Optimal Control & Viscosity Solutions

Tutorial Slides from
Banff International Research Station Workshop 11w5086:
Advancing Numerical Methods for
Viscosity Solutions and Applications

Ian M. Mitchell

mitchell@cs.ubc.ca

http://www.cs.ubc.ca/~mitchell

University of British Columbia
Department of Computer Science

February 2011

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

Control Theory

- Control theory is the mathematical study of methods to steer the evolution of a dynamic system to achieve desired goals
  - For example, stability or tracking a reference
- **Optimal control** is a branch of control theory that seeks to steer the evolution so as to optimize a specific objective functional
  - There are close connections with calculus of variations
- Mathematical study of control requires predictive models of the system evolution
  - Assume **Markovian** models: everything relevant to future evolution of the system is captured in the current state
- Many classes of models, but we will talk primarily about deterministic, continuous state, continuous time systems
  - Other continuous models: stochastic DEs, delay DEs, differential algebraic equations, differential inclusions, ...
  - Other classes of dynamic evolution: discrete time (eg: discrete event), discrete state (eg: Markov chains), ...

System Models

- Deterministic, continuous state, continuous time systems are often modeled with ordinary differential equations (ODEs)

$$\dot{x}(t) = \frac{dx(t)}{dt} = f(x(t), u(t))$$

with state $x(t) \in \mathbb{R}^{d_x}$, input $u \in \mathcal{U} \subseteq \mathbb{R}^{d_u}$, and initial condition $x(0) = x_0$

  - To ensure that trajectories are well-posed (they exist and are unique), it is typically assumed that f is bounded and Lipschitz continuous with respect to x for fixed u
  - The field of **system identification** studies how to determine f

- An important subclass of system dynamics are linear

$$\dot{x}(t) = \mathbf{A}x + \mathbf{B}u$$

with $\mathbf{A} \in \mathbb{R}^{d_x \times d_x}$ and $\mathbf{B} \in \mathbb{R}^{d_x \times d_u}$

  - Unless specifically described as "nonlinear control," most engineering control theory (academic and practical) assumes linear systems

Optimal Control Objectives

- Choose input signal

$$u(\cdot) \in \mathfrak{U} \triangleq \{u : [0, \infty[\to \mathcal{U} \mid u(\cdot) \text{ is measureable}\}$$

  to minimize the **cost functional** $J(x, u(\cdot))$ or $J(x, t, u(\cdot))$
- Many possible cost functionals exist, such as:
  - **Finite horizon**: given horizon $T > 0$, running cost ℓ and terminal cost g

$$J(x(t), t, u(\cdot)) \triangleq \int_t^T \ell(x(s), u(s))\, ds + g(x(T))$$

  - **Minimum time**: given target set $\mathcal{T} \subset \mathbb{R}^{d_x}$

$$J(x_0, u(\cdot)) \triangleq \begin{cases} \min\{t \mid x(t) \in \mathcal{T}\}, & \text{if } \{t \mid x(t) \in \mathcal{T}\} \neq \emptyset; \\ +\infty, & \text{otherwise} \end{cases}$$

  - **Discounted infinite horizon**: given discount factor $\lambda > 0$ and running cost ℓ

$$J(x_0, u(\cdot)) \triangleq \int_0^\infty \ell(x(s), u(s)) e^{-\lambda s}\, ds$$

- Alternatively, "maximize payoff functionals" or "optimize objective functionals"

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

Value Functions

- The **value function** specifies the best possible value of the cost functional starting from each state (and possibly time)

$$V(x) = \inf_{u(\cdot) \in \mathfrak{U}} J(x, u(\cdot)) \qquad \text{or} \qquad V(x,t) = \inf_{u(\cdot) \in \mathfrak{U}} J(x, t, u(\cdot))$$

  - Infimum may not be achievable
  - If infimum is attained then the (possibly non-unique) optimal input is often designated $u^*(\cdot)$, and sometimes the corresponding optimal trajectory is designated $x^*(\cdot)$
- Intuitively, to find the best trajectory from a point x, go to a neighbour $\hat{x}$ of x which minimizes the sum of the cost from x to $\hat{x}$ and the cost to go from $\hat{x}$.
  - This intuition is formalized in the **dynamic programming principle**

Dynamic Programming Principle

- For concreteness, we assume a finite horizon objective with horizon T, running cost $\ell(x, u)$ and terminal cost $g(x)$
- Dynamic Programming Principle (DPP): for each $h > 0$ small enough that $t + h < T$

$$V(x, t) = \inf_{u(\cdot) \in \mathfrak{U}} \left[\int_t^{t+h} \ell(x(s), u(s))\, ds + V(x(t+h), t+h) \right]$$

  - Similar DPP can be formulated for other objective functionals
- Proof [Evans, chapter 10.3.2] in two parts: For any $\epsilon > 0$
  - Show that $V(x, t) \leq \inf_{u(\cdot)} \left[\int_t^{t+h} \ell(x(s), u(s))\, ds + V(x(t+h), t+h) \right] + \epsilon$
  - Show that $V(x, t) \geq \inf_{u(\cdot)} \left[\int_t^{t+h} \ell(x(s), u(s))\, ds + V(x(t+h), t+h) \right] - \epsilon$

Proof of DPP (upper bound part 1)

Consider $V(\hat{x}, t)$

- Choose any $u_1(\cdot)$ and define the trajectory

$$\dot{x}_1(s) = f(x_1(s), u_1(s)) \text{ for } s > t \text{ and } x_1(t) = \hat{x}$$

- Fix $\epsilon > 0$ and choose $u_2(\cdot)$ such that

$$V(x_1(t+h), t+h) + \epsilon \geq \int_{t+h}^{T} \ell(x_2(s), u_2(s))\, ds + g(x_2(T))$$

  where

$$\dot{x}_2(s) = f(x_2(s), u_2(s)) \text{ for } s > t+h \text{ and } x_2(t+h) = x_1(t+h)$$

- Define a new control

$$u_3(s) = \begin{cases} u_1(s), & \text{if } s \in [t, t+h[; \\ u_2(s), & \text{if } s \in [t+h, T] \end{cases}$$

  which gives rise to trajectory

$$\dot{x}_3(s) = f(x_3(s), u_3(s)) \text{ for } s > t \text{ and } x_3(t) = \hat{x}$$

Proof of DPP (upper bound part 2)

- By uniqueness of solutions of ODEs

$$x_3(s) = \begin{cases} x_1(s), & \text{if } s \in [t, t+h]; \\ x_2(s), & \text{if } s \in [t+h, T] \end{cases}$$

- Consequently

$$\begin{aligned} V(\hat{x}, t) &\leq J(\hat{x}, t, u_3(\cdot)) \\ &= \int_t^T \ell(x_3(s), u_3(s))\, ds + g(x_3(T)) \\ &= \int_t^{t+h} \ell(x_1(s), u_1(s))\, ds + \int_{t+h}^T \ell(x_2(s), u_2(s))\, ds + g(x_2(T)) \\ &\leq \int_t^{t+h} \ell(x_1(s), u_1(s))\, ds + V(x_1(t+h), t+h) + \epsilon \end{aligned}$$

- Since $u_1(\cdot)$ was arbitrary, it must be that

$$V(\hat{x}, t) \leq \inf_{u(\cdot) \in \mathfrak{U}} \left[\int_t^{t+h} \ell(x(s), u(s))\, ds + V(x(t+h), t+h) \right] + \epsilon$$

Proof of DPP (lower bound)

- Fix $\epsilon > 0$ and choose $u_4(\cdot)$ such that

$$V(\hat{x}, t) \geq \int_t^T \ell(x_4(s), u_4(s))\, ds + g(x_4(T)) - \epsilon$$

  where

$$\dot{x}_4(s) = f(x_4(s), u_4(s)) \text{ for } s > t \text{ and } x_4(t) = \hat{x}$$

- From the definition of the value function

$$V(x_4(t+h), t+h) \leq \int_{t+h}^T \ell(x_4(s), u_4(s))\, ds + g(x_4(T))$$

- Consequently

$$V(\hat{x}, t) \geq \inf_{u(\cdot) \in \mathfrak{U}} \left[\int_t^{t+h} \ell(x(s), u(s))\, ds + V(x(t+h), t+h) \right] + \epsilon$$

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

A Formal Derivation of the Hamilton-Jacobi PDE (part 1)

- Assume that $V(x,t)$ is smooth
- Start from rearranged DPP

$$\inf_{u(\cdot)\in\mathfrak{U}} \left[V(x(t+h), t+h) - V(x,t) + \int_t^{t+h} \ell(x(s), u(s))\, ds \right] = 0$$

- Divide through by $h > 0$

$$\inf_{u(\cdot)\in\mathfrak{U}} \left[\frac{V(x(t+h), t+h) - V(x,t)}{h} + \frac{1}{h}\int_t^{t+h} \ell(x(s), u(s))\, ds \right] = 0$$

- Let $h \to 0$

$$\inf_{u(\cdot)\in\mathfrak{U}} \left[\frac{d}{dt} V(x,t) + \ell(x(t), u(t)) \right] = 0$$

- Apply chain rule on first term

$$\inf_{u(\cdot)\in\mathfrak{U}} \left[D_t V(x,t) + D_x V(x,t) \cdot \frac{d}{dt} x(t) + \ell(x(t), u(t)) \right] = 0$$

A Formal Derivation of the Hamilton-Jacobi PDE (part 2)

- Introduce system dynamics $\dot{x} = f(x, u)$

$$\inf_{u(\cdot) \in \mathfrak{U}} \left[D_t V(x,t) + D_x V(x,t) \cdot f(x(t), u(t)) + \ell(x(t), u(t)) \right] = 0$$

- Observe that only dependence on $u(\cdot) \in \mathfrak{U}$ is $u(t) = u \in \mathcal{U}$

$$\inf_{u \in \mathcal{U}} \left[D_t V(x,t) + D_x V(x,t) \cdot f(x, u) + \ell(x, u) \right] = 0$$

  - If $\mathcal{U}$ is compact, infimum becomes minimum
- Arrive at (time-dependent) Hamilton-Jacobi(-Bellman) PDE

$$D_t V(x,t) + H\left(x, D_x V(x,t)\right) = 0$$

with Hamiltonian

$$H(x,p) = \inf_{u \in \mathcal{U}} \left[p \cdot f(x,u) + \ell(x,u) \right]$$

and terminal conditions (choose $t = T$ in definition of V)

$$V(x,T) = g(x)$$

No Classical Solutions

- Unfortunately, even for smooth terminal conditions, running cost and dynamics, solution of HJ PDE may not remain differentiable for all time
  - A rigorous derivation must take into account that the value function may not be differentiable, and that the optimal input and/or trajectory may not be unique or may not exist
- Search for well-posed weak solutions included the **vanishing viscosity** solution
  - For $\epsilon > 0$, the semilinear or quasilinear parabolic PDE

$$D_t V(x, t) + H\left(t, x, D_x V(x, t)\right) = \epsilon \Delta V(x, t)$$

    has a smooth solution for all time
  - The vanishing viscosity solution is the limiting solution as $\epsilon \to 0$
  - Unfortunately, it does not always exist

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

Viscosity Solutions

- Crandall & Lions (1983) propose the “viscosity solution”
  - Under reasonable conditions there exists a unique viscosity solution
  - Anywhere that V is differentiable, it solves the HJ PDE in the classical sense
  - If there exists a vanishing viscosity solution, then it is the same as the viscosity solution
- Original definition has been supplanted by an equivalent definition from Crandall, Evans & Lions (1984): $V(x,t)$ is a **viscosity solution** of the terminal value HJ PDE

$$D_t V(x,t) + H(x, D_x V(x,t)) = 0$$
$$V(x,T) = g(x)$$

  if V satisfies the terminal conditions and for each smooth $\phi(x,t)$
  - if $V(x,t) - \phi(x,t)$ has a local maximum then

$$D_t \phi(x,t) + H(x, D_x \phi(x,t)) \geq 0$$

  - if $V(x,t) - \phi(x,t)$ has a local minimum then

$$D_t \phi(x,t) + H(x, D_x \phi(x,t)) \leq 0$$

- For initial value HJ PDE, reverse the inequalities

Assumptions and Bounds

- Assume that dynamics, running and terminal costs are bounded and Lipschitz continuous: there exists a constant C such that for fixed u

$$
\begin{aligned}
|f(x,u)| &\leq C & \qquad |f(x,a) - f(\hat{x},a)| &\leq C|x-\hat{x}| \\
|\ell(x,u)| &\leq C & \qquad |\ell(x,a) - \ell(\hat{x},a)| &\leq C|x-\hat{x}| \\
|g(x)| &\leq C & \qquad |g(x) - g(\hat{x})| &\leq C|x-\hat{x}|
\end{aligned}
$$

- This assumption implies continuity properties for the Hamiltonian, but more generally we could assume such properties: there exists a constant C such that

$$
\begin{aligned}
|H(x,p) - H(x,\hat{p})| &\leq C|p-\hat{p}| \\
|H(x,p) - H(\hat{x},p)| &\leq C|x-\hat{x}|(1+|p|)
\end{aligned}
$$

- Then it can be shown that the value function is bounded and Lipschitz continuous: there exists a constant C

$$
\begin{aligned}
|V(x,t)| &\leq C \\
|V(x,t) - V(\hat{x},\hat{t})| &\leq C(|x-\hat{x}| + |t-\hat{t}|)
\end{aligned}
$$

Proof: Value Function is the Viscosity Solution (terminal condition and local maximum part 1)

- From the definition of the value function and objective functional

$$V(x,T) = \inf_{u(\cdot)} J(x,T,u(\cdot)) = \int_T^T \ell(x(s),u(s))\,ds + g(x(T)) = g(x)$$

- Choose smooth ϕ and assume that $V - \phi$ has a local maximum at $(\hat{x},\hat{t})$
  - Then we must show

$$D_t\phi(\hat{x},\hat{t}) + \min_{u\in\mathcal{U}}\left[D_x\phi(\hat{x},\hat{t})\cdot f(\hat{x},u) + \ell(\hat{x},u)\right] \geq 0$$

  - Since $V - \phi$ has a local maximum, choose $\delta > 0$ such that for all $|x-\hat{x}| + |t-\hat{t}| \leq \delta$

$$(V-\phi)(x,t) \leq (V-\phi)(\hat{x},\hat{t})$$

  - Proof proceeds by contradiction: if the inequality is false then there exist $\hat{u}\in\mathcal{U}$ and $\xi > 0$ such that for all $|x-\hat{x}| + |t-\hat{t}| \leq \delta$ we have

$$D_t\phi(x,t) + D_x\phi(x,t)\cdot f(x,\hat{u}) + \ell(x,\hat{u}) \leq -\xi < 0$$

  - Choose constant control $u(\cdot) = \hat{u}$ and define the corresponding trajectory

$$\dot{x}(s) = f(x(s),\hat{u}) \text{ for } s > \hat{t} \text{ and } x(t) = \hat{x}$$

Proof: Value Function is the Viscosity Solution (local maximum part 2)

- Working on contradiction if $V - \phi$ has a local maximum at $(\hat{x}, \hat{t})$
  - Choose $h \in [0, \delta]$ small enough that $|x(s) - \hat{x}| \leq \delta$ for $s \in [\hat{t}, \hat{t} + h]$ so that

$$D_t \phi(x(s), s) + D_x \phi(x(s), s) \cdot f(x(s), \hat{u}) + \ell(x(s), \hat{u}) \leq -\xi$$

  - Because $V - \phi$ has a local maximum

$$\begin{aligned} V(x(\hat{t}+h), \hat{t}+h) - V(\hat{x}, \hat{t}) &\leq \phi(x(\hat{t}+h), \hat{t}+h) - \phi(\hat{x}, \hat{t}) \\ &= \int_{\hat{t}}^{\hat{t}+h} \frac{d}{ds} \phi(x(s), s)\, ds \\ &= \int_{\hat{t}}^{\hat{t}+h} D_t \phi(x(s), s) + D_x \phi(x(s), s) \cdot f(x(s), \hat{u})\, ds \end{aligned}$$

  - From the DPP

$$V(\hat{x}, \hat{t}) \leq \int_{\hat{t}}^{\hat{t}+h} \ell(x(s), \hat{u})\, ds + V(x(\hat{t}+h), \hat{t}+h)$$

  - Therefore we arrive at the contradiction

$$0 \leq \int_{\hat{t}}^{\hat{t}+h} D_t \phi(x(s), s) + D_x \phi(x(s), s) \cdot f(x(s), \hat{u}) + \ell(x(s), \hat{u})\, ds \leq -\xi h$$

Proof: Value Function is the Viscosity Solution (local minimum part 1)

- Choose smooth ϕ and assume that $V - \phi$ has a local minimum at $(\hat{x}, \hat{t})$
  - Then we must show

$$D_t \phi(\hat{x}, \hat{t}) + \min_{u \in \mathcal{U}} \left[D_x \phi(\hat{x}, \hat{t}) \cdot f(\hat{x}, u) + \ell(\hat{x}, u) \right] \leq 0$$

  - Since $V - \phi$ has a local minimum, choose $\delta > 0$ such that for all $|x - \hat{x}| + |t - \hat{t}| \leq \delta$

$$(V - \phi)(x, t) \geq (V - \phi)(\hat{x}, \hat{t})$$

  - Proof proceeds by contradiction: if the inequality is false then there exists $\xi > 0$ such that for all $\hat{u} \in \mathcal{U}$ and $|x - \hat{x}| + |t - \hat{t}| \leq \delta$ we have

$$D_t \phi(x, t) + D_x \phi(x, t) \cdot f(x, \hat{u}) + \ell(x, \hat{u}) \geq \xi > 0$$

  - For any control $u(\cdot) \in \mathfrak{U}$ choose $h \in [0, \delta]$ small enough that $|x(s) - \hat{x}| \leq \delta$ for $s \in [\hat{t}, \hat{t} + h]$ and the corresponding trajectory

$$\dot{x}(s) = f(x(s), u(s)) \text{ for } s > \hat{t} \text{ and } x(t) = \hat{x}$$

Proof: Value Function is the Viscosity Solution (local minimum part 2)

- Working on contradiction if $V - \phi$ has a local minimum at $(\hat{x}, \hat{t})$
  - Because $V - \phi$ has a local minimum

$$
\begin{aligned}
V(x(\hat{t}+h), \hat{t}+h) - V(\hat{x}, \hat{t}) &\geq \phi(x(\hat{t}+h), \hat{t}+h) - \phi(\hat{x}, \hat{t}) \\
&= \int_{\hat{t}}^{\hat{t}+h} \frac{d}{ds}\phi(x(s), s)\, ds \\
&= \int_{\hat{t}}^{\hat{t}+h} D_t\phi(x(s), s) + D_x\phi(x(s), s) \cdot f(x(s), u(s))\, ds
\end{aligned}
$$

  - From the DPP we can choose a control $u(\cdot) \in \mathfrak{U}$ such that

$$
V(\hat{x}, \hat{t}) \geq \int_{\hat{t}}^{\hat{t}+h} \ell(x(s), u(s))\, ds + V(x(\hat{t}+h), \hat{t}+h) - \frac{\xi h}{2}
$$

  - Therefore we arrive at the contradiction

$$
\frac{\xi h}{2} \geq \int_{\hat{t}}^{\hat{t}+h} D_t\phi(x(s), s) + D_x\phi(x(s), s) \cdot f(x(s), u(s)) + \ell(x(s), u(s))\, ds \geq \xi h
$$

Synthesizing an Optimal Control

- Given (viscosity) solution $V(x, t)$, the optimal control is

$$u^*(x, t) \in \arg\min_{u \in \mathcal{U}} \left[D_x V(x, t) \cdot f(x, u) + \ell(x, u) \right]$$

  - Such a control is called a time-dependent **feedback control** since it depends on the current time and state
  - Optimal choice may not be unique
  - Issues arise when $V(x, t)$ is not differentiable, gradient is zero and/or Hamiltonian is (locally) independent of input

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

Hamilton-Jacobi Equations for Discounted Infinite Horizon

- Given discount factor $\lambda > 0$ and running cost ℓ, objective is

$$J(x_0, u(\cdot)) = \int_0^\infty \ell(x(s), u(s)) e^{-\lambda s}\, ds$$

- The value function $V(x) = \inf_{u(\cdot) \in \mathfrak{U}} J(x, u(\cdot))$ satisfies the dynamic programming principle

$$V(x) = \inf_{u(\cdot) \in \mathfrak{U}} \left[\int_0^h \ell(x(s), u(s) e^{-\lambda s}\, ds + V(x(h)) e^{-\lambda h} \right]$$

  and static HJ PDE

$$\lambda V(x) - \min_{u \in \mathcal{U}} \left[D_x V(x) \cdot f(x, u) + \ell(x, u) \right] = 0 \qquad \text{for } x \in \mathbb{R}^{d_x}$$

- Another relatively well behaved problem
  - Similar results to finite horizon problem: viscosity solution $V(x)$ is bounded and continuous but not necessarily differentiable
  - Optimal feedback input is time-independent

Hamilton-Jacobi Equations for Minimum Time

- Given target $\mathcal{T}$, objective is

$$J(x_0, u(\cdot)) = \begin{cases} \min\{t \mid x(t) \in \mathcal{T}\}, & \text{if } \{t \mid x(t) \in \mathcal{T}\} \neq \emptyset; \\ +\infty, & \text{otherwise} \end{cases}$$

- Let $\Omega = \{x \mid V(x) < \infty\}$ be the set of states that give rise to trajectories which can reach the target set in finite time
- The value function $V(x) = \inf_{u(\cdot) \in \mathfrak{U}} J(x, u(\cdot))$ satisfies the dynamic programming principle for $x \in \Omega$

$$V(x) = \inf_{u(\cdot) \in \mathfrak{U}} [h + V(x(h))] \text{ if } h < V(x)$$

and static boundary value HJ PDE

$$H(x, D_x V(x)) = \min_{u \in \mathcal{U}} [D_x V(x) \cdot f(x, u) - 1] = 0 \quad \text{for } x \in \Omega \setminus \mathcal{T}$$
$$V(x) = 0 \quad \text{for } x \in \mathcal{T}$$

Small Time Local Controllability and the Static HJ PDE

- A system is **small time locally controllable** (STLC) at a state x if the set of states which give rise to trajectories which reach x contains x in its interior for all positive times
  - Intuitively, the system can move in any direction
  - Many important types of system are not STLC
- If dynamics are STLC everywhere then the static HJ PDE is relatively well behaved: the viscosity solution $V(x)$ is bounded and continuous (but not necessarily differentiable) and $\Omega = \mathbb{R}^{d_x}$
- If dynamics are not STLC then there may not be a bounded continuous viscosity solution which solves the PDE and/or Ω must be determined

Disturbance Parameters

Sometimes the dynamics are influenced by additional parameters

$$\dot{x} = f(x, u, v)$$

where $v \in \mathbb{R}^{d_v}$ are not known and are not controllable. There are two typical ways of treating these **disturbance inputs**

- Stochastic: $v(t) \sim \mathcal{V}$ where $\mathcal{V}$ is some distribution
  - Modelled by stochastic differential equations (SDEs) in continuous case, or various probabilistic models in discrete settings (Markov chains, discrete state Poisson processes, etc)
  - Optimal control of SDEs leads to Fokker-Plank or Kolmogorov PDEs: second order versions of the HJ PDE
- Bounded value: $v(t) \in \mathcal{V}$ where $\mathcal{V} \subseteq \mathbb{R}^{d_v}$ is a specified set
  - Modelled by standard ODEs with multiple inputs
  - Robust or worst-case treatment of disturbance input is modelled by two player zero sum games and HJ PDE with nonconvex Hamiltonians

Dynamics, Objective Functional and Player Knowledge in Differential Games

- Dynamics and objective functional are almost the same as in the single input case; for example

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

$$J(x(t), t, u(\cdot), v(\cdot)) = \int_t^T \ell(x(s), u(s), v(s))\, ds + g(x(T))$$

  - Control input $u(\cdot) \in \mathfrak{U}$ attempts to minimize
  - Disturbance input $v(\cdot) \in \mathfrak{V}$ attempts to maximize
- In a differential setting, how much does each player know about the other's choice of input?
  - A **non-anticipative strategy** allows one player to know the other player's current input value
  - However, the player with the additional knowledge must declare their strategy (reaction to every input) in advance
  - For example, the disturbance can be given the advantage by permitting it a non-anticipative strategy γ

$$\gamma \in \Gamma(t) = \left\{ \zeta : \mathfrak{U} \to \mathfrak{V} \;\middle|\; \begin{array}{l} u(r) = \hat{u}(r) \text{ for almost every } r \in [t, T] \\ \quad \implies \zeta[u](r) = \zeta[\hat{u}](r) \text{ for almost every } r \in [t, T] \end{array} \right\}$$

Hamilton-Jacobi(-Isaacs) Equations for Differential Games

- Value function is then an optimization over the appropriate strategy and input signal; for example

$$V(x,t) = \sup_{\gamma \in \Gamma(t)} \inf_{u(\cdot) \in \mathfrak{U}} J(x, t, u(\cdot), \gamma[u(\cdot)](\cdot))$$

  - This choice is called the **upper value function** because the maximizing disturbance is given the advantage of the non-anticipative strategy
  - A dual **lower value function** can be defined
  - If the upper and lower value functions are equivalent, then both optimal inputs can be synthesized without strategies as pure state feedback
- The value function satisfies the DPP

$$V(x,t) = \sup_{\gamma \in \Gamma(t)} \inf_{u(\cdot) \in \mathfrak{U}} \left[\int_t^{t+h} \ell(x(s), u(s), \gamma[u](s))\, ds + V(x(t+h), t+h) \right]$$

and the HJ PDE

$$D_t V(x,t) + H\left(x, D_x V(x,t)\right) = 0$$

$$H(x,p) = \min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} \left[p \cdot f(x,u,v) + \ell(x,u,v) \right]$$

  - Optimization in Hamiltonian requires no special treatment of strategies, but it is nonconvex

Fokker-Planck or Kolmogorov Equations for Optimal Stochastic Control

- For system dynamics given by the (Itô) stochastic ordinary differential equation (SDE)

$$dx(t) = f(x(t), u(t))dt + \sigma(x(t))dW(t)$$

  where the (controlled) "drift term" f is the same as in the deterministic ODE case and the "diffusion term" providing the stochastic disturbance is $\sigma : \mathbb{R}^{d_x} \to \mathbb{R}^{d_W}$ and a d_W dimensional Wiener process $W(t)$
- For the finite horizon objective, the value function satisfies a Fokker-Planck or backward Kolmogorov PDE

$$D_t V(x,t) + \min_{u \in \mathcal{U}} \left[D_x V(x,t) \cdot f(x,u) + \ell(x,u) \right] + \tfrac{1}{2} \sigma(x) \sigma^T(x) D_x^2 V(x,t) = 0$$

  - If $d_W = d_x$ and σ is full rank then the PDE is semilinear or quasilinear and under mild assumptions has a classical solution
  - Otherwise the PDE is **degenerate parabolic** and a viscosity solution is the appropriate weak solution
  - Note that solution evolution is no longer governed entirely by characteristics

Other Control Applications with HJ PDEs

- State estimation / observation
  - In most real systems we can only observe sensor outputs—the true state is not directly observable
  - State estimation can be formulated as various types of HJ PDE, depending on the noise model
  - Optimal control subject to state uncertainty can be formulated as an infinite dimensional HJ equation
- Optimal stopping times
  - In some problems the control (or disturbance) can choose the stopping time
  - Can be formulated as a **variational inequality**; for example, for finite horizon objective functional with stopping / terminal cost $g(x)$

$$\max\left[D_t V(x,t) + H(x, D_x V(x,t)), V(x,t) - g(x)\right] = 0$$

- Reachability
  - Next set of slides
- Many more...

Outline

- Optimal control: models of system dynamics and objective functionals
- The value function and the dynamic programming principle
- A formal derivation of the Hamilton-Jacobi(-Bellman) equation
- Viscosity solutions and a rigorous derivation
- Other types of Hamilton-Jacobi equations in control
- Optimal control problems with analytic solutions
- References

Finite Horizon: LQR Formulation

- In the **Linear Quadratic Regulator** (LQR) problem
  - The dynamics are linear
    $$\dot{x} = \mathbf{A}x + \mathbf{B}u$$
    with $u \in \mathcal{U} = \mathbb{R}^{d_u}$
  - The finite horizon objective is quadratic
    $$J(x, t, u(\cdot)) = u^T(T)\mathbf{Q}_f u(T) + \int_t^T x^T(s)\mathbf{Q}x(s) + u^T(s)\mathbf{R}u(s)\, ds$$
    where $\mathbf{Q}_f = \mathbf{Q}_f^T \geq 0$, $\mathbf{Q} = \mathbf{Q}^T \geq 0$, and $\mathbf{R} = \mathbf{R}^T > 0$ are the **terminal state cost**, the **running state cost**, and the **input cost** matrices respectively
  - It can be shown that the value function is quadratic in the state
    $$V(x, t) = \inf_{u(\cdot) \in \mathfrak{U}} J(x, t, u(\cdot)) = x^T \mathbf{P}(t) x$$

Finite Horizon: LQR Solution (part 1)

- Analytic solution can be constructed from a dynamic programming argument
  - Start at state $\hat{x}$ and take $u(s) = \hat{u}$ fixed over a small time interval $s \in [t, t+h]$
  - Cost incurred is

$$\int_t^{t+h} x^T(s)\mathbf{Q}x(s) + u^T(s)\mathbf{R}u(s)\,ds \approx h(\hat{x}^T\mathbf{Q}\hat{x} + \hat{u}^T\mathbf{R}\hat{u})$$

  - State after that time period is $x(t+h) \approx \hat{x} + h(\mathbf{A}\hat{x} + \mathbf{B}\hat{u})$
  - Value function at that new state is

$$\begin{aligned} V(x(t+h), t+h) &= x^T(t+h)\mathbf{P}(t+h)x(t+h) \\ &\approx (\hat{x} + h(\mathbf{A}\hat{x} + \mathbf{B}\hat{u}))^T(\mathbf{P}(t) + h\dot{\mathbf{P}}(t))(\hat{x} + h(\mathbf{A}\hat{x} + \mathbf{B}\hat{u})) \\ &\approx \hat{x}^T\mathbf{P}(t)\hat{x} + h\begin{pmatrix} (\mathbf{A}\hat{x} + \mathbf{B}\hat{u})^T\mathbf{P}(t)\hat{x} \\ +\hat{x}\mathbf{P}(t)(\mathbf{A}\hat{x} + \mathbf{B}\hat{u}) + \hat{x}^T\dot{\mathbf{P}}(t)\hat{x} \end{pmatrix} \end{aligned}$$

Finite Horizon: LQR Solution (part 2)

- Dynamic programming derivation of LQR solution
  - Dynamic programming principle

$$V(\hat{x}, t) = \min_{u(\cdot) \in \mathfrak{U}} \left[\int_t^{t+h} x^T(s)\mathbf{Q}x(s) + u^T(s)\mathbf{R}u(s)\, ds + V(x(t+h), t+h) \right]$$

$$\hat{x}^T\mathbf{P}(t)\hat{x} = \min_{\hat{u} \in \mathbb{R}^{d_u}} \left[\begin{array}{c} h(\hat{x}^T\mathbf{Q}\hat{x} + \hat{u}^T\mathbf{R}\hat{u}) + \hat{x}^T\mathbf{P}(t)\hat{x} \\ +h \left(\begin{array}{c} (\mathbf{A}\hat{x} + \mathbf{B}\hat{u})^T\mathbf{P}(t)\hat{x} \\ +\hat{x}\mathbf{P}(t)(\mathbf{A}\hat{x} + \mathbf{B}\hat{u}) + \hat{x}^T\dot{\mathbf{P}}(t)\hat{x} \end{array} \right) \end{array} \right]$$

$$0 = \min_{\hat{u} \in \mathbb{R}^{d_u}} h \left[\begin{array}{c} \hat{x}^T\mathbf{Q}\hat{x} + \hat{u}^T\mathbf{R}\hat{u} + (\mathbf{A}\hat{x} + \mathbf{B}\hat{u})^T\mathbf{P}(t)\hat{x} \\ +\hat{x}\mathbf{P}(t)(\mathbf{A}\hat{x} + \mathbf{B}\hat{u}) + \hat{x}^T\dot{\mathbf{P}}(t)\hat{x} \end{array} \right]$$

  - Set derivative with respect to $\hat{u}$ to be zero to find optimal $\hat{u}$

$$2h(\hat{u}\mathbf{R} + \hat{x}^T\mathbf{P}(t)\mathbf{B}) = 0$$

$$\hat{u}^* = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}(t)\hat{x}$$

  - Substitute $\hat{u}^*$ into dynamic programming equation and solve for $\dot{\mathbf{P}}(t)$ to find **Riccati differential equation**

$$-\dot{\mathbf{P}}(t) = \mathbf{A}^T\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A} - \mathbf{P}(t)\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}(t) + \mathbf{Q}$$

    with terminal condition $\mathbf{P}(T) = \mathbf{Q}_f$

(In)Finite Horizon: Steady State LQR

- In conclusion: LQR value function is $V(x,t) = x^T \mathbf{P}(t) x$ where $\mathbf{P}(t)$ is the solution to a terminal value (matrix) ODE
- In practice, $\mathbf{P}(t)$ and $\hat{u}^*$ rapidly converge to steady state values
  - Solve (continuous time) **algebraic Riccati equation** for steady state $\mathbf{P}$

$$\mathbf{A}^T \mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T \mathbf{P} + \mathbf{Q} = 0$$

  - Time-independent state feedback given by

$$u(t) = \mathbf{K} x(t) \text{ where } \mathbf{K} = -\mathbf{R}^{-1}\mathbf{B}^T \mathbf{P}$$

- See Stanford's EE363: Linear Dynamical Systems (Stephen Boyd) `http://www.stanford.edu/class/ee363/`
  - This and several more derivations given in lecture notes 4 (Continuous LQR)
  - Other lectures discuss discrete time, Kalman filter (eg: LQR for state estimation), ...
- See any textbook on "state space" / "modern" control

Minimum Time: Double Integrator

- The **double integrator** is one of the simplest systems which is not STLC
  - System states are position x_1 and velocity x_2, and the input is the acceleration $u \in \mathcal{U} = [-1, +1]$

$$f(x, u) = \begin{bmatrix} x_2 \\ u \end{bmatrix}$$

- If target is the origin

$$V(x) = \begin{cases} x_2 + \sqrt{4x_1 + 2x_2^2}, & \text{if } x_1 > \frac{1}{2} x_2 |x_2|; \\ -x_2 + \sqrt{-4x_1 + 2x_2^2}, & \text{if } x_1 < \frac{1}{2} x_2 |x_2|; \\ |x_2|, & \text{if } x_1 < \frac{1}{2} x_2 |x_2| \end{cases}$$

  - Dynamics are small time controllable at the origin, so value function is continuous, but **not** Lipschitz continuous
  - Optimal trajectories / characteristics travel along the curve where Lipschitz continuity fails
- If target set is not a circle, value function is discontinuous
- See *Optimal Control*, Athans & Falb (1966) or *Applied Optimal Control*, Bryson & Ho (1975) or many others

Known Solutions for More Complex Dynamics (part 1)

- Unicycle model:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \theta \end{bmatrix} \qquad \dot{x} = \begin{bmatrix} v\cos\theta \\ v\sin\theta \\ \omega \end{bmatrix}$$

  where (x_1, x_2) is position in the plane, θ is heading, v is linear velocity and ω is angular velocity

- Dubins' car: Unicycle with fixed positive linear velocity and bounded angular velocity
  - Alternative viewpoint: unicycle with minimum turn radius
  - Minimum time to reach is generally discontinuous
  - Extensive study of combinatorial aspects of optimal paths in robotics literature: optimal paths include CCC or CSC forms, where C is a minimum radius left or right arc of a circle (possibly of zero length) and S is a straight segment
  - For example, see Bui, Boissonnat, Souères & Laumond, "Shortest Path Synthesis for Dubins Non-holonomic Robot," ICRA 1994

Known Solutions for More Complex Dynamics (part 2)

- Game of two identical vehicles: Collision avoidance with two adversarial Dubins' cars
  - Solved in relative coordinate system, so state space remains three dimensional
  - Reachability problem becomes a two player zero sum differential game becomes a HJI PDE
  - Analytic optimal trajectories can also be enumerated and points on the boundary of the reachable set determined
  - Optimal characteristics both converge and diverge, causing challenges for Lagrangian approaches
  - More details in subsequent set of slides
  - See Mitchell, "Games of Two Identical Vehicles," Stanford University Department of Aeronautics and Astronautics Report (SUDAAR) 740 (2001).
- In summary, there is no shortage of toy optimal control problems with analytic solutions
  - On the other hand, there is no shortage of real optimal control problems without analytic solutions

Viscosity Solution & Control References

- Crandall & Lions (1983) original publication
- Crandall, Evans & Lions (1984) current formulation
- Evans & Souganidis (1984) for differential games
- Crandall, Ishii & Lions (1992) "User's guide" for viscosity solutions of degenerate ellipic and parabolic equations (dense reading)
- *Viscosity Solutions & Applications* Springer Lecture Notes in Mathematics (1995), featuring Bardi, Crandall, Evans, Soner & Souganidis (Capuzzo-Dolcetta & Lions eds)
- *Optimal Control & Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Bardi & Capuzzo-Dolcetta (1997)
- *Partial Differential Equations*, Evans (3rd ed, 2002)