6.231 Dynamic Programming and Stochastic Control
Fall 2008

6.231 DYNAMIC PROGRAMMING

LECTURE 5

LECTURE OUTLINE

- Deterministic continuous-time optimal control
- Examples
- Connection with the calculus of variations
- The Hamilton-Jacobi-Bellman equation as a continuous-time limit of the DP algorithm
- The Hamilton-Jacobi-Bellman equation as a sufficient condition
- Examples

PROBLEM FORMULATION

- Continuous-time dynamic system:

$$\dot{x}(t) = f\big(x(t), u(t)\big), \quad 0 \le t \le T, \quad x(0): \text{ given,}$$

where

  - $x(t) \in \Re^n$: state vector at time t
  - $u(t) \in U \subset \Re^m$: control vector at time t
  - U: control constraint set
  - T: terminal time

- Admissible control trajectories $\big\{u(t) \mid t \in [0, T]\big\}$: piecewise continuous functions $\big\{u(t) \mid t \in [0, T]\big\}$ with $u(t) \in U$ for all $t \in [0, T]$; uniquely determine $\big\{x(t) \mid t \in [0, T]\big\}$

- Problem: Find an admissible control trajectory $\big\{u(t) \mid t \in [0, T]\big\}$ and corresponding state trajectory $\big\{x(t) \mid t \in [0, T]\big\}$, that minimizes the cost

$$h\big(x(T)\big) + \int_0^T g\big(x(t), u(t)\big) dt$$

- f, h, g are assumed continuously differentiable

EXAMPLE I

- Motion control: A unit mass moves on a line under the influence of a force u
- $x(t) = \big(x_1(t), x_2(t)\big)$: position and velocity of the mass at time t
- Problem: From a given $\big(x_1(0), x_2(0)\big)$, bring the mass "near" a given final position-velocity pair $(\overline{x}_1, \overline{x}_2)$ at time T in the sense:

$$\text{minimize } \big|x_1(T) - \overline{x}_1\big|^2 + \big|x_2(T) - \overline{x}_2\big|^2$$

subject to the control constraint

$$|u(t)| \leq 1, \qquad \text{for all } t \in [0, T]$$

- The problem fits the framework with

$$\dot{x}_1(t) = x_2(t), \qquad \dot{x}_2(t) = u(t),$$

$$h\big(x(T)\big) = \big|x_1(T) - \overline{x}_1\big|^2 + \big|x_2(T) - \overline{x}_2\big|^2,$$

$$g\big(x(t), u(t)\big) = 0, \qquad \text{for all } t \in [0, T]$$

EXAMPLE II

- A producer with production rate $x(t)$ at time t may allocate a portion $u(t)$ of his/her production rate to reinvestment and $1-u(t)$ to production of a storable good. Thus $x(t)$ evolves according to

$$\dot{x}(t) = \gamma u(t) x(t),$$

where $\gamma > 0$ is a given constant

- The producer wants to maximize the total amount of product stored

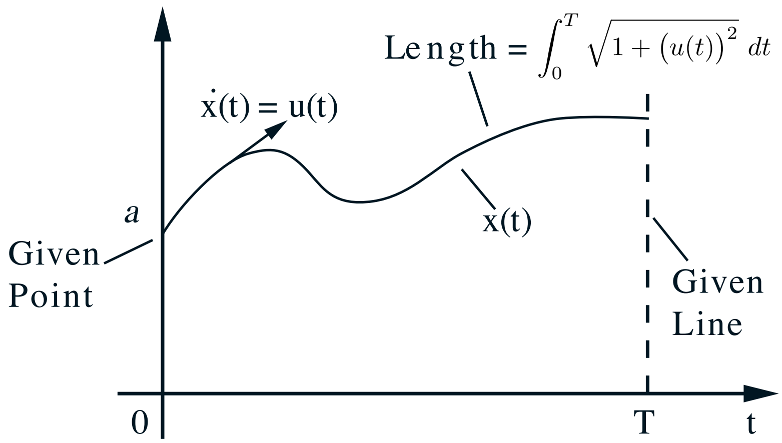
$$\int_0^T \big(1 - u(t)\big) x(t) dt$$

subject to

$$0 \le u(t) \le 1, \qquad \text{for all } t \in [0, T]$$

- The initial production rate $x(0)$ is a given positive number

EXAMPLE III (CALCULUS OF VARIATIONS)



- Find a curve from a given point to a given line that has minimum length
- The problem is

$$\text{minimize} \quad \int_0^T \sqrt{1 + \big(\dot{x}(t)\big)^2}\, dt$$

$$\text{subject to} \quad x(0) = \alpha$$

- Reformulation as an optimal control problem:

$$\text{minimize} \quad \int_0^T \sqrt{1 + \big(u(t)\big)^2}\, dt$$

subject to $\dot{x}(t) = u(t), \; x(0) = \alpha$

HAMILTON-JACOBI-BELLMAN EQUATION I

- We discretize $[0, T]$ at times $0, \delta, 2\delta, \ldots, N\delta$, where $\delta = T/N$, and we let

$$x_k = x(k\delta), \quad u_k = u(k\delta), \qquad k = 0, 1, \ldots, N$$

- We also discretize the system and cost:

$$x_{k+1} = x_k + f(x_k, u_k) \cdot \delta, \quad h(x_N) + \sum_{k=0}^{N-1} g(x_k, u_k) \cdot \delta$$

- We write the DP algorithm for the discretized problem

$$\tilde{J}^*(N\delta, x) = h(x),$$

$$\tilde{J}^*(k\delta, x) = \min_{u \in U} \Big[g(x, u) \cdot \delta + \tilde{J}^* \big((k+1) \cdot \delta, x + f(x, u) \cdot \delta \big) \Big].$$

- Assume $\tilde{J}^*$ is differentiable and Taylor-expand:

$$\tilde{J}^*(k\delta, x) = \min_{u \in U} \Big[g(x, u) \cdot \delta + \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x) \cdot \delta + \nabla_x \tilde{J}^*(k\delta, x)' f(x, u) \cdot \delta + o(\delta) \Big]$$

- Cancel $\tilde{J}^*(k\delta, x)$, divide by δ, and take limit

HAMILTON-JACOBI-BELLMAN EQUATION II

- Let $J^*(t,x)$ be the optimal cost-to-go of the continuous problem. Assuming the limit is valid

$$\lim_{k\to\infty,\,\delta\to 0,\,k\delta=t} \tilde{J}^*(k\delta,x) = J^*(t,x), \qquad \text{for all } t,x,$$

we obtain for all t,x,

$$0 = \min_{u\in U}\big[g(x,u) + \nabla_t J^*(t,x) + \nabla_x J^*(t,x)' f(x,u)\big]$$

with the boundary condition $J^*(T,x) = h(x)$

- This is the *Hamilton-Jacobi-Bellman (HJB) equation* – a *partial* differential equation, which is satisfied for all time-state pairs (t,x) by the cost-to-go function $J^*(t,x)$ (assuming J^* is differentiable and the preceding informal limiting procedure is valid)

- Hard to tell *a priori* if $J^*(t,x)$ is differentiable

- So we use the HJB Eq. as a verification tool; if we can solve it for a differentiable $J^*(t,x)$, then:
  - J^* is the optimal-cost-to-go function
  - The control $\mu^*(t,x)$ that minimizes in the RHS for each (t,x) defines an optimal control

VERIFICATION/SUFFICIENCY THEOREM

- Suppose $V(t,x)$ is a solution to the HJB equation; that is, V is continuously differentiable in t and x, and is such that for all t, x,

$$0 = \min_{u \in U} \big[g(x,u) + \nabla_t V(t,x) + \nabla_x V(t,x)' f(x,u) \big],$$

$$V(T,x) = h(x), \qquad \text{for all } x$$

- Suppose also that $\mu^*(t,x)$ attains the minimum above for all t and x
- Let $\{x^*(t) \mid t \in [0,T]\}$ and $u^*(t) = \mu^*\big(t, x^*(t)\big)$, $t \in [0,T]$, be the corresponding state and control trajectories
- Then

$$V(t,x) = J^*(t,x), \qquad \text{for all } t, x,$$

and $\{u^*(t) \mid t \in [0,T]\}$ is optimal

PROOF

Let $\{(\hat{u}(t), \hat{x}(t)) \mid t \in [0,T]\}$ be any admissible control-state trajectory. We have for all $t \in [0,T]$

$$0 \le g\big(\hat{x}(t), \hat{u}(t)\big) + \nabla_t V\big(t, \hat{x}(t)\big) + \nabla_x V\big(t, \hat{x}(t)\big)' f\big(\hat{x}(t), \hat{u}(t)\big).$$

Using the system equation $\dot{\hat{x}}(t) = f\big(\hat{x}(t), \hat{u}(t)\big)$, the RHS of the above is equal to

$$g\big(\hat{x}(t), \hat{u}(t)\big) + \frac{d}{dt}\big(V(t, \hat{x}(t))\big)$$

Integrating this expression over $t \in [0,T]$,

$$0 \le \int_0^T g\big(\hat{x}(t), \hat{u}(t)\big)dt + V\big(T, \hat{x}(T)\big) - V\big(0, \hat{x}(0)\big).$$

Using $V(T,x) = h(x)$ and $\hat{x}(0) = x(0)$, we have

$$V\big(0, x(0)\big) \le h\big(\hat{x}(T)\big) + \int_0^T g\big(\hat{x}(t), \hat{u}(t)\big)dt.$$

If we use $u^*(t)$ and $x^*(t)$ in place of $\hat{u}(t)$ and $\hat{x}(t)$, the inequalities becomes equalities, and

$$V\big(0, x(0)\big) = h\big(x^*(T)\big) + \int_0^T g\big(x^*(t), u^*(t)\big)dt$$

EXAMPLE OF THE HJB EQUATION

Consider the scalar system $\dot{x}(t) = u(t)$, with $|u(t)| \leq 1$ and cost $(1/2)\big(x(T)\big)^2$. The HJB equation is

$$0 = \min_{|u|\leq 1} \big[\nabla_t V(t,x) + \nabla_x V(t,x) u\big], \qquad \text{for all } t, x,$$

with the terminal condition $V(T,x) = (1/2)x^2$

- Evident candidate for optimality: $\mu^*(t,x) = -\text{sgn}(x)$. Corresponding cost-to-go

$$J^*(t,x) = \frac{1}{2}\Big(\max\big\{0, \, |x| - (T-t)\big\}\Big)^2.$$

- We verify that J^* solves the HJB Eq., and that $u = -\text{sgn}(x)$ attains the min in the RHS. Indeed,

$$\nabla_t J^*(t,x) = \max\big\{0, |x| - (T-t)\big\},$$

$$\nabla_x J^*(t,x) = \text{sgn}(x) \cdot \max\big\{0, \, |x| - (T-t)\big\}.$$

Substituting, the HJB Eq. becomes

$$0 = \min_{|u|\leq 1} \big[1 + \text{sgn}(x) \cdot u\big] \max\big\{0, |x| - (T-t)\big\}$$

LINEAR QUADRATIC PROBLEM

Consider the n-dimensional linear system

$$\dot{x}(t) = Ax(t) + Bu(t),$$

and the quadratic cost

$$x(T)'Q_T x(T) + \int_0^T \big(x(t)'Qx(t) + u(t)'Ru(t)\big)dt$$

The HJB equation is

$$0 = \min_{u \in \Re^m} \Big[x'Qx + u'Ru + \nabla_t V(t,x) + \nabla_x V(t,x)'(Ax+Bu)\Big],$$

with the terminal condition $V(T,x) = x'Q_T x$. We try a solution of the form

$$V(t,x) = x'K(t)x, \qquad K(t): n \times n \text{ symmetric},$$

and show that $V(t,x)$ solves the HJB equation if

$$\dot{K}(t) = -K(t)A - A'K(t) + K(t)BR^{-1}B'K(t) - Q$$

with the terminal condition $K(T) = Q_T$