Pontryagin's Maximum Principle

an introduction

Boualem Djehiche
KTH, Stockholm

September 28, 2013

Deterministic optimal control problem:

Minimize or maximize

$$J(u) = \int_0^T f(x(t), u(t)) dt + h(x(T)), \qquad (1)$$

with respect to $u : [0, T] \longrightarrow U$, subject to

$$\begin{cases} \dot{x}(t) = b(x(t), u(t)), & 0 < t \leq T, \\ x(0) = x_0. \end{cases} \qquad (2)$$

where, U is a given set of controls.

Stochastic optimal control problem:

Minimize or maximize

$$J(u) = E\left[\int_0^T f(x(t), u(t))dt + h(x(T))\right], \qquad (3)$$

with respect to $u : [0, T] \longrightarrow U$, subject to

$$\begin{cases} dx(t) = b(x(t), u(t))dt + \sigma(x(t), u(t))dW(t), \\ x(0) = x_0. \end{cases} \qquad (4)$$

Stochastic optimal control problem for systems of mean-field type

Minimize or maximize

$$J(u) = E\left[\int_0^T f(x(t), E[x(t)], u(t))dt + h(x(T), E[x(T)])\right], \quad (5)$$

with respect to $u : [0, T] \longrightarrow U$, subject to

$$\begin{cases} dx(t) = b(x(t), E[x(t)], u(t))dt + \sigma(x(t), E[x(t)], u(t))dW(t), \\ x(0) = x_0. \end{cases}$$
$$(6)$$

Example:

$$J(u) = Var(x(T)) = E[x^2(T)] - (E[x(T)])^2.$$

Optimal control theory tries to answer two questions:

- Existence of a minimum/maximum of the performance functional J.
- Explicit computation/characterization of such a minimum/maximum.
    - The Bellman principle which yields the Hamilton-Jacobi-Bellman equation (HJB) for the value function;
    - Pontryagin's maximum principle which yields the Hamiltonian system for "the derivative" of the value function.

Features of the Bellman principle and the HJB equation

- ▶ The Bellman principle is based on the "law of iterated conditional expectations".
    - ▶ It does not apply for dynamics of mean-filed type:

$$J(u) = E\left[\int_0^T f(x(t), E[x(t)], u(t))dt + h(x(T), E[x(T)])\right].$$

- ▶ The HJB equation is a nonlinear PDE, valid only for "Markovian systems", where the coefficients b, σ, h and f are deterministic functions of (t, x).
- ▶ An eventual optimal control should be of "feedback form": $\bar{u}(t) = v(t, X_t)$, where v is a deterministic function of (t, x).

Features of the Pontryagin's maximum principle

- ▶ Pontryagin's principle is based on a "perturbation technique" for the control process, that does not put "structural" restrictions on the dynamics of the controlled system.
- ▶ It seems well suited for
  - ▶ Non-Markovian systems. i.e. where the coefficients b, σ, h and f are random and not necessarily deterministic functions of (t, x).
  - ▶ Systems of mean-field type.

A heuristic derivation of the Pontryagin's principle

Minimize

$$J(u) = \int_0^T f(x(t), u(t))dt + h(x(T)), \qquad (7)$$

with respect to $u : [0, T] \longrightarrow U$, subject to

$$\begin{cases} \dot{x}(t) = b(x(t), u(t)), & 0 < t \leq T, \\ x(0) = x_0. \end{cases} \qquad (8)$$

Using the Lagrangian multiplier method, we minimize the Lagrangian functional

$$\mathcal{L}(x, u, p) := \int_0^T \left(f(x(t), u(t)) - p(t) \cdot (\dot{x}(t) - b(x(t), u(t)) \, dt \right. \\ + h(x(T)), \qquad (9)$$

Introducing the so-called Hamiltonian

$$H(x, u, p) := p \cdot b(x, u) + f(x, u), \qquad (10)$$

the Lagrangian functional becomes

$$\mathcal{L}(x, u, p) := \int_0^T \left(H(x(t), u(t), p(t)) - p(t) \cdot \dot{x}(t) \right) dt + h(x(T)).$$

If $(\bar{x}, \bar{u}, p)$ is a minimizer of $\mathcal{L}$, performing a first-order Taylor expansion with perturbation $(\delta x, \delta u, \delta p)$, such that $\delta x(0) = 0$:

$$
\begin{aligned}
\delta \mathcal{L} \quad &:= \mathcal{L}(\bar{x} + \delta x, \bar{u} + \delta u, p + \delta p) - \mathcal{L}(\bar{x}, \bar{u}, p) \\
&\approx \int_0^T \left(H_x \cdot \delta x + H_u \cdot \delta u + H_p \cdot \delta p - p \cdot \tfrac{d}{dt}(\delta x) - \delta p \cdot \dot{\bar{x}} \right) dt \\
&\quad + h_x \cdot \delta x(T),
\end{aligned}
$$

Integrating by parts, we obtain

$$
\begin{aligned}
\delta \mathcal{L} &\approx \int_0^T \left((H_x + \dot{p}) \cdot \delta x + H_u \cdot \delta u + \left(H_p - \dot{\bar{x}} \right) \cdot \delta p - \tfrac{d}{dt}(p \cdot \delta x) \right) dt \\
&\quad + h_x \cdot \delta x(T).
\end{aligned}
$$

Requiring $\delta\mathcal{L} = 0$ and noting that $\delta x(0) = 0$, yields

$$0 = \int_0^T \left((H_x + \dot{p}) \cdot \delta x + H_u \cdot \delta u + \left(H_p - \dot{x}\right) \cdot \delta p \right) dt$$
$$+ (h_x - p(T)) \cdot \delta x(T).$$

This in turn suggests that the extremal $(\bar{x}, \bar{u}, p)$ satisfies the so-called *Hamiltonian system* associated with the optimal control problem (7)-(8):

$$\begin{cases} \bar{x}(0) = x_0, & \text{(Initial value)} \\ \dot{\bar{x}}(t) = H_p(\bar{x}(t), p(t), \bar{u}(t)), & \text{(Controlled dynamics)} \\ \dot{p}(t) = -H_x(\bar{x}(t), p(t), \bar{u}(t)), & \text{(Adjoint equation)} \\ p(T) = h_x(\bar{x}(T)), & \text{(Adjoint terminal value)} \\ \bar{u}(t) = \arg\min_{u \in U} H(\bar{x}(t), u, p(t)), & \text{(Maximum Principle)} \end{cases}$$

The adjoint equation

$$\begin{cases} \dot{p}(t) = -H_x(\bar{x}(t), p(t), \bar{u}(t)), & 0 \le t < T, \\ p(T) = h_x(\bar{x}(T)). \end{cases} \tag{11}$$

Pontryagin's Maximum Principle. If $(\bar{x}, \bar{u})$ is an optimal solution of the control problem (7)-(8), then there exists a function p solution of the adjoint equation (11) and for which

$$\bar{u}(t) = \arg\min_{u \in U} H(\bar{x}(t), u, p(t)), \quad 0 \le t \le T. \quad \text{(Maximum Principle)}$$

- ▶ This perturbation method, implicitly assumes that the set U of control is *linear* i.e. if $u \in U$ then also $u + \delta u \in U$.
- ▶ But, in many practical examples, the set/space U is rather convex or a general metric space.
- ▶ Device a new perturbation method which should be compatible with the structure of the set of controls.

If the set of controls U is convex, then a natural perturbation method would be a convex perturbation of a given optimal control $\bar{u}$:

$$\delta u := \epsilon(u - \bar{u}), \quad u \in U,$$

where, ϵ will tend to 0. Hence, if $\bar{u}$ is optimal for J, we have

$$J(\bar{u} + \epsilon(u - \bar{u})) - J(\bar{u}) \geq 0,$$

and thus, we can write

$$(J'(\bar{u}), u - \bar{u}) \geq 0, \qquad \forall u \in U,$$

provided that J is Gateaux differentiable.

A local maximum principle

Mimicking the previous formal computations we would expect the following (local) version of the maximum principle.

$$\frac{dH}{du}(\bar{x}(t), \bar{u}(t), p(t)) \cdot (u - \bar{u}(t)) \geq 0, \qquad 0 \leq t \leq T, \quad \forall u \in U.$$

If the set U is not necessarily convex, but a *general separable metric space*,

Pontryagin's approach suggests the following perturbation method called *spike variation* would be appropriate.

For $\epsilon > 0$, pick a subset $E_\epsilon \subset [0, T]$ such that $|E_\epsilon| = \epsilon$. The control process u^ϵ is a spike variation of u if

$$u^\epsilon(t) := \begin{cases} u(t), & t \in E_\epsilon, \\ \bar{u}(t), & t \in E_\epsilon^c, \end{cases} \tag{12}$$

where, $u \in U$ is an arbitrary control.

Using the spike variation technique, Pontryagin's main contribution is the following key relation between the performance functional J and the associated Hamiltonian H:

$$J(u^{\epsilon}) - J(\bar{u}) = \int_0^T \delta H(s) I_{E_{\epsilon}}(s)\, ds + \circ(\epsilon), \qquad (13)$$

for arbitrary $u \in U$, where,

$$\delta H(s) = H(\bar{x}(s), u(s), p(s)) - H(\bar{x}(s), \bar{u}(s), p(s)).$$

We can choose it of the form $E_\epsilon := [\bar{t}, \bar{t} + \epsilon]$, for arbitrarily chosen $\bar{t} \in [0, T]$. This yields

$$0 \leq J(u^\epsilon) - J(\bar{u}) = \int_{\bar{t}}^{\bar{t}+\epsilon} \delta H(t) \, dt + \circ(\epsilon).$$

Dividing by ϵ and then sending ϵ to zero together with the separability of U we obtain

$$\bar{u}(t) = \arg\min_{u \in U} H(\bar{x}(t), u, p(t)), \quad 0 \leq t \leq T. \quad \text{(Maximum Principle)}$$

Relation to the Hamilton-Jacobi-Bellman's equation

The Hamilton-Jacobi-Bellman equation is a nonlinear backward PDE

$$\begin{cases} \frac{\partial V}{\partial t}(t,x) + \min_u H(x, V_x(t,x), u) = 0, \\ V(x, T) = h(x). \end{cases} \quad (14)$$

satisfied by the value function $V(x,t)$ defined by

$$V(t,x) = \min_{u:x(t)=x} \int_t^T f(x(s), u(s))dt + h(x(T)) \quad (15)$$

The associated optimal control is of *feedback* type:

$$u^*(t,x) = \arg\min_u H(x, V_x(t,x), u),$$

which depends on $\nabla_x V(t,x)$ but not on V.

An informal comparison between the adjoint equation and the HJB equation suggests that

$$p(t) = V_x(t, x),$$

provided that the HJB equation admits a smooth solution V. Unfortunately, smooth solutions of HJB equations are obtained in very few cases. Most of real world problems exhibit non-smooth solutions of the HJB equation.

References

Yong, J. and Zhou, XY (1999): Stochastic Controls. *Springer*.

Papers by BD with co-authors.