Game Theory 2014

Extra Lecture 1 (BoB)

- Differential games
- Tools from optimal control
- Dynamic programming
- Hamilton-Jacobi-Bellman-Isaacs' equation
- Zerosum linear quadratic games and H_∞ control

Baser/Olsder, pp. 233-246, 265-288, 310-333, 342-350

To understand

- *Definition of difference/differential games*
- *Optimal Control*
- *Dynamic programming and HJBI equation*
- *Open-loop and feedback Nash Equilibria*

Material

- Copies from Baser/Olsder

Difference Games

State equation

$$x_{k+1} = f(x_k, u_1, u_2, \ldots, u_N), \quad x_1 \text{ given}$$

u_i determined by player i

Observations $y_k^i = h_k^i(x_k), \quad i = 1, \ldots, N$

Information structure, for each player a subset η_k^i of

$$\{y_1^1, \ldots, y_k^1; \ldots; y_1^N, \ldots, y_k^N; u_1^1, \ldots, u_{k-1}^1; \ldots; u_1^N, \ldots, u_{k-1}^N\}$$

Goal: Player i wants to minimize some functional, L_i, of x and u given his available information.

Solution concepts: As before, Nash, Stackelberg, etc

Differential Games

$$\frac{d}{dt}x(t) = f(x(t), u_1(t), u_2(t), \ldots, u_N(t)), \quad x(t_0) \text{ given}$$

Goal: to minimize

$$L^i = \int_0^T g^i(x, u^1, \ldots, u^N)\, dt$$

T either fixed, or given by an implicit equation

$$T = \min_t \{t; l(x(t)) \leq 0\}$$

Example, Robust Control H_∞ control

Assume u^1, u^2 and y are related via

$$
\begin{aligned}
\dot{x} &= Ax + B_1u_1 + B_2u_2 \\
z &= C_1x \\
y &= C_2x + Du_2
\end{aligned}
$$

Typical H_∞ question: Is $\int z^2(t) + u_1^2(t)dt \le \gamma^2 \int u_2^2(t)dt$?

Differential game between

controller, $u_1(t)$ function of $y([0,t])$

disturbance, $u_2(t)$ function of $x([0,t])$ and $u_1([0,t])$

Performance criterium

$$\min_{u_1} \max_{u_2} \int_0^\infty (z^2 + u_1^2 - \gamma^2 u_2^2)dt$$

Robust/H_∞ Control

The interpretation is that u^1 is the control signal and u^2 is a worst-case disturbance signal.

Introduce the L_2 (or energy) norm of the signal w

$$||w||_2 = \left(\int_0^\infty |w(t)|^2 \, dt \right)^{1/2}$$

The H_∞ norm of a linear system $G(s)$ is defined by

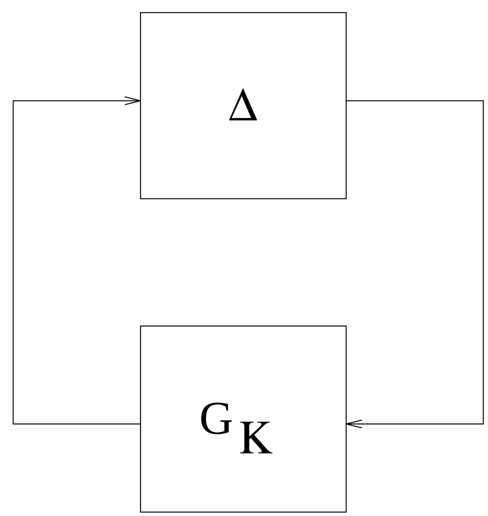
$$||G||_\infty = \sup_{w \neq 0} \frac{||Gw||}{||w||}$$

Robustness

The H_∞ norm measures the largest amplification of energy by the system.

Minimization of the norm is clearly interesting if w is a disturbance signal. Another motivation is given by the so called small-gain theorem

Theorem A closed loop system is stable for all perturbations Δ with norm $||\Delta||_\infty < \gamma$ if and only if $||G_K||_\infty \leq 1/\gamma$



H_∞ control

$$
\begin{aligned}
&||G_K||_\infty < \gamma \Leftrightarrow \\
&\exists K : ||z||^2 - \gamma^2 ||w||^2 < 0 \Leftrightarrow \\
&\min_u \max_w (||z||^2 - \gamma^2 ||w||^2) < 0
\end{aligned}
$$

This is exactly (the upper value of) an affine quadratic game

The relation between H_∞ control and game theory was noted rather late (end of 80s).

For details see Section 6.6

Pursuit Evasion Games

Tools from One-Person Optimization

Dynamic programming

The (maximum) minimum principle

Baser/Olsder Ch 5.5

Dynamic Programming, discrete time

$$
\begin{aligned}
x_{k+1} &= f_k(x_k, u_k), \quad u_k \in U_k \\
L(u) &= \sum_{k=1}^{K} g_k(x_{k+1}, u_k, x_k)
\end{aligned}
$$

f_k, g_k, U_k, K, x_1 given.

Want $u_k = \gamma_k(x_k)$ that minimizes L

Idea: Generalize the problem; Calculate the value function

$$
V(k, x) = \min_{\gamma_k, \ldots, \gamma_K} \sum_{i=k}^{K} g_i(x_{i+1}, u_i, x_k)
$$

for all initial conditions $x_k = x$.

Principle of Optimality

An optimal control sequence $u_1, \ldots, u_K$ should be sequentially optimal. The only coupling between the optimization problems on time horizons $[1, k-1]$ and $[k, K]$ is via the state x_k.

This leads to

$$V(k, x) = \min_{u_k \in U_k} [g_k(f_k(x, u_k), u_k, x) + V(k+1, f_k(x, u_k))]$$

with the final condition $V(K, x) = \min_{u_K} g_k(x_{K+1}, u_K, x_K)$.

Example – Affine quadratic problems

$$
\begin{aligned}
x_{k+1} &= A_k x_k + B_k u_k + c_k \\
L(u) &= \frac{1}{2}\sum_{k=1}^{K}(x'_{k+1} Q x_{k+1} + u'_k R u_k)
\end{aligned}
$$

has the solution

$$
\begin{aligned}
V(k,x) &= \frac{1}{2}x' S_k x + x' s_k + q_k \\
u_k^* &= -P_k S_{k+1} A_k x_k - P_k(s_{k+1} + S_{k+1} c_k)
\end{aligned}
$$

Formulas for P_k, S_k, s_k are given on p.234-5

Dynamic Programming, continuous-time

The same reasoning leads to a PDE equation called the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned}
\dot{x}(t) &= f(x(t), u(t)) \\
u(t) &= \gamma(x(t)) \\
L(u) &= \int_0^T g(x(t), u(t)dt + q(T, x(T))
\end{aligned}$$

The final time T can be either fixed known, or given implicitly by

$$T = \min_{t \geq 0} \{t \ : \ l(x(t)) = 0\}$$

HJB equation

$$V(t,x) = \min_{\{u(s), s\in[t,T]\}} \int_t^T g(x(s),u(s))ds + q(T,x(T))$$
$$V(T,x) = q(T,x) \quad \text{along} \;\; l(x) = 0$$

Principle of optimality shows that if V is C^1 then

$$-\frac{\partial V(t,x)}{\partial t} = \min_{u\in U}\left[\frac{\partial V(t,x)}{\partial x} f(x,u) + g(x,u)\right]$$

($+$ final condition when $t = T$)

Theorem of Sufficiency

Theorem 5.3 p 237 If a C^1 function V can be found that satisfies the HJB equation and boundary conditions above then it generates the optimal strategy u^* through the pointwise optimization problem defined by the right hand side.

Proof see p. 237

Example: Affine-quadratic problems has a quadratic function $V(t,x) = \frac{1}{2}x'S(t)x + k'(t)x + m(t)$, see p.238-9

The Minimum Principle

Introduce the costate vector $p'(t) = \partial V(t, x^*(t))/\partial x$ where x^* denotes the optimal trajectory corresponding to u^*, i.e. $\dot{x}^*(t) = f(x^*(t), u^*(t))$. Also define (the Hamiltonian)

$$H(t, p, x, u) = g(x, u) + p'(t) f(x, u)$$

Theorem 5.4

$$\begin{aligned} p'(T) &= \frac{\partial q(T(x^*), x^*)}{\partial x} \quad \text{along } l(T, x) = 0 \\ \dot{p}'(t) &= \frac{\partial H(t, p, x^*, u^*)}{\partial x} \\ u^* &= \arg\min_{u \in U} H(t, p, x^*, u) \end{aligned}$$

See also the discrete-time counterpart Theorem 5.5

N person difference games

How does this translate to difference games with $N > 1$ players?

$$\begin{aligned} x_{k+1} &= f_k(x_k, u_k^1, \ldots, u_k^N) \\ J^i(u^1, \ldots, u^N) &= \sum_{k=1}^{K} g_k^i(x_{k+1}, u_k^1, \ldots, u_k^N, x_k) \end{aligned}$$

The control laws $u_i = \gamma^{i*}$ constitute a Nash equilibrium if

$$J^i(\gamma^*) \le J^i(\{\gamma_{-i}^*, \gamma_i\}) \quad \forall i$$

Open-loop information: u_k^i is an open-loop function of k

Closed-loop information: $u_k^i(x_k)$ is allowed to be a function of x_k. Hence changing u^i will result in changes in u^j

Open Loop Dynamic Games

See Section 6.2.1

Idea: Player i solves a one-player problem. His choice u^i will not influence the other u^j, so these can be treated as given functions of time.

Then the results from the $N = 1$ apply directly

Open Loop Theorems

Theorem 6.1 If $u^* = \gamma^*$ provides an open-loop Nash equilibrium, then there exists a sequence of costate vectors p^i such that

$$
\begin{aligned}
x^*_{k+1} &= f_k(x^*_k, u^*_k) \\
\gamma^{i*}_k &= \arg\min_{u^i_k} H^i_k(p^i_{k+1}, \{\gamma^{-i*}_k, u^i_k\}, x^*_k) \\
p^i_k &= \frac{\partial}{\partial x_k} f'_k \left[p^i_{k+1} + \left(\frac{\partial}{\partial x_{k+1}} g^i_k \right)' \right] + \left(\frac{\partial}{\partial x_k} g^i_k \right)'
\end{aligned}
$$

where $H^i = g^i + p^{i'} f$. For details see p. 268

The affine case results in coupled Riccati equations. The special zero-sum case for $N = 2$ is given in Theorem 6.3, and 6.4. Only one Riccati equation must be solved. The condition for existence of Nash equilibrium has the form $I - B^{2'}_k S_{k+1} B^2_k > 0$.

Feedback Solutions

Theorem 6.6 The set of strategies γ^* provides a feedback Nash equilibrium if and only if there exists functions $V^i(k,x)$ such that

$$\begin{aligned} V^i(k,x) &= \arg\min_{u_k^i}[g_k^i(-i*, u_k^i) + V^i(k+1, \tilde{f}_k^{i*}(x, u_k^i))] \\ &= g_k^i(-i*, \gamma_k^{i*}) + V^i(k+1, \tilde{f}_k^{i*}(x, \gamma_k^{i*})) \end{aligned}$$

where $\tilde{f}_k^{i*}(x, u_k^i) = f_k(x, \{\gamma^{-i*}(x), u_k^i\})$

Feedback Solutions, Zero-sum case $N = 2$

See Corollary 6.2 p. 282 The set of strategies γ^{1*}, γ^{2*} provides a feedback saddle-point solution if and only if there exists functions $V(k,x)$ such that for all k

$$
\begin{aligned}
V(k,x) & \\
&= \min_{u_k^1} \max_{u_k^2} \left[g_k(f_k(x,u_k^1,u_k^2), u_k^1, u_k^2, x) + V(k+1, f_k(x,u_k^1,u_k^2) \right] \\
&= \max_{u_k^2} \min_{u_k^1} \left[g_k(f_k(x,u_k^1,u_k^2), u_k^1, u_k^2, x) + V(k+1, f_k(x,u_k^1,u_k^2) \right]
\end{aligned}
$$

$$
= g_k(f_k(x,\gamma_k^{1*},\gamma_k^{2*}), \gamma_k^{1*}(x), \gamma_k^{2*}(x), x) + V(k+1, f_k(x, \gamma_k^{1*}(x), \gamma_k^{2*}(x)))
$$

Differential Games Open-loop Nash Equilibria

Want to present counterparts for continuous time.

Existence of smooth cost function V is more problematic.

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), u^1(t), \ldots, u^N(t)) \\
L^i(u^1, \ldots, u^N) = &= \int_0^T g^i(x(t), u^1(t), \ldots, u^N(t))\, dt + q^i(x(T))
\end{aligned}
$$

Open-loop Nash Equilibria

Theorem 6.11 (Some smoothness assumptions) If $u_i^* = \gamma^{i*}(t, x_0)$ provides an open-loop Nash equilibrium (with corresponding x^*) then there exist N costate functions $p^i(t)$ such that

$$
\begin{aligned}
\dot{x}^*(t) &= f(x^*(t), u^{1*}(t), \ldots, u^{N*}(t)) \\
\gamma^{i*} &= \arg\min_{u^i \in U^i} H^i(p^i(t), x^*(t), \{u^{-i*}, u^i\}) \\
\dot{p}^{i'}(t) &= \frac{\partial}{\partial x} H^i(p^i(t), x^*, u^*(t)) \\
\dot{p}^{i'}(T) &= \frac{\partial}{\partial x} q^i(x^*(T))
\end{aligned}
$$

where $H^i(p^i, x, u) = g^i(x, u) + p^{i'} f(x, u)$.

The Affine-Quadratic Case

$$
\begin{aligned}
f(x,u) &= A(t)x + \sum B^i(t)u^i + c(t) \\
g^i(x,u) &= \frac{1}{2}x'Q^i(t)x + \sum u^{j'}R^{ij}u^j \\
q^i(x(T)) &= \frac{1}{2}x'(T)Q_f^i x(T)
\end{aligned}
$$

with $Q^i \geq 0, Q_f^i \geq 0, R^{ii} > 0$.

The Affine-Quadratic Case

Theorem 6.12+p.316 There is an open-loop Nash equilibrium solution to the affine-quadratic game for any $T \in [0, T_f]$ if and only if the following coupled matrix Riccati differential equations have solutions for any $T \in [0, T_f]$

$$\dot{M}^i + M^i A + A' M^i + Q^i - M^i \sum_j B^j R^{jj}(t)^{-1} B^{j'} M^j = 0$$

$$M^i(T) = Q^i_f$$

The NE is given by $u^i = -R^{ii}(t)^{-1} B^{i'} [M^i x^*(t) + m^i]$ where the feedforward signal m^i is given by (6.51)

$N = 2$ Zero-Sum Case

$P1$ minimizes, $P2$ maximizes.

$$L(u^1, u^2) = \frac{1}{2}\int_0^T x'Qx + u^{1'}u^1 - u^{2'}u^2 \, dt + \frac{1}{2}x'(T)Q_f x(T)$$

with $Q \geq 0, \;\; Q_f \geq 0$.

Assume there exists a unique bounded symmetric solution $S(\cdot)$ to the matrix equation

$$\dot{S} + A'S + SA + Q + SB^2B^{2'}S = 0, \qquad S(T) = Q_f$$

on the interval $[0, T]$.

Then there exists a solution to

$$\dot{M} + A'M + MA + Q - M(B^1 N^{1'} - N^2 B^{2'})M = 0; M(T) = Q_f$$

and the game admits a unique open-loop saddle-point given by

$$\gamma^{1*}(t, x_0) = -B^1(t)'[M(t)x^*(t) + m(t)]$$
$$\gamma^{2*}(t, x_0) = B^2(t)'[M(t)x^*(t) + m(t)]$$

Closed-loop Feedback Nash Equilibria

Will only give the result for $N = 2$ and zero-sum situation

Corollary 6.6, p326 A pair of strategies $\{\gamma^{i*}\}$ provides a feedback saddle-point solution if there exists a function V satisfying the PDE

$$
\begin{aligned}
-\frac{\partial V(t,x)}{\partial t} &= \min_{u^1}\max_{u^2}\left[\frac{\partial V(t,x)}{\partial x} f(x,u^1,u^2) + g(x,u^1,u^2)\right] \\
&= \max_{u^2}\min_{u^1}\left[\frac{\partial V(t,x)}{\partial x} f(x,u^1,u^2) + g(x,u^1,u^2)\right] \\
&= \left[\frac{\partial V(t,x)}{\partial x} f(x,\gamma^{1*},\gamma^{2*}) + g(x,\gamma^{1*},\gamma^{2*})\right]
\end{aligned}
$$

The value of the game is $V(0, x_0)$

This is the famous Hamilton-Jacobi-Bellman-Isaacs' equation obtained by Isaacs in 1950s.

$N = 2$, Zero-sum Affine Case

$$\begin{aligned}\dot{x} &= Ax + B^1u^1 + B^2u^2 + c \\ L &= \frac{1}{2}\int_0^T x'Qx + u^{1'}u^1 - u^{2'}u^2\,dt\end{aligned}$$

Theorem 6.17 If there exists a unique symmetric bounded solution to the Riccati equation

$$\dot{Z} + A'Z + ZA + Q - Z(B^1B^{1'} - B^2B^{2'})Z = 0, \quad Z(T) = Q_f$$

then the two-player zero-sum game admits a unique feedback saddle-point given by (for details see p. 327)

$$\gamma^{i*}(t,x) = (-1)^i B^i(t)'[Z(t)x(t) + \zeta(t)]$$