OPTIMAL CONTROL VIA DYNAMIC PROGRAMMING

Deterministic approach

Andrea Scalenghe

Tesi magistrale

Introduction Optimal Control Problems

Notably, a dynamical system (physical, social, biological, etc.) can be described by its derivatives. Control theory assumes the derivative to be influenced by external factors, by controls.

It aims at finding the ”best” system behavior under a control. Optimality will be in the form of minimizing a cost function.

Control Problem Formulation

Consider a finite interval $I = [t, t_1] \subset \mathbb{R}$ as the operating time of a system, where at any time $s \in I$, the system is described by $x(s) \in O \subseteq \mathbb{R}^m$ and controlled by $u(s) \in U \subseteq \mathbb{R}^n$, known as control and the control space. The system is defined by:

$$\begin{cases} \dot{x}(s) = f(s, x(s), u(s)), & s \in I \\ x(t) = x \end{cases}$$

With f satisfying the Lipschitz condition to ensure a unique solution. Controls $u(\cdot)$ are IN $L^\infty([t, t_1]; U)$.

Optimality

Optimality is measured by a payoff J, incorporating continuoud running cost L and terminal cost Ψ, defined as:

$$J(t,x;u) = \int_t^{\tau} L(s,x(s),u(s))\,ds + \Psi(\tau,x(\tau)),$$

where τ is the exit time of $(s,x(s))$ from $Q = [t,t_1] \times O$. The terminal cost Ψ has the form:

$$\Psi(t,x) = \begin{cases} g(t,x) & \text{if } (t,x) \in [t,t_1) \times O \\ \psi(x) & \text{if } (t,x) \in \{t_1\} \times O \end{cases}$$

Optimal control theory aims at minimizing J across admissible controls.

Dynamic Programming

The value function V encapsulates the essence of dynamic programming:

$$V(t,x) = \inf_{u \in \mathcal{U}(t,x)} J(t,x;u).$$

This framework enables the subdivision of the control problem into smaller, more manageable problems, leading to the Hamilton-Jacobi-Bellman (HJB) equation as a crucial tool for finding V.

Dynamic Programming Principle

The dynamic programming principle enables the just cited subdivision. Indeed the following holds true.

Theorem

For any $(t, x) \in \overline{Q}$ and any $r \in I$ then:

$$V(t,x) = \inf_{u \in \mathcal{U}(t,x)} \left\{ \int_t^{r \wedge \tau} L(s, x(s), u(s))\, ds + g(\tau, x(\tau))\chi_{\tau < r} + V(r, x(r))\chi_{r \leq \tau} \right\} \tag{1}$$

Dynamic Programming Equation

From the dynamic programming principle, we get a necessary condition for a value function to be optimal. Formally, we take $h > 0$ and rewrite 1 as:

$$\inf_{u \in \mathcal{U}} \{ \frac{1}{h} \int_t^{(t+h)\wedge\tau} L(s, x(s), u(s))\, ds + \frac{1}{h} g(\tau, x(\tau)) \chi_{\tau < t+h} \\ + \frac{1}{h} \left[V(t+h, x(t+h)) \chi_{\tau \geq t+h} - V(t, x) \right] \} = 0.$$

Then letting $h \to 0^+$:

$$\inf_{u \in \mathcal{U}} \{ L(t, x(t), u(t)) + \partial_t V(t, x(t)) + D_x V(t, x(t)) \cdot f(t, x(t), u(t)) \} = 0 \quad (2)$$

Dynamic Programming Equation

We can generally define:

$$-\frac{\partial}{\partial t}V(t,x) + H(t,x,D_xV(t,x)) = 0 \tag{3}$$

Where for $(t,x,p) \in \overline{Q} \times \mathbb{R}^n$ the Hamiltonian is defined as:

$$H(t,x,p) = \sup_{v \in \mathbb{R}^n} \left\{-p \cdot f(t,x,v) - L(t,x,v)\right\}. \tag{4}$$

Equation 3 turns out to be the main sufficient condition for the value function to be optimal.

Theorem (Verification Theorem)
Let $W \in C^1(\overline{Q})$ satisfy 3 and the boundary conditions then:

$$W(t,x) \leq V(t,x) \ \forall \ (t,x) \in \overline{Q}$$

Moreover, there exists $u^* \in \mathcal{U}$ such that:

$$\begin{cases} L(s, x^*(s), u^*(s)) + f(s, x^*, u^*(s)) \cdot D_x W(s, x^*(s)) = -H(s, x^*(s), D_x W(s, x^*(s)) \\ W(\tau^*, x^*(\tau^*)) = g(\tau^*, x^*(\tau^*)) \end{cases} \tag{5}$$

if and only if u^* is optimal and $W = V$.

Pontryagin's Principle

Porntryagin's principle gives another perspective on the problem. It asserts the existence of a *costate* variable that satisfies certain, similar, conditions under optimality of the value function.

The *control state Hamiltonian* is defined as:

$$H(s,x,u,p) = -p \cdot f(s,x,u) - L(s,x,u),$$

with p representing the system's *costate*.

Pontryagin's Principle

Theorem

Let u^* be an optimal control and x^* its corresponding trajectory. Then there exists a function $p^* : [t, t_1] \to O$ such that:

$$\dot{x}^*(s) = D_p H(s, x^*(s), u^*(s), p^*(s)) \tag{6}$$

$$\dot{p}^*(s) = -D_x H(s, x^*(s), u^*(s), p^*(s)) \tag{7}$$

And also:

$$H(s, x^*(s), u^*(s), p^*(s)) = \sup_{v \in U} H(s, x^*(s), v, p^*(s)) \tag{8}$$

With:

$$p^*(t_1) = D\psi(x^*(t_1)) \tag{9}$$

Proof Idea

General form can be seen as no running cost ($L \equiv 0$), it is proved in this setting. The costate is set to be the solution of the Adjoint dynamic 7 with transversality condition 9. To show maximality 6 we set $v \in U$ and take small "variation" u_ϵ of the optimal control such that $u_\epsilon(r) = v$ for some $r > 0$. Then:

$$
\begin{aligned}
0 &\leq \frac{d}{d\epsilon} J(t, x; u_\epsilon)\Big|_{\epsilon=0} = \frac{d}{d\epsilon}\psi(x_\epsilon(t_1))\Big|_{\epsilon=0} && (10)\\
&= \frac{d}{d\epsilon}\psi(x(t_1) + \epsilon y(t_1) + o(\epsilon)) = D\psi(x(t_1)) \cdot y(t_1) && (11)\\
&= p(t_1) \cdot y(t_1) = p(r) \cdot [f(r, x(r), v) - f(r, x(r), u(r))] && (12)
\end{aligned}
$$

where y is a suitable path, governed by the adjoint dynamic.

Connection to Dynamic Programming

Pontryagin's Principle and dynamic programming, though seemingly different, are closely linked.

Theorem
Let u^* be an optimal right-continuous control and x^* its corresponding trajectory. Assume that the value function V is differentiable at $(s, x^*(s))$ for $s \in [t, t_1)$. If we define:

$$p(s) = D_x V(s, x^*(s)) \tag{13}$$

Then $p(s)$ satisfies 7, 8 and 9.

Existence Theorem for Optimal Controls

We now prove an existence theorem for optimal controls. We study the fixed time interval case with $O = \mathbb{R}^n$ and the function f linear in v. Furthermore, we impose convexity of L in v. Under these assumptions a classical variational argument proves the optimal control existence.

Theorem

Let U compact and convex, $f_1, f_2 \in C^1(\overline{Q} \times U)$ such that $f(t, x, v) = f_1(t, x) + f_2(t, x)v$ and $\partial_x f_1, \partial_x f_2, f_2$ bounded. Let also $L \in C^1(\overline{Q} \times U)$ and $L(t, x, \cdot)$ be convex for all $(t, x) \in \overline{Q}$ and the temrinal cost $\phi \in C(\mathbb{R}^n)$. Then there exist an optimal control $u^*(\cdot)$.

Proof idea

Proof Strategy:
- Start with a minimizing sequence u_n such that:

$$\lim J(t, x; u_n) = V(t, x).$$

- Show weak convergence of u_n and uniform convergence (U compact and convex) of the state trajectories x_n (via Ascoli-Arzela).
- Show x^* to be solution of the control problem defined by f and u^*.
- Show u^* to be an optimal control.