

Differential Games & iLQG

Neuro-fuzzy Control Project



National Technical University of Athens
Electrical and Computer Engineering

Κόρδας Νικόλαος
Αριθμός Μητρώου: 03121032

Abstract

Contents

1	Introduction	2
1.1	Background & Tools	2
1.1.1	LQR Control Problem	2
1.1.2	General Sum Differential Games	3

1 Introduction

1.1 Background & Tools

1.1.1 LQR Control Problem

The Linear Quadratic Regulator problem is an optimal control problem concerning dynamic systems with linear dynamics and quadratic cost functions. The objective is to find the optimal control u that minimizes the cost J [2].

- A linear time-varying system:

$$\dot{x} = A(t)x + B(t)u, \quad x(t_0) = x_0, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m$$

- Quadratic cost function:

$$J(x, u, t) = \int_{t_0}^{t_f} (x^T(t)Qx(t) + u^T(t)Ru(t))dt + x^T(t_f)Q_fx(t_f),$$

$$Q = Q^T \geq 0, \quad Q_f = Q_f^T \geq 0, \quad R = R^T > 0$$

The running cost's matrices Q and R penalize the size of the state and the control effort respectively.

We derive the necessary conditions for the minimization of J by applying Pontryagin's Maximum Principle. This yields that the optimal control is a linear state feedback law of the form:

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

where P must be a solution of the matrix differential equation (Riccati differential equation (RDE)):

$$\dot{P}(t) = -P(t)A(t) - A^T(t)P(t) - Q(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t), \quad P(t_f) = Q_f$$

A special case of the LQR problem described above is the infinite-horizon LQR problem. Now, we assume that A , B , Q and R are constant (transient phenomena have subsided) and as $t_f \rightarrow \infty$ the terminal cost becomes negligible, thus $Q_f = 0$. So, the dynamics of the control system become time-invariant and the cost function becomes $J(x, u, t) = \int_{t_0}^{\infty} (x^T(t)Qx(t) + u^T(t)Ru(t))dt$. The optimal solution is given by:

1. $u^*(t) = -R^{-1}B^TPx^*(t)$
2. The limit $P = \lim_{t_f \rightarrow \infty} P(t_0, t_f)$ of the solution of the RDE is a constant matrix that satisfies the Algebraic Riccati Equation (ARE):

$$PA + A^TP + Q - PBR^{-1}B^TP = 0$$

3. The optimal cost is: $J(u^*) = x_0^TPx_0$
4. The closed-loop system $\dot{x}^* = (A - BR^{-1}B^TP)x^*$ is exponentially stable, as long as the couple (A, \sqrt{Q}) is detectable.

The theoretical foundation of the LQR problem is necessary for the understanding of dynamic environments featuring multiple agents. In classical optimal control, our objective is to minimize a unique cost function. In contrast, in multi-player games, the optimization process is coupled. Actions of one agent influence the performance and the state of the others. This transforms the Riccati equation to a system of coupled Riccati equations. Their solutions lead to a Nash equilibrium point, not a minimum. This framework is the basis of the iLQG algorithm which utilizes LQR iteratively by approximating non-linear dynamics as LQ problems.

1.1.2 General Sum Differential Games

Before developing a theoretical framework in the context of general-sum differential games, it is advisable to examine several simpler cases.

Zero-sum Matrix Games

To analyze multi-agent interactions, we first examine two-player zero-sum games, which are effectively represented in bi-matrix form [1]. Let P_1 and P_2 be the participants, choosing from m and n available strategies, respectively. The game is defined by a pair of matrices (A, B) , where each entry $(a_{ij} = J_1(u_1, u_2), b_{ij} = J_2(u_1, u_2))$ represents the cost incurred by P_1 and P_2 for a specific strategy pair (i, j) . In a zero-sum game, the players' interests are diametrically opposed, satisfying the condition $B = -A$. Consequently, for every outcome, the sum of their costs is $J_1(u_1, u_2) + J_2(u_1, u_2) = 0$, for every strategy pair (u_1, u_2) implying that any advantage gained by one agent results in an equivalent loss for the other.

P_1/P_2	a	b
c	2, -2	0, 0
d	4, -4	-4, 4

(a) Game 1

P_1/P_2	a	b
c	-2, 2	0, 0
d	2, -2	-2, 2

(b) Game 2

Table 1: Examples of zero-sum games

In table 1 above, matrices A and B are merged into a single cost matrix C . We suppose that P_1 can play either c or d and P_2 can play either a or b . Both P_1 and P_2 try to minimize their cost functions.

In the game in table 1a, P_2 should rationally play strategy a because it yields a lower cost for P_2 regardless of what P_1 plays. Knowing P_2 is rational and will play a , P_1 plays strategy c every time, choosing the lesser cost. Even if P_2 were not guaranteed to play a , P_1 's security strategy is to play c because that is the strategy that ensures the smallest cost (2). This "rational" solution is called a *saddle-point* solution.

In our second example (1b), there is no saddle-point solution. P_1 's security strategy is to play c (this strategy minimizes his maximum loss - 0). P_2 's security strategies are both a and b (same smallest maximum loss). However, if P_2 plays securely (b) and P_1 knows this, P_1 has an incentive to switch to b . Conversely, if P_1 plays c (security strategy) and P_2 knows this, P_2 has to choose b from his security strategies. Thus, unlike the first game, there is not a Nash equilibrium.

A formal definition of Nash equilibrium can be found in [3].

Suppose a game with N players. If $J_1(u_1, \dots, u_N), \dots, J_N(u_1, \dots, u_N)$ are their respective cost functions, then the strategy set $\{u_1^*, \dots, u_N^*\}$ is a Nash equilibrium if for every $i = 1, \dots, N$:

$$J_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*) \geq J_i(u_1^*, \dots, u_N^*)$$

According to [3], all Nash equilibria are equivalent in a zero-sum game. Thus, every Nash equilibrium's strategy set is interchangeable.

In zero-sum games, players' interests are diametrically opposed—one's gain is the other's loss—so there is no possibility of mutual gain. Consequently, the primary solution concept of interest is a Nash (or minimax) equilibrium, where each player's strategy is optimal given the other's.

Nonzero-sum Matrix Games

In this case, the sum of the cost functions is not zero for every possible outcome dictated by the strategies selected.

P_1/P_2	Not Betray	Betray
Not Betray	2, 2	10, 1
Betray	1, 10	5, 5

Table 2: Prisoner's Dilemma

Table 2 is the famous game Prisoner's Dilemma in bi-matrix form. Here, the cost functions correspond to years in prison for each player (P_1 and P_2). The only Nash equilibrium solution is $(u_1, u_2) = (\text{Betray}, \text{Betray})$. But, it is not optimal in the sense that $(u_1, u_2) = (\text{Not Betray}, \text{Not Betray})$ yields a better result for both players. This illustrates the possibility of mutual interest in nonzero-sum games. A solution produced by the coalition of all the players is called pareto-optimal solution [3].

In addition, in nonzero-sum games, every Nash equilibrium strategy is not equivalent to every other (cost-wise).

While the preceding discussion of matrix and nonzero-sum games provides a foundational understanding of strategic interaction in discrete-time, finite-action settings, many real-world problems involve dynamics, continuous time, and evolving states. To model strategic decision-making in such contexts—where players' actions influence not only immediate payoffs but also future states and opportunities—we extend the framework to differential games. These games generalize optimal control theory to multi-agent settings, incorporating differential equations to describe state evolution and enabling the analysis of continuous-time strategic behavior over a horizon. In what follows, we introduce the core formulation of differential games, including dynamics, cost functionals, information structures, and solution concepts such as open-loop and feedback Nash equilibria

References

- [1] Tamer Başar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory, 2nd Edition*. Society for Industrial and Applied Mathematics, 1998.
- [2] Daniel Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, Princeton, NJ, 2012.
- [3] Alan W. Starr and Y. C. Ho. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3:184–206, 1969.