1)*}, \ldots, u^{N*}), \forall u^i \in U_i$$

6.4 Feedback Nash Solution

- The feedback solution is obtained when the set of trajectories are functions of both time and the current state of the system. In this way, the maximum principle is given as follows:

$$H^i(x, u^1, u^2, \ldots, u^N, \lambda^i) = F^i + \lambda^i f,$$
$$\dot{\lambda}^i = -H_x^i - \sum_{j=1}^{N} H_{u^j}^i u_x^j = -H_x^i - \sum_{j=1, j \neq i}^{N} H_{u^j}^i u_x^j, \; \lambda^i(T) = S_x^i(x(T)),$$
$$H^i(u^{1*}, u^{2*}, \ldots, u^{N*}) \geq H^i(u^{1*}, u^{2*}, \ldots, u^{(i-1)*}, u^i, u^{(i+1)*}, \ldots, u^{N*}), \forall u^i \in U_i$$