Optimal Control Theory
Bellman's Dynamic Programming

Alessandra Buratto

.

DIPARTIMENTO MATEMATICA

Università degli Studi di Padova

Bellman's Dynamic Programming

let $x(\cdot; u)$ be the state function
solution to the Cauchy problem determined by control u

notation:
let us further use y for the state variable
and x for specific real values of state

$$
\begin{aligned}
\text{maximize} \quad J(u) &= \int_{t_0}^{t_1} f_0(y(t;u), u(t), t)\, dt + S(y(t_1;u)) \\
\text{subject to} \quad \dot{y}(t) &= f(y(t), u(t), t) \\
y(t_0) &= x^0 \\
u(t) &\in \Omega
\end{aligned}
$$

Principle of Optimality

An optimal policy has the property that,
whatever the initial state and initial decisions are,
the remaining decisions must constitute an optimal policy
with regard to the state resulting from the first decision.

R. Bellman (1957)

Bellman's Principle of Optimality

for all $t \in [t_0, t_1]$ and $x \in \mathbb{R}$ consider problem

$$\text{maximize} \quad \int_{\boxed{t}}^{t_1} f_0 \left(y \left(s; u \right), u \left(s \right), s \right) ds + S \left(y \left(t_1; u \right) \right)$$

$$\text{subject to} \quad \begin{aligned} \dot{y} \left(s \right) &= f \left(y \left(s \right), u \left(s \right), s \right) \\ y \left(\boxed{t} \right) &= \boxed{x} \\ u \left(s \right) &\in \Omega \end{aligned}$$

call it problem $P_{t,x}$

original problem: P_{t_0,x^0} embedded in $\left\{ P_{t,x} \middle| \, t \in [t_0, t_1], \; x \in \mathbb{R} \right\}$

interpretation of problem $P_{t,x}$

. . . you have fallen asleep at time t_0.

Suddenly you wake up, noticing that the time now is t
and that your state process while you were asleep
has moved to the point x.

You now try to do as well as possible under the circumstances,
so you want to maximize your utility over the remaining time, given the fact that you start at time t in the state x.

Value function

Definition

If $P_{t,x}$ has an optimal solution, then let $V(t,x)$ *be the optimal value*

$$V(t,x) = \max_{u(\cdot)} \left\{ \int_t^{t_1} f_0\left(y(s;u), u(s), s\right) ds + S\left(y(t_1;u)\right) \right\}$$

where $y(s;u)$ is the unique solution to

$$\begin{cases} \dot{y}(s) = f\left(y(s), u(s), s\right) \\ y(t) = x \end{cases}$$

associated with the control $u(\cdot)$

we call $V(t,x)$ the *value function*

Value function

notice that

$$V(t_1, x) = S(x)$$

in fact

$$V(t_1, x) = \max_{u(\cdot)} \left\{ \int_{t_1}^{t_1} f_0(y(s;u), u(s), s)\, ds + S(y(t_1;u)) \right\}$$

$$= S(y(t_1;u)) = S(x)$$

Optimal control problem: Hamilton Jacobi Bellman

Theorem (Optimality Principle)

Let the value function $V(t, x)$ be
defined and continuously differentiable
at all $t \in [t_0, t_1]$, $x \in \mathbb{R}$

then it is a solution of the partial differential equation

$$-\frac{\partial V(t, x)}{\partial t} = \max_{\alpha \in \Omega} \left\{ f_0(x, \alpha, t) + \frac{\partial V(t, x)}{\partial x} f(x, \alpha, t) \right\}$$

the *Hamilton-Jacobi-Bellman equation*

Optimal control problem HJB

if $P_{t,x}$ has an optimal solution, for all $t \in [t_0, t_1]$, $x \in \mathbb{R}$
if $V(t,x)$ is continuously differentiable
then $V(t,x)$ satisfies the PDE

$$\frac{\partial V(t,x)}{\partial t} + \max_{\alpha \in \Omega} \left\{ f_0(x,\alpha,t) + \frac{\partial V(t,x)}{\partial x} f(x,\alpha,t) \right\} = 0$$

$$(t,x) \in [t_0, t_1) \times \mathbb{R}$$

and the final time condition

$$V(t_1, x) = S(x) \qquad x \in \mathbb{R}$$

Feedback control strategies

strategic idea:

look at the system at any time $t \in [t_0, t_1]$
observe its state $x(t)$
then choose the control strategy $u(t)$ as a function of the state at the same time (and possibly the time)

$$u(t) = \varphi(t, x(t))$$

Definition

A function $\varphi : [t_0,\ t_1] \times \mathbb{R} \to \mathbb{R}$ *is called an admissible feedback function for the control problem*
if, for all $(t_0,\ x^0)$ *initial data, the Cauchy problem*

$$\begin{cases} \dot{x}(t) = f(t, x(t), \varphi(t, x(t)))\,, & t \in [t_0,\ t_1] \\ x(t_0) = x^0 \end{cases}$$

has a unique solution – denote it by $x(\cdot; t_0,\ x^0, \varphi)$
and the corresponding control function

$$u(t) = \varphi(t, x(t; t_0,\ x^0, \varphi))\,, \qquad t \in [t_0,\ t_1],$$

is admissible (continuous and takes values in Ω*)*

Theorem (Verification)

Let $V(t,x)$ be a continuously differentiable function
let $V(t,x)$ satisfy HJB equation and terminal condition;
let there exist a continuous feedback control function

$$\begin{array}{rccc} \varphi: & [t_0,t_1]\times\mathbb{R} & \to & K \\ & (t,x) & \mapsto & \varphi(t,x) \end{array}$$

such that

$$\frac{\partial V(t,x)}{\partial t} + f_0(x,\varphi(t,x),t) + \frac{\partial V(t,x)}{\partial x}\, f(x,\varphi(t,x),t) \;=\; 0$$

then

$$u^*(t) = \varphi(t, y(t;u^*))$$

is an optimal (feedback) strategy
moreover $V(t,x)$ is the optimal value of problem $P_{t,x}$

remark: stating that

$$\frac{\partial V(t,x)}{\partial t} + f_0(x,\varphi(t,x),t) + \frac{\partial V(t,x)}{\partial x}\ f(x,\varphi(t,x),t) \ = \ 0$$

equivalent to

$$\max_{\alpha\in\Omega}\left\{ f_0(x,\alpha,t) + \frac{\partial V(t,x)}{\partial x}\ f(x,\alpha,t)\right\} \ =$$

$$= \ f_0(x,\varphi(t,x),t) + \frac{\partial V(t,x)}{\partial x}\ f(x,\varphi(t,x),t)$$

Optimal control problem - discount factor

$$\text{maximize} \quad \int_{t_0}^{t_1} e^{-\rho t} F_0\left(x(t), u\left(t\right), t\right) dt + e^{-\rho\left(t_1 - t_0\right)} S\left(y\left(t_1; u\right)\right)$$

$$\begin{aligned} \text{subject to} \quad \dot{x}\left(t\right) &= f\left(x\left(t\right), u\left(t\right), t\right) \\ y\left(t_0\right) &= x \\ u\left(t\right) &\in \Omega \end{aligned}$$

Looking for feedback strategies $\varphi\left(t, y\left(t; u^*\right)\right)$

Value function $V\left(t, x\right)$

Optimal control problem HJB equation - discount factor

HJB equation:

$$\rho V(t,x) - \frac{\partial V(t,x)}{\partial t} = \max_{\alpha \in \Omega} \left\{ F_0(x,\alpha,t) + \frac{\partial V(t,x)}{\partial x} f(x,\alpha,t) \right\}$$

$$(t,x) \in [t_0, t_1) \times \mathbb{R}$$

and the final time condition

$$V(t_1, x) = S(x) \qquad x \in \mathbb{R}$$

Optimal control problem infinite horizon

$$
\begin{aligned}
\text{maximize} \quad & \int_{t_0}^{+\infty} e^{-\rho t} F_0(\underline{x}(t), \underline{u}(t), t)\, dt \\
\text{subject to} \quad & \underline{\dot{x}}(t) = \underline{f}(\underline{x}(t), \underline{u}(t), t) \\
& \underline{x}(t_0) = \underline{x}^0 \\
& \underline{u}(t) \in \Omega
\end{aligned}
$$

Looking for steady state solution (not depending on t) $x(t) \equiv x^{SS}$
Looking for stationary feedback strategies (not depending on t) $\varphi\left(x^{SS}\right)$

Value function $V(x)$ not depending on t

Optimal control problem HJB equation - Infinite Horizon

HJB equation:

$$\rho V(t,x) - \overbrace{\frac{\partial V(t,x)}{\partial t}}^{=0} = \max_{\alpha \in \Omega} \left\{ F_0(x,\alpha,t) + \frac{\partial V(t,x)}{\partial x} \, f(x,\alpha,t) \right\}$$

$$(t,x) \in [t_0, t_1) \times \mathbb{R}$$

Sufficient conditions:

$$\lim_{T \to +\infty} \inf e^{-\rho T}(V(x(T),T) - V(x^*(T),T)) \geq 0$$