Optimal Control-PMP- and Games

Rafael Murrieta Cid

Centro de Investigación en Matemáticas (CIMAT)

October 2019

Outline

Optimal Control-PMP-

Optimal Control

- Optimal control theory has as a main objective to determine the optimal controls $u(t)^*$ and optimal trajectory $x(t)^*$ for a dynamical system.
- Indeed, the optimal control when applied yields the optimal trajectory.

Functional to be optimized

The functional to be optimized has the following form:

$$J(x,u) = \int_{t_s}^{t_f} L(x(t), u(t))\, \mathrm{d}t + G(x(t_f)) \tag{1}$$

One wants to find

$$\min_{u \in U} \quad J(x,u) \tag{2}$$

$$\text{under} : \dot{x} = f(x,u)$$

$$\text{with} : x(t_s) = x_s$$

where:

- x is the state variable, such that $x(t) \in \Omega$ is an open set of $\mathbb{R}^n$ or $x(t) \in M$ is a n-dimensional manifold, for all t.
- u is the control variable, such that $u(t) \in U$ subset of $\mathbb{R}^m$ of admisible controls, for all t.
- L is the running cost function.
- $G(x(t_f))$ is the terminal cost function.

System Hamiltonian

The system Hamiltonian is defined as follows:

Definition

$$H(x, u, \psi) = \langle \psi(t), f(x, u) \rangle + L(x, u) \tag{3}$$

$\psi(t)$ is the adjoint or costate variable.

Pontryagin Maximum Principle

Pontryagin Maximum Principle gives necessary conditions for the optimal control problem.

Theorem (PMP)

Let $x^*(t)$ and $u^*(t)$,the optimal trajectory and control respectively of the functional (1). Then, there exists $\psi^*(t)$ such that:

- $\dot{\psi}^* = -\frac{\partial H(x^*, u^*, \psi^*)}{\partial x}$ *with* $\psi^*(t_f) = \frac{\partial G(x^*(t_f))}{\partial x}$. *This is known as the adjoint equation.*
- $\dot{x}^*(t) = f(x^*(t), u^*(t))$.
- $u^*(t)$ *optimizes the Hamiltonian, that is, when one wants to maximize the Hamiltonian, one has* $H(x^*, u^*, \psi^*) \geq H(x^*, u, \psi^*)$, *and when one wants to minimize the Hamiltonian, one has* $H(x^*, u^*, \psi^*) \leq H(x^*, u, \psi^*)$. *Furthermore,* $H(x^*(t), u^*(t), \psi^*(t))$ *is constant in the domain of* t.

Geometric interpretation of the PMP

- It is convient to consider functional J as a state variable.
- Then one has $\hat{x}(t) = (J(x(t), u(t)), x(t))$, and its velocity vector is $\dot{\hat{x}}(t) = \hat{f}(x, u) = (L(x(t), u(t)), f(x(t), u(t)))$.
- Let D be the line passing through $(0, x_f)$ parallel to the axis $\hat{x}_0$.
- Among all the trajectories starting at $\hat{x}_s = (0, x_s)$ and reaching D, one looks for the one that maximizes the first coordinate x^0 of the intersection point $x = (x^0, x_f)$ with D.
- For the optimality principle, if $u^*(t)$ is the optimal control then for all moment t, the trajectory associate to $u^*(t)$ is at $\partial\mathcal{R}(\hat{x}_s, t)$, the frontier of $\mathcal{R}(\hat{x}_s, t)$, the reachable region of $\hat{x}_s$ at time t.
- Consider all other trajectories of control $\hat{u}(t)$, such that those trajectories are different from the one generated by the optimal control $u^*(t)$ by small time intervals and which are infinitesimal close to $u^*(t)$.
- The set of states reached by those *perturbed* trajectories constitute a convex cone C with vertex $x(t, u^*(t))$ and within $\mathcal{R}(x_s, t)$.

Geometric interpretation PMP

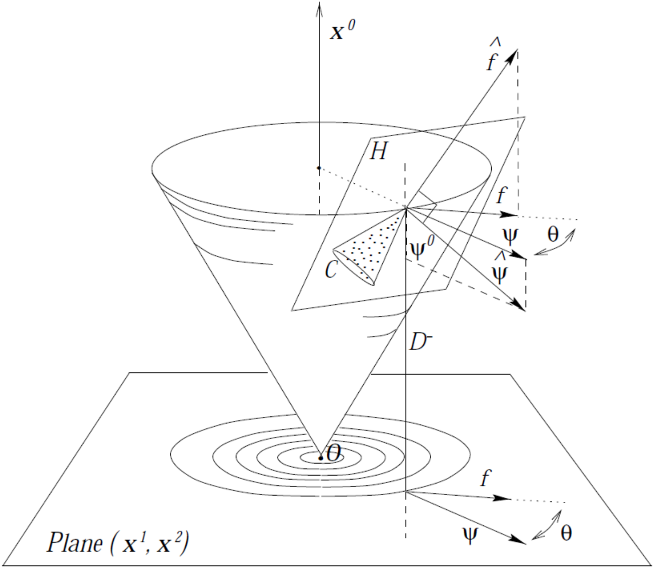


Figure: x^0, $f(x, u)$ and Ψ .

Geometric interpretation of the PMP

- This cone C can be locally approximated to the set $\mathcal{R}(x_s, t)$.
- It does not intersect la half line D^- on the vector $\hat{f}(x(t), u^*(t))$ tangent to trajectory $\hat{x}(t)$ for all t, otherwise it contradicts the optimality of $u^*(t)$.
- It is possible to find a hyperplane H tangent to C at point $x(t, u^*(t))$, dividing C and the half line D^-, containing vector $\hat{f}(x(t), u^*(t))$.
- Since the adjoint vector, $\hat{\psi} = (\psi_0, \psi^*)$, is the one that optimizes and makes constant the Hamiltonian, then $\hat{\psi}$ is ortogonal to al hyperplane H at all time.
- Following the principle of optimal evolution, the projection of vector $\hat{f}(x(t), u^*(t))$ on the line parallel to $\hat{\psi}(t)$, passing through $x(t, u(t))$, should be maximal for control $u(t)$.
- This previous point is equivalent to have a maximal projection of $f(x(t), u^*(t))$ on $\psi(t)$.

Games Concepts

Definitions

Payoff

A standard representation [Isaacs65, Basar95] of the payoff is

$$J(\mathbf{x}(t_s), u, v) \quad = \quad \int_{t_s}^{t_f} \underbrace{L(\mathbf{x}(\bar{t}), u(\bar{t}), v(\bar{t}))}_{\text{running cost}} d\bar{t} \quad + \quad \underbrace{G(\mathbf{x}(t_f))}_{\text{terminal cost}}$$

For problems of *minimum time* [Isaacs65, Basar95], as in this game, $L(\mathbf{x}(t), u(t), v(t)) = 1$ and $G(t_f, \mathbf{x}(t_f)) = 0$. Therefore in our game, the payoff is represented as

$$J(\mathbf{x}(t_s), u, v) = \int_{t_s}^{t_f(\mathbf{x}(t_s),u,v)} d\bar{t} = t_f(\mathbf{x}(t_s), u, v) - t_s \tag{4}$$

Note that $t_f(\mathbf{x}(t_s), u, v)$ depends on the sequence of controls u and v applied to reach the point $\mathbf{x}(t_f)$ from the point $\mathbf{x}(t_s)$.

Definitions

Value of the game

For a given state of the system $\mathbf{x}(t_s)$, $V(\mathbf{x}(t_s))$ represents the outcome if the players implement their optimal strategies starting at the point $\mathbf{x}(t_s)$, and it is called the *value of the game* or the *value function* at $\mathbf{x}(t_s)$ [Isaacs65, Basar95]

$$V(\mathbf{x}(t_s)) = \min_{u(t)\in\widehat{U}} \max_{v(t)\in\widehat{V}} J(\mathbf{x}(t_s), u, v) \tag{5}$$

where $\widehat{U}$ and $\widehat{V}$ are the set of valid values for the controls at all time t. $V(\mathbf{x}(t))$ is defined over the entire state space.

Definitions

Open and closed-loop strategies

Let $\gamma_p(\mathbf{x}(t))$ and $\gamma_e(\mathbf{x}(t))$ denote the closed-loop strategies of the DDR and the evader, respectively, therefore $u(t) = \gamma_p(\mathbf{x}(t))$ and $v(t) = \gamma_e(\mathbf{x}(t))$.
A strategy pair $(\gamma_p^*(\mathbf{x}(t)), \gamma_e^*(\mathbf{x}(t)))$ is in closed-loop (saddle-point) equilibrium [Basar95] if

$$\begin{aligned} J(\gamma_p^*(\mathbf{x}(t)), \gamma_e(\mathbf{x}(t))) &\leq J(\gamma_p^*(\mathbf{x}(t)), \gamma_e^*(\mathbf{x}(t))) \\ &\leq J(\gamma_p(\mathbf{x}(t)), \gamma_e^*(\mathbf{x}(t))) \forall \gamma_p(\mathbf{x}(t)), \gamma_e(\mathbf{x}(t)) \end{aligned} \tag{6}$$

where J is the payoff of the game in terms of the strategies. An analogous relation exists for open-loop strategies.

Necessary Conditions for Saddle-Point Equilibrium Strategies

Theorem (Pontryagin's Maximum Principle - PMP)

Suppose that the pair $\{\gamma_p^*, \gamma_e^*\}$ *provides a saddle-point solution in closed-loop strategies, with* $\mathbf{x}^*(t)$ *denoting the corresponding state trajectory. Furthermore, let its open-loop representation* $\{u^*(t) = \gamma_p(\mathbf{x}^*(t)), v^*(t) = \gamma_e(\mathbf{x}^*(t))\}$ *also provide a saddle-point solution (in open-loop polices). Then there exists a costate function* $p(\cdot) : [0, t_f] \rightarrow R^n$ *such that the following relations are satisfied:*

$$\dot{\mathbf{x}}^*(t) = f(\mathbf{x}^*(t), u^*(t), v^*(t)),\ \mathbf{x}^*(0) = \mathbf{x}(t_s) \quad (7)$$

$$H(p(t), \mathbf{x}^*(t), u^*(t), v(t)) \leq H(p(t), \mathbf{x}^*(t), u^*(t), v^*(t)) \leq H(p(t), \mathbf{x}^*(t), u(t), v^*(t)) \quad (8)$$

$$p(t) = \nabla V(\mathbf{x}(t)) \quad (9)$$

$$\dot{p}^T(t) = -\frac{\partial}{\partial x} H(p(t), \mathbf{x}^*(t), u^*(t), v^*(t)) \ \textbf{(Adjoint Equation)} \quad (10)$$

$$p^T(t_f) = \frac{\partial}{\partial x} G(t_f, \mathbf{x}^*(t_f)) \ along \ \zeta(\mathbf{x}^*(t)) = 0 \quad (11)$$

where

$$H(p(t), \mathbf{x}(t), u(t), v(t)) = p^T(t) \cdot f(\mathbf{x}(t), u(t), v(t)) + L(\mathbf{x}(t), u(t), v(t)) \ \textbf{(Hamiltonian)} \quad (12)$$

and T *denotes the transpose operator.*

Pursuit-Evasion Games

Variants

Target capturing in an environment without obstacles

- R. Isaacs. Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization. John Wiley and Sons. Inc., 1965.
- Y.C. Ho et. al., Differential Games and Optimal Pursuit-Evasion Strategies, IEEE Transactions on Automatic Control, 1965.

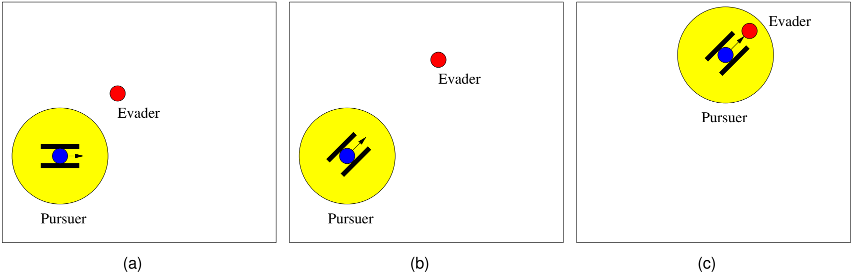


Figure: Target capturing.

Variants

Target tracking in an environment with obstacles

- S. M. LaValle et. al., "Motion strategies for maintaining visibility of a moving target", in Proc. IEEE Int. Conf. on Robotics and Automation, 1997.
- H.H. González-Baños et. al., Motion Strategies for Maintaining Visibility of a Moving Target In Proc IEEE Int. Conf. on Robotics and Automation, 2002.
- R. Murrieta-Cid et. al., Surveillance Strategies for a Pursuer with Finite Sensor Range, International Journal on Robotics Research, Vol. 26, No 3, pages 233-253, March 2007.
- S. Bhattacharya and S. Hutchinson , On the Existence of Nash Equilibrium for a Two Player Pursuit-Evasion Game with Visibility Constraints, The International Journal of Robotics Research, December, 2009.

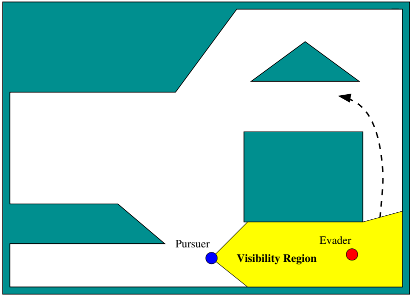


Figure: Target tracking.

Variants

Target finding in an environment with obstacles

- V. Isler et. al., “Randomized pursuit-evasion in a polygonal enviroment”, IEEE Transactions on Robotics, vol. 5, no. 21, pp. 864-875, 2005.
- R. Vidal et. al., “Probabilistic pursuit-evasion games: Theory, implementation, and experimental evaluation”, IEEE Transactions on Robotics and Automation, vol. 18, no. 5, pp. 662-669, 2002.

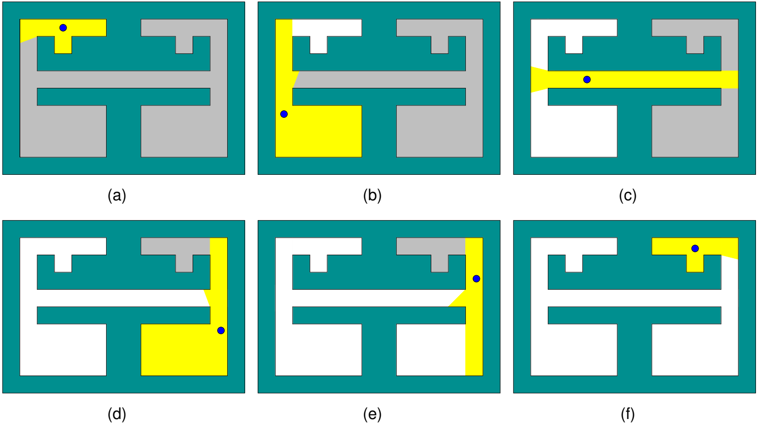
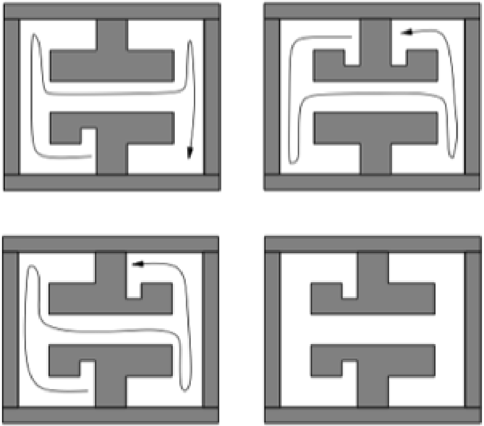


Figure: Target finding.

Target finding in an environment with obstacles

- L. Guibas, J.-C. Latombe, S. LaValle, D. Lin and R. Motwani, “Visibility-based pursuit-evasion in a polygonal environment”, International Journal of Computational Geometry and Applications, vol. 9, no. 4/5, pp. 471-494, 1999.



Thanks... Questions?

murrieta@cimat.mx

T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory, 2nd Ed*. SIAM Series in Classics in Applied Mathematics, Philadelphia, 1995.

R. Isaacs. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. Wiley, New York, 1965.

L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* JohnWiley, 1962.

P. Souères and J. D. Boissonnat, *Optimal Trajectories for Nonholonimic Mobile Robots*. J.P. Laumond, Editor, Springer, 1990.