ENEE 664, ~~05~~. Spring 2004

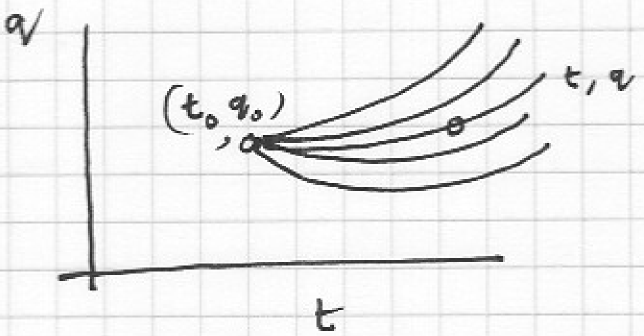
Lecture 13 — Sufficiency & Optimal Control

In classical mechanics, for a system with Lagrangian L, one is interested in the extremals of

$$\int_{t_1}^{t_2} L(t, q(t), \dot{q}(t))\, dt$$

along curves $t \mapsto q(t)$ with $q(t_1) = q_1$ and $q(t_2) = q_2$ fixed.

Extremals satisfy E–L equations. Given (t_0, q_0), and a choice of $\dot{q}_0$, one can integrate E–L to produce an extremal curve (not necessarily passing through specified end-points.), generating a field of extremals. A given extremal γ passing through (t, q) is contained in a central field of extremals if the map $\dot{q}_0 \to q$ is nonsingular.



In that case, one can ~~st.~~ define correctly

$$S_{(t_0,q_0)}(t,q) = \int_{t_0}^{t} L(\sigma, q(\sigma), \dot{q}(\sigma))\,d\sigma$$

where $\sigma \to q(\sigma)$ is an <u>extremal</u>

connecting (t_0, q_0) to (t, q)

We call S the action function (with parameter (t_0, q_0)).

<u>Lemma</u>

$$dS = p\,dq - H\,dt$$

where $p = \frac{\partial L}{\partial \dot{q}}$, $H = p\dot{q} - L$

<u>Remark</u> Proof is an application of Stokes theorem.

Now $dS = \frac{\partial S}{\partial t} dt + \frac{\partial S}{\partial q} dq$.

The lemma says.

$$p = \frac{\partial S}{\partial q}$$

$$\frac{\partial S}{\partial t} + H\left(t, q, \frac{\partial S}{\partial q}\right) = 0$$

This last equation is the Hamilton–Jacobi equation of mechanics. It has a very nice generalization to control theory called the Hamilton–Jacobi–Bellman equation (or simply Bellman equation for short).

We approach this by a discretization of the problem of continuous time optimal control

$$(*) \quad \begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \\ x(0) = x_0 \end{cases}$$

$$\min_{u(\cdot)} \quad \phi(x(T)) + \int_0^T L(t, x(t), u(t))\, dt$$

subject to $(*)$.

This is a free-end point problem with a terminal cost $\phi(x(T))$.

Approximate $(*)$ by the discrete time system

$$x_{k+1} = x_k + \delta \cdot \underset{\delta}{f}(k, x_k, u_k)$$

$$k = 0, 1, 2, \ldots, N-1, \quad \delta = T/N$$

where $x_k \triangleq x(k\delta)$, $u_k \triangleq u(k\delta)$.

The cost is approximated by

$$\phi(x_N) + \sum_{k=0}^{N-1} L(k, x_k, u_k)\, \delta$$

Let $J^*(k, x)$

= optimal cost-to-go at time $k\delta$ and state x

We also call this the value function. Clearly

$$J^*(N, x) = \phi(x)$$ (end state $\Rightarrow$ no decision)

$$J^*(k, x) = \min_v \left\{ L(k, x, v)\delta + J^*(k+1, x + f^{\delta}(k, x, v)) \right\}$$

$k = 0, 1, 2, \ldots, N-1$

These are the equations of dynamic programming (follow from the principle of optimality). We have a functional equation for unknown J^*. Solve it backwards from $k = N$.

Suppose J^* has sufficient differentiability (in the continuous time and space setting).

HERE:

$\nabla_t = \frac{\partial}{\partial t}$

$\nabla_x = \frac{\partial}{\partial x}$

Then one can write,

$$J^*((k+1), x + f(k, x, v)\delta)$$

$$= J^*(k, x) + \nabla_t J^*(k, x) \cdot \delta$$

$$+ \nabla_x J^*(k, x) \cdot f^k(x, v)\delta$$

$$+ o(\delta)$$

$$\Rightarrow \quad J^*(k, x) = \min_v \Big\{ L(k, x, v)\delta + J^*(k, x)$$

$$+ \nabla_t J^*(k, x)\delta$$

$$+ \nabla_x J^*(k, x) \cdot f(k, x, v)\delta$$

$$+ o(\delta)\Big\}$$

Cancelling $J^*(k, x)$ from both sides, dividing by δ and sending $\delta \to 0$ we get

$$0 = \min_v \Big\{ L(t, x, v) + \nabla_t J^*(t, x) + \nabla_x J^*(t, x) \cdot f(t, x, v) \Big\}$$

$\leftrightarrow$

$$\nabla_t J^*(t, x) = - \min_v \Big\{ L(t, x, v) + \nabla_x J^*(t, x) \cdot f(t, x, v) \Big\}$$

This is the Hamilton – Jacobi – Bellman (HJB) eqn or simply the Bellman equation in continuous time.

Sufficiency Theorem.

Suppose, $V(t,x)$ is a solution to

(HJB) $$0 = \min_{v} \left\{ L(t,x,v) + \nabla_t V(t,x) + \nabla_x V(t,x) \cdot f(t,x,v) \right\} \quad \forall\, t, x$$

$$V(T,x) = \phi(x) \qquad \forall\, x .$$

Suppose that $\mu^*(t,x)$ attains the minimum in HJB, $\forall$ t and x. Let $\{ x^*(t) \mid t \in [0,T] \}$ be the state trajectory satisfying

$$\left.\begin{aligned} \dot{x}^*(t) &= f(t, x^*(t), \mu^*(t, x^*(t))) \\ x^*(0) &= x_0 \end{aligned}\right\}$$

Suppose that this closed loop dynamics has a unique solution starting from at any pair (t,x) and that the control trajectory $\mu^*(t, x^*(t))$ is piecewise continuous as a function t. Then V is the unique solution to HJB, equal to the optimal cost-to-go (value) function

$$V(t,x) = J^*(t,x) \qquad \forall\, t, x$$

where $$J^*(t,x) = \min_{u(\cdot)} \left[\int_t^T L(\sigma, x(\sigma), u(\sigma))\, d\sigma + \phi(x(T)) \right]$$

subject to $\dot{x}(\sigma) = f(\sigma, x(\sigma), u(\sigma)) \qquad \sigma \geq t$

$x(t) = x$.

Furthermore $u^*(t) = \mu^*(t, x^*(t))$ is

optimal $0 \leq t \leq T$.

Proof Let $\bar{u}(\cdot)$ be any admissible control trajectory and let $\bar{x}(t)$ $0 \leq t \leq T$ be state corresponding to $\bar{u}(\cdot)$. From HJB

$$0 \leq L(t, \bar{x}(t), \bar{u}(t)) + \nabla_t V(t, \bar{x}(t)) + \nabla_x V(t, \bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t))$$

$$= L + \nabla_t V + \nabla_x V \cdot \dot{\bar{x}}$$

$$= L(t, \bar{x}(t), \bar{u}(t)) + \frac{d}{dt} V(t, \bar{x}(t)) .$$

integrate both sides to get

$$0 \leq \int_0^T L(\sigma, \bar{x}(\sigma), \bar{u}(\sigma))\, d\sigma + V(T, \bar{x}(T)) - V(0, \bar{x}(0))$$

$$= \int_0^T + \phi(\bar{x}(T)) - V(0, x_0)$$

$$\Rightarrow \quad V(0, x_0) \le \phi(\bar{x}(T)) + \int_0^T L(\sigma, \bar{x}(\sigma), \bar{u}(\sigma))\, d\sigma$$

If we $u^*(t)$ and $x^*(t)$ instead of $\bar{u}(t)$ and $\bar{x}(t)$, the preceding inequalities become equalities (starting from HJB).

$$\Rightarrow \quad V(0, x_0) = h(x^*(T)) + \int_0^T L(\sigma, x^*(\sigma), u^*(\sigma))\, d\sigma$$

Thus the cost corresponding to (u^*, x^*) is $\le$ cost corresp. $(\bar{u}, \bar{x})$ for any admissible $\bar{u}$.

$\Rightarrow$ $u^*(t)$ is optimal

and $V(0, x_0) = J^*(0, x_0)$

In the preceding arguments we can replace $(0, x_0)$ by (t, x) with no change.

$$\Rightarrow \quad V(t, x) = J^*(t, x) \qquad \blacksquare$$