

Solving Nonzero-Sum Nash Differential Game via Variation and Pseudospectral Methods

Guang-Yan XU^{1,a}, Biao ZHOU^{2,b,*}, Hong-Mei ZHANG^{3,c} and Dan ZHAO^{4,d}

^{1,3}School of Aerospace Engineering, Shenyang Aerospace University, Shenyang, 110136, China

^{2,4}School of Automation, Shenyang Aerospace University, Shenyang, 110136, China

^aguangyan_xu@163.com, ^bzhoubiaoxt@163.com, ^czhang.hongmei@163.com,

^dzhaodan19900406@163.com

*Corresponding author

Keywords: Nonzero-sum Nash game, Differential game, Multi-player, Legendre Pseudospectral Method, SNOPT.

Abstract. A numerical methods are presented in this paper to solve the nonzero-sum N-player Nash differential game. Variation methods are used to convert the original game into a regular optimal control problem which consists of one objective function, the state equations of all the players with initial conditions and the other necessary conditions derived. Then the later optimization problem is interpolated through the Legendre-pseudospectral method(LPM) and solved by applying SNOPT algorithm to get the optimal state trajectory. As an illustration, the air combat between a superior fighter and an inferior fighter is formulated as a nonzero-sum differential game. The states, traces, costs of both and the relative distance between them are displayed. The results show that numerical solutions converge to the saddle-points successfully, which show the feasibility and effectiveness of the proposed method in solving the nonzero-sum differential game.

Introduction

Over the past few decades, many researchers have studied a variety of the Nash differential games, especially the pursuit-evasion game[1] in depth, and a large number of valuable results have been achieved. In the qualitative analysis, a tail-chase air combat problem was solved by using differential dynamic programming[2]. A coplanar pursuit-evasion game in the atmosphere was studied with assuming a constant speed for evader[3]. A pursuit–evasion problem between missile and aircraft with realistic dynamics was solved by using an indirect, multiple shooting method[4,5]. Iterating a direct method[6] and nonlinear MPC method[7] was also applied for the pursuit-evasion game. Besides, in recently, a variety of evolutionary algorithms have also been used to solve this kind of differential game problem[8,9]. In addition, a named semi-direct method was introduced for pursuit–evasion dynamic game[10] and applied for the games between missiles[11] and spacecrafts[12].

This research concentrates on solving the nonzero-sum N-player differential game. This is the first time for that game to be solved accurately by the variation method and the LPM method[13,14,15]. In the beginning of the content, the conversation by variation method is deduced in details. Then a nonzero-sum pursuit–evasion game with two Bolza-type objective functions is taken as an illustration. Simulations are performed at last to support the conclusions.

Nonzero-sum Nash differential game and optimal control problem

In this section, the nonzero-sum n -player Nash differential game is converted to the optimal control problem. With regard to that game, every player possesses one objective function. All of the objective functions can be described as the kind of the Bolza-type. Thus, the game can be depicted as:

$$J_i = \Phi_i(\mathbf{x}_{1,f}, \dots, \mathbf{x}_{n,f}, t_f) + \int_{t_0}^{t_f} L_i(\mathbf{x}_i, \mathbf{u}_i, t) dt. \quad i \in \{1, \dots, n\} \quad (1)$$

subject to the uncoupled state equations:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i, t). \quad i \in \{1, \dots, n\} \quad (2)$$

with the initial conditions $\mathbf{x}_i(t_0) = \mathbf{x}_{i,0}$. To state the necessary conditions, the Hamiltonians of $n-1$ players except the first one are introduced as follows:

$$H_i(\mathbf{x}_i, \mathbf{u}_i, \boldsymbol{\lambda}_i, t) \triangleq L_i + \boldsymbol{\lambda}_i^T \mathbf{f}_i. \quad i \in \{2, \dots, n\} \quad (3)$$

The adjoint equations of the $n-1$ players are:

$$\dot{\boldsymbol{\lambda}}_i = -\frac{\partial H_i}{\partial \mathbf{x}_i} = -\frac{\partial L_i}{\partial \mathbf{x}_i} - \frac{\partial \mathbf{f}_i^T}{\partial \mathbf{x}_i} \boldsymbol{\lambda}_i. \quad i \in \{2, \dots, n\} \quad (4)$$

The boundary constraints of the adjoint variables mentioned in relationship are:

$$\boldsymbol{\lambda}_i(t_f) = \frac{\partial \Phi_i}{\partial \mathbf{x}_i(t_f)}. \quad i \in \{2, \dots, n\} \quad (5)$$

As unbounded control variables are assumed for the problem, the following first-order optimality and the second-order conditions must be satisfied by the control variables for the $n-1$ players:

$$\frac{\partial H_i}{\partial \mathbf{u}_i} = \frac{\partial L_i}{\partial \mathbf{u}_i} + \frac{\partial \mathbf{f}_i^T}{\partial \mathbf{u}_i} \boldsymbol{\lambda}_i = 0 \quad \text{and} \quad H_{\mathbf{u}_i, \mathbf{u}_i} = \frac{\partial^2 H_i}{\partial \mathbf{u}_i^2} \geq 0. \quad i \in \{2, \dots, n\} \quad (6)$$

According to Eq. 7, the controls of the second to nth player can be written as functions of \mathbf{x}_i , $\boldsymbol{\lambda}_i$, and t : $\mathbf{u}_i = \mathbf{u}_i(\mathbf{x}_i, \boldsymbol{\lambda}_i, t)$, for $i \in \{2, \dots, n\}$. After bringing them into Eq. 2, the original dynamic equations can be rewritten as:

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \mathbf{u}_1, t) \quad \dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i, \boldsymbol{\lambda}_i, t). \quad i \in \{2, \dots, n\} \quad (7)$$

Apparently, only one control variable \mathbf{u}_1 is remained for the system with that the new state vector is introduced as: $\mathbf{y} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T, \boldsymbol{\lambda}_2^T, \dots, \boldsymbol{\lambda}_n^T]^T \triangleq [\mathbf{y}_1^T, \dots, \mathbf{y}_{2n-1}^T]^T$. Up to now, the original nonzero-sum n -player Nash differential game can be equivalent to the following optimal control problem:

$$\min_{\mathbf{u}_1} J_1 = \Phi_1(\mathbf{y}_{1,f}, \dots, \mathbf{y}_{n,f}, t_f) + \int_{t_0}^{t_f} L_1(\mathbf{y}_1, \mathbf{u}_1, t) dt. \quad (8)$$

s.t.

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{u}_1, t). \quad (9)$$

$$\mathbf{y}_i(t_0) = \mathbf{x}_i(t_0) = \mathbf{x}_{i,0}. \quad i \in \{1, \dots, n\} \quad (10)$$

$$\mathbf{y}_{i+n-1}(t_f) = \boldsymbol{\lambda}_i(t_f) = \partial \Phi_i / \partial \mathbf{x}_i(t_f). \quad i \in \{2, \dots, n\} \quad (11)$$

Here, J_1 is retained as the ultimate objective function. It is worth noting that it is not the only option. Theoretically, any player's objective function can be selected as the final performance index. In specific issues, in order to easy the conversion, a good selection is such a player who has more constraints or whose constraints are difficult to deal with. Then, LPM is applied for converting the optimal problem into the NLP problem. In this method, the system dynamics are collocated at Legendre-Gauss-Lobatto (LGL) points. The integral term involved in the objective function is discretized based on Gauss-Lobatto quadrature rule. The boundary conditions can be used to assign

initial values of unknown parameters, or be discretized as equality or inequality constraints. The optimal control problem is thereby converted into a mathematical programming problem.

A pursuit-evasion instance

In this section, a point-mass model between an evasive fighter and a pursuing fighter is used to describe the flight trajectories in two-dimensional space. The geometry of planar pursuit defining the state variables of the game is depicted in Fig. 1:

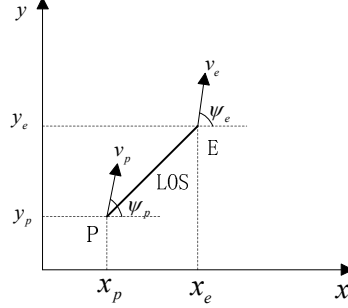


Fig. 1 Geometrical positions and angles of P and E

where the LOS is the line of sight from P to E ; ψ_i is the angle between the velocity vector v_i and the horizontal axis in the positive direction. Their velocity v_p and v_e are assumed as constants. (x_p, y_p) and (x_e, y_e) indicate their position in the two-dimensional plane. The dynamic system is governed by the following equations of motion:

$$\dot{x}_i = v_i \cos \psi_i \quad \dot{y}_i = v_i \sin \psi_i \quad \dot{\psi}_i = \omega_i \quad i \in \{p, e\} \quad (12)$$

in conjunction with the initial conditions:

$$x_i(t_0) = x_{i0} \quad y_i(t_0) = y_{i0} \quad \psi_i(t_0) = \psi_{i0} \quad (13)$$

That v_p is larger than v_e is assumed in order to match the real air combat. Two Bolza-type objective functions are:

$$J_p = \Phi + \int_{t_0}^{t_f} K_p \omega_p^2 dt \quad \text{and} \quad J_e = -\Phi + \int_{t_0}^{t_f} K_e \omega_e^2 dt \quad (14)$$

where, $\Phi = [(x_p(t_f) - x_e(t_f))^2 + (y_p(t_f) - y_e(t_f))^2] / 2$ indicates the relative distance index. The integral term is on fuel consumption penalties. Apparently, both of the players expect to decrease their cost functions, respectively. The Hamiltonian H_p is introduced as:

$$H_p = K_p \omega_p^2 + \lambda_1 v_p \cos \psi_p + \lambda_2 v_p \sin \psi_p + \lambda_3 \omega_p \quad (15)$$

where $\lambda_p = [\lambda_1 \quad \lambda_2 \quad \lambda_3]^T$ is the adjoint vector for P . The adjoint equations of pursuer are as follows:

$$\dot{\lambda}_1 = -\frac{\partial H_p}{\partial x_p} = 0 \quad \dot{\lambda}_2 = -\frac{\partial H_p}{\partial y_p} = 0 \quad \dot{\lambda}_3 = -\frac{\partial H_p}{\partial \psi_p} = \lambda_1 v_p \sin \psi_p - \lambda_2 v_p \cos \psi_p \quad (16)$$

According to the variation method, the values of the adjoint variables for P in the terminal time are:

$$\lambda_1(t_f) = x_p(t_f) - x_e(t_f) \quad \lambda_2(t_f) = y_p(t_f) - y_e(t_f) \quad \lambda_3(t_f) = 0 \quad (17)$$

And the control variable ω_p must satisfy the following conditions:

$$\frac{\partial H_p}{\partial \omega_p} = 2K_p \omega_p + \lambda_3 = 0 \quad \text{and} \quad \frac{\partial^2 H_p}{\partial \omega_p^2} = 2K_p \geq 0 \quad (18)$$

According to the Eq. 18, the optimal control for the player P is shown below:

$$\omega_p = -\lambda_3(t)/2K_p . \quad (19)$$

Eq. 19 is used to replace the ω_p in Eq. 12 of P . So far, the original nonzero-sum two-player game has been converted to the following optimal control problem:

$$V = \min_{\omega_E} J_E . \quad (20)$$

subject to Eq. 12 and Eq. 16 as differential equations, Eq. 13 and Eq. 17 as the boundary conditions, with the terminal time specified and terminal states free.

Experiments and Conclusions

The optimal control problem depicted in Eq. 20 is converted, by applying LPM method, into a NLP problem which is solved by an accurate commercial NLP solver, *snopt*. Simulations are performed on the *matlab* platform, including the *tomlab* toolbox. The state variables are normalized by a reference length of 50km and a reference angle of π , which avoids reducing the accuracy of the variables by keeping them to similar order and improves the robustness of the NLP problem solver. Constants are assigned as follows: $v_p = 0.3 \text{ km/s}$, $v_E = 0.2 \text{ km/s}$, $K_p = 100$, $K_E = 100$.

The terminal time is specified with 45 seconds. At the initial time, the pursuer is put at the (-5km, 0) in the plane coordinate system, and the evader is at (5km, 0); the pursuer flies along the horizontal axis in the positive direction, and the evader along the longitudinal axis toward the positive direction. Thus, the initial condition for the nominal case is $[x_{P0}, y_{P0}, \psi_{P0}, x_{E0}, y_{E0}, \psi_{E0}] = [-0.1, 0.0, 0.0, 0.1, 0.0, 0.5]$. The number of configuration nodes in LPM method is 26.

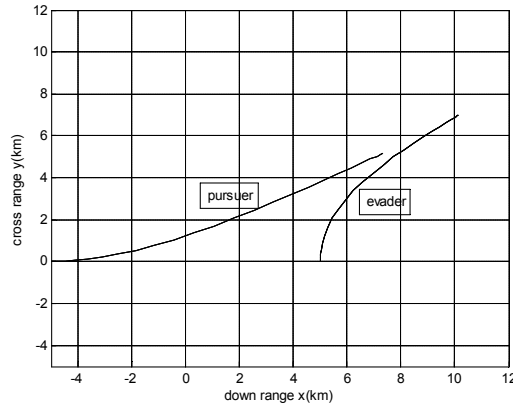


Fig. 2 Saddle-point trajectories of the two fighters

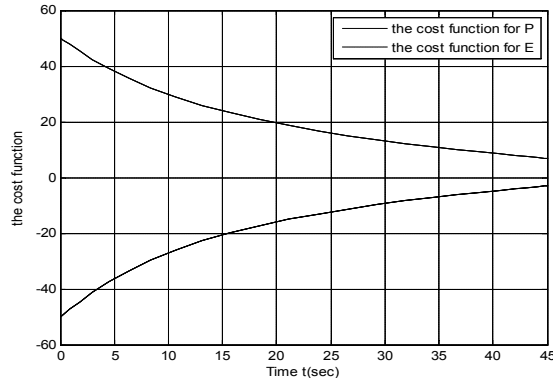


Fig. 3 Cost functions of the two fighters

Fig. 2 shows the saddle-point trajectories of the two fighters intuitively. It is easy to see that the angle of evader changes faster than the pursuer's at beginning, which indicates the evader is trying to

get rid of the pursuer by maneuvering in order to delay the interception. the pursuer tries to make its heading angle consistent with the evader's, i.e., the pursuer tends to be the behind of the evader. The cost functions of the two fighters are displayed in Fig. 3. The combined results just start to zero and then are greater than zero, which is because of the integral terms in the two cost functions and reflects that it is not a zero-sum game.

The results show the feasibility and effectiveness of the proposed method for a non-sum Nash differential game. As future work, the condition of game time free may be considered.

Acknowledgement

This research was financially supported by the National Science Foundation of China (61074159).

References

- [1] R. Isaacs, *Games of Pursuit*, Rand Corp., Santa Monica, 1951, P-257.
- [2] B. Jarmark, A.W. Merz, J. Breakwell, The Variable Speed Tail-Chase Aerial Combat Problem, *Journal of Guidance, Control, and Dynamics*. 4 (1981) 323–328.
- [3] M. Guelman, J. Shinar, A. Green, Qualitative study of a planar pursuit evasion game in the atmosphere, *Journal of Guidance, Control, and Dynamics*. 13 (1990) 1136–1142.
- [4] M.H. Breitner, H.J. Pesch, W. Grimm, Complex Differential Games of Pursuit–Evasion Type with State Constraints, Part 1: Necessary Conditions for Open-Loop Strategies, *Journal of Optimization Theory and Applications*. 78 (1993) 419–441.
- [5] M.H. Breitner, H.J. Pesch, W. Grimm, Complex Differential Games of Pursuit-Evasion Type with State Constraints, Part 2: Numerical Computation of Open-Loop Strategies, *Journal of Optimization Theory and Applications*. 78 (1993) 443–463.
- [6] T. Raivio, H. Ehtamo, Visual Aircraft Identification as a Pursuit–Evasion Game, *Journal of Guidance, Control, and Dynamics*. 23 (2000) 701–708.
- [7] S. Kang, H.J. Kim, M.J. Tahk, Aerial Pursuit-Evasion Game Using Nonlinear Model Predictive Guidance, *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario Canada, 2-5 August, 2010-7880.
- [8] Z. Li, X.J. Yu, Solution of Pursuit/Evasion Differential Games Using Genetic Algorithms, *World Congress on Intelligent Control and Automation*, ChongQing, China, 25-27 June, 2008.
- [9] H.Y. Shi, Z.Q. Shang, Study on a Solution of Pursuit-evasion Differential Game Based on Artificial Fish School Algorithm, *Chinese Control and Decision Conference*, XuZhou, China, 26-28 May, 2010.
- [10] K. Horie, B.A. Conway, Optimal Fighter Pursuit–Evasion Maneuvers Found via Two-Sided Optimization, *Journal of Guidance, Control, and Dynamics*. 29 (2006) 105–112.
- [11] M. Pontani, B.A. Conway, Optimal Interception of Evasive Missile Warheads: Numerical Solution of the Differential Game, *Journal of Guidance, Control, and Dynamics*. 13 (2008) 1111–1122.
- [12] M. Pontani, B.A. Conway, Numerical solution of the three-dimensional orbital pursuit-evasion game, *Journal of Guidance, Control, and Dynamics*. 32 (2009) 474–487.
- [13] G. Elnagar, M.A. Kazemi, M. Razzaghi, The Pseudospectral Legendre Method for Discretizing Optimal Control Problems, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*. 40 (1995) 1793-1796.

- [14] M. Ross, F. Fahroo, Pseudospectral Knotting Methods for Solving Optimal Control Problems, *Journal of Guidance, Control, and Dynamics*. 27 (2004) 397-405.
- [15] G.Y. Xu, C.M. Sun, Vibration suppressive control of flexible-joint space manipulators based on Legendre pseudospectral method, Chinese Control and Decision Conference, Changsha, China, 31 May-2 June, 2014.