

# 基于自适应最优控制的有限时间微分对策制导律

陈燕妮, 刘春生<sup>†</sup>, 孙景亮

(南京航空航天大学 自动化学院, 江苏 南京 211106)

**摘要:** 针对固定末端时刻拦截机动目标的制导系统, 本文首先构建了非线性有限时间微分对策框架, 将导弹拦截非线性系统的最优问题转化为一般非线性系统的最优控制问题, 并通过自适应动态规划算法(adaptive dynamic programming, ADP)获得近似最优值函数与最优控制策略. 为了有效实现该算法, 本文利用一个具有时变权值和激活函数的评价网络来逼近Hamilton-Jacobi-Isaacs(HJI)方程的解, 并在线更新. 通过李雅普诺夫法来证明本文提出的控制策略可保证闭环微分对策系统稳定性和评价网络权值近似误差的有界性. 最后给出一个非线性导弹拦截目标系统的仿真例子验证了该方法的可行性和有效性.

**关键词:** 非线性微分对策; 有限时间; 自适应动态规划; 制导律

**引用格式:** 陈燕妮, 刘春生, 孙景亮. 基于自适应最优控制的有限时间微分对策制导律. 控制理论与应用, 2019, 36(6): 877 – 884

DOI: 10.7641/CTA.2018.80019

## Finite-time differential guidance law based on adaptive optimal control

CHEN Yan-ni, LIU Chun-sheng<sup>†</sup>, SUN Jing-liang

(College of Automation Engineering, Nanjing Aeronautic and Astronautic University, Nanjing Jiangsu 211106, China)

**Abstract:** In this paper, the problem of intercepting a maneuvering target within a fixed final time is posed in a nonlinear finite-time differential game framework. Then, we convert the optimal problem of nonlinear guidance system into optimal control problem of general nonlinear system. For this system, the approximate optimal function and optimal control strategy are found by solving the finite-time differential game problem via adaptive dynamic programming (ADP) technique. To implement the algorithm effectively, the single critic network with time-varying weights and activation functions is constructed to estimate the solution of associated time-varying Hamilton-Jacobi-Isaacs (HJI) equation, and update it online. By utilizing the Lyapunov stability theorem, the closed-loop differential game system are proved to be stable and the estimation weight error of the critic network are proved to be uniformly ultimately bounded. Finally, a simulation of a nonlinear missile-target interception system shows the feasibility and effectiveness of the proposed method.

**Key words:** nonlinear differential games; finite-time; adaptive dynamic programming; guidance law

**Citation:** CHEN Yanni, LIU Chunsheng, SUN Jingliang. Finite-time differential guidance law based on adaptive optimal control. *Control Theory & Applications*, 2019, 36(6): 877 – 884

## 1 引言

导弹拦截目标问题是一类具有利益冲突双方之间的对抗问题, 代表着此类双方对抗型最优控制问题. 典型制导律大多基于比例制导律、最优制导律, 目标多为静止或低速运动, 对机动目标的情形研究较少. 在设计具有末端约束的制导律过程中, 研究最广泛的是最优制导律方法, 主要思想为将末端约束的制导律设计转化为最优控制问题<sup>[1]</sup>. 而相比最优控制理论,

微分对策理论在拦截机动目标上具有明显的优势, 并在多个领域获得广泛应用<sup>[2]</sup>. 导弹在拦截过程中, 会遇到燃料消耗导致飞行器质量变化、姿态快速机动等情况, 这些变化都会引起控制系统中参数的变化. 从拦截制导控制系统时间优化的角度来看, 使闭环系统有限时间收敛的控制方法才是时间最优的控制方法, 且有限时间控制可使得闭环系统具有良好的鲁棒性和抗扰动性能<sup>[3]</sup>. 文献[4–5]考虑利用连续有限时间控

收稿日期: 2018–01–08; 录用日期: 2018–09–04.

<sup>†</sup>通信作者. E-mail: liuchsh@nuaa.edu.cn; Tel.: +86 13814541683.

本文责任编辑: 潘泉.

南京航空航天大学研究生创新基地(实验室)开放基金项目(kfjj20170320, kfjj20170304), 中央高校基本科研业务费专项基金项目资助.

Supported by the Foundation of Graduate Innovation Center in NUA A (kfjj20170320, kfjj20170304) and the Fundamental Research Funds for the Central Universities.

制方法设计控制器. 文献[6]基于有限时间稳定性定理, 给出了有限时间稳定的高阶滑模导引规律, 但给出的方法比较繁琐. 动态规划可有效处理最优控制, 但在实际求解非线性系统最优控制问题时, 容易出现“维数灾”<sup>[7-8]</sup>的问题. 而自适应动态规划(adaptive dynamic programming, ADP)将动态规划与人工神经网络有机结合, 利用非线性函数拟合方法逼近动态规划的性能指标, 是解决复杂非线性系统最优控制的重要理论和方法<sup>[9-11]</sup>. 文献[12]针对有限时间不确定非线性系统, 在不含策略迭代的情况下利用 HJI 方程对零和双方进行优化设计. 文献[13]通过迭代算法求解有限时间随机线性二次型(linear quadratic, LQ)最优控制问题, 并给出了最优控制可解性的充分条件. 文献[14]基于二维平面导弹-目标相对运动模型, 利用有限时间稳定性理论<sup>[15]</sup>, 得到了视线角速率有限时间收敛的充分条件. 针对有限时间微分对策制导问题中难以避免的非线性 Hamilton-Jacobi-Isaacs(HJI)方程, 采用自适应的方法求解微分对策鞍点问题.

本文针对一类平面导弹拦截目标问题, 在微分对策理论框架下, 研究有限时间制导律. 首先将导弹-目标运动微分方程描述为非线性仿射方程, 利用微分对策理论, 得出追逃双方的一对纳什均衡解. 基于具有时变权值和激活函数的评价网络, 在线求解双方最优控制策略. 该算法的优点是避免了直接求解非线性偏微分方程, 满足了末端约束并实现了闭环系统有限时间内收敛.

## 2 有限时间微分对策框架

将导弹与目标之间的对抗看作是一种博弈策略问题, 系统的状态方程可表示为如下连续时间非线性仿射系统:

$$\dot{x} = f(x) + g(x)u(t) + k(x)v(t), \quad (1)$$

式中:  $x \in \mathbb{R}^n$  为状态向量;  $u(t) \in \mathbb{R}^m$ ,  $v(t) \in \mathbb{R}^k$  分别为导弹和目标的控制向量;  $f(x) \in \mathbb{R}^n$ ,  $g(x) \in \mathbb{R}^{n \times m}$  和  $k(x) \in \mathbb{R}^{n \times k}$  代表系统内部动力学系数矩阵.

**假设 1**  $f(x)$ ,  $g(x)$  和  $k(x)$  均为局部李普希兹连续的, 并且有  $f(0) = 0$ .

对于导弹和目标这样对抗的双方, 给出如下在  $[t_0, t_f]$  范围内的有限时间代价函数:

$$V(x, t_0) = \psi(x(t_f), t_f) + \int_{t_0}^{t_f} (Q(x) + u^T R_1 u - v^T R_2 v) dt, \quad (2)$$

式中:  $t_0$  和  $t_f$  分别为初始与末端时刻,  $\psi(x(t_f), t_f)$  为末端  $t_f$  时刻含有某些性能指标的末端约束,  $Q$  为包含状态的正半定函数,  $R_1, R_2$  为具有合适维数的对称正定矩阵.

当  $t_0 = t_f$  时, 微分对策末端约束代价函数表示为

$$V(x, t_f) = \psi(x(t_f), t_f). \quad (3)$$

**假设 2** 输入矩阵  $g(x)$  和  $k(x)$  有界, 即对所有  $x$ , 存在正常数  $g_M$  和  $k_M$  使得  $\|g(x)\| \leq g_M$ ,  $\|k(x)\| \leq k_M$ .

定义系统(1)的哈密顿函数为

$$H(x, u, v) = V_t + Q(x, t) + u^T R_1 u - v^T R_2 v + V_x^T (f(x) + g(x)u + k(x)v), \quad (4)$$

式中:  $V_t = \frac{\partial V(x, t)}{\partial t}$ ,  $V_x = \frac{\partial V(x, t)}{\partial x}$ , 其中  $V_t$  为有限时间相关项.

**注 1** 在固定末端时间的条件下, 代价函数的形式为时变代价函数. 因此, 与无限时间的情况<sup>[19]</sup>相比, 式(4)表示的哈密顿函数中出现了时间的显函数部分, 即为时变的哈密顿函数, 一般求解时不变哈密顿函数的方法则很难对其进行求解.

根据 Hamilton-Jacobi-Isaacs (HJI) 理论, 代价函数的最优解  $V^*(x)$  可通过求解下列 HJI 方程获得

$$0 = \max_v \min_u H(x, u, v, V^*(x)). \quad (5)$$

根据 Nash-Pontryagin 极小值原理, 若存在控制策略  $u = u^*$  和  $v = v^*$  满足

$$\begin{aligned} H(x, u^*, v, V^*(x, t)) &\leq \\ H(x, u^*, v^*, V^*(x, t)) &\leq H(x, u, v^*, V^*(x, t)), \end{aligned} \quad (6)$$

则称  $u^*$  和  $v^*$  为微分对策鞍点, 假设式(5)的鞍点存在且唯一, 那么上述微分对策问题的最优策略为

$$\begin{cases} u^*(x) = -\frac{1}{2} R_1^{-1} g^T(x) V_x^*, \\ v^*(x) = \frac{1}{2} R_2^{-1} k^T(x) V_x^*. \end{cases} \quad (7)$$

根据假设条件, 式(7)中的代价函数未知, 求解过程困难, 无法给出具体最优策略的解. 因此, 下面采用自适应动态规划的方法在线求解此类问题.

## 3 自适应动态规划求解鞍点

根据前馈神经网络的全局逼近性质, 时变代价函数可近似为

$$V(x, t) = W_c^T(t) \sigma_c(x, t_f - t) + \varepsilon(x, t), \quad (8)$$

其中: 理想权值向量  $W_c(t) \in \mathbb{R}^L$ ,  $L$  为隐含层神经元个数,  $\sigma_c(x, t_f - t) \in \mathbb{R}^L$  为有限时间激活函数,  $\varepsilon(x, t)$  为神经网络近似误差.

进一步, 末端约束代价函数可表示为

$$V(x, t_f) = W_c^T(t_f) \sigma_c(x(t_f), 0) + \varepsilon(x, t_f). \quad (9)$$

**假设 3** 理想权值  $W_c$  及其导数  $\dot{W}_c$  有界, 即  $\|W_c\| \leq W_M$ ,  $\|\dot{W}_c\| \leq W_N$ , 其中:  $W_M > 0$ ,  $W_N > 0$ .

激活函数及其变化律局部有界, 即  $\|\sigma_c(x)\| < \sigma_M$ ,  $\|\nabla_x \sigma_c(x)\| < \sigma_{xM}$ . 存在正常数  $b_\varepsilon$ , 使得对所有  $x$  有  $\|\varepsilon(x)\| \leq b_\varepsilon$ ,  $\|\nabla_x \varepsilon(x)\| \leq b_{\varepsilon_x}$ , 其中  $b_{\varepsilon_x} > 0$ <sup>[17]</sup>.

$V(x, t)$ 对 $x$ 和 $t$ 的偏导可以分别表示为

$$V_x = \nabla_x^T \sigma_c(x, t_f - t) W_c(t) + \nabla_x \varepsilon(x, t), \quad (10)$$

$$V_t = \dot{W}_c^T(t) \sigma_c(x, t_f - t) + \nabla_t^T \sigma_c(x, t_f - t) W_c(t) + \nabla_t \varepsilon(x, t), \quad (11)$$

式中:

$$\nabla_x \sigma_c(x, t_f - t) = \frac{\partial \sigma_c(x, t_f - t)}{\partial x},$$

$$\nabla_x \varepsilon(x, t) = \frac{\partial \varepsilon(x, t)}{\partial x},$$

$$\nabla_t \sigma_c(x, t_f - t) = \frac{\partial \sigma_c(x, t_f - t)}{\partial t},$$

$$\nabla_t \varepsilon(x, t) = \frac{\partial \varepsilon(x, t)}{\partial t}.$$

那么, 最优策略可表示为

$$\begin{cases} u^* = \\ -\frac{1}{2} R_1^{-1} g^T(x) \nabla_x^T \sigma_c(x, t_f - t) W_c(t) + \varepsilon_{u^*}, \\ v^* = \frac{1}{2} R_2^{-1} k^T(x) \nabla_x^T \sigma_c(x, t_f - t) W_c(t) + \varepsilon_{v^*}, \end{cases} \quad (12)$$

其中:

$$\varepsilon_{u^*} = -\frac{1}{2} R_1^{-1} g^T(x) \nabla_x \varepsilon(x, t),$$

$$\varepsilon_{v^*} = \frac{1}{2} R_2^{-1} k^T(x) \nabla_x \varepsilon(x, t).$$

HJI方程可替换为

$$\begin{aligned} H(x, W_c, u^*, v^*) = & \dot{W}_c^T(t) \sigma_c(x, t_f - t) + \nabla_t \sigma_c(x, t_f - t) W_c(t) + \\ & Q(x, t) + W_c^T(t) \nabla_x \sigma_c(x, t_f - t) f(x) - \\ & \frac{1}{4} W_c^T(t) G W_c(t) + \frac{1}{4} W_c^T(t) K W_c(t) + \\ & \varepsilon_{HJI} = 0, \end{aligned} \quad (13)$$

式中:

$$G = \nabla_x \sigma_c(x, t_f - t) g(x) R_1^{-1} g^T(x) \nabla_x^T \sigma_c(x, t_f - t),$$

$$K = \nabla_x \sigma_c(x, t_f - t) k(x) R_2^{-1} k^T(x) \nabla_x^T \sigma_c(x, t_f - t),$$

$\varepsilon_{HJI}$ 为神经网络近似误差产生的残余误差,

$$\begin{aligned} \varepsilon_{HJI} = & \nabla_t \varepsilon(x, t) + \nabla_x^T \varepsilon(x, t) (f(x) + \\ & g(x)u + k(x)v). \end{aligned} \quad (14)$$

理想权值通常未知, 不可应用于实际控制过程中, 因此利用一个评价网络来近似有限时间代价函数,

$$\hat{V}(x, t) = \hat{W}_c^T(t) \sigma_c(x, t_f - t), \quad (15)$$

末端约束即为

$$\hat{V}(x, t_f) = \hat{W}_c^T(t_f) \sigma_c(\hat{x}(t_f), 0), \quad (16)$$

其中:  $\hat{V}(x, t)$ 为 $V(x, t)$ 近似代价函数,  $\hat{V}(x, t_f)$ 为末端时刻 $t_f$ 时的近似代价函数,  $\sigma_c(\hat{x}(t_f), 0) \in P^L$ 为近似状态 $\hat{x}_{t_f}$ 的激活函数,  $\hat{x}_{t_f}$ 可由系统动力学和实时状态计算得到. 同时, 定义权值估计误差为

$$\tilde{W}_c(t) = W_c(t) - \hat{W}_c(t), \quad (17)$$

则有

$$\dot{\tilde{W}}_c(t) = \dot{W}_c(t) - \dot{\hat{W}}_c(t). \quad (18)$$

**注 2** 现有文献中对逼近函数的设定, 理想权值多为常值<sup>[17-19]</sup>. 本文中考虑时变权值, 与时不变权值相比, 更符合神经网络的固有性质, 可更为准确地逼近代价函数, 得到更准确的最优策略解, 使导弹在拦截过程中可以更好地适应快速变化的环境.

类似地, 有

$$\hat{V}_x = \nabla_x^T \sigma_c(x, t_f - t) \hat{W}_c(t), \quad (19)$$

$$\begin{aligned} \hat{V}_t = & \dot{\hat{W}}_c^T(t) \sigma_c(x, t_f - t) + \nabla_t^T \sigma_c(x, t_f - t) \hat{W}_c(t), \end{aligned} \quad (20)$$

近似最优控制策略即为

$$\begin{cases} \hat{u}(x) = -\frac{1}{2} R_1^{-1} g^T(x) \nabla_x^T \sigma_c(x, t_f - t) \hat{W}_c(t), \\ \hat{v}(x) = \frac{1}{2} R_2^{-1} k^T(x) \nabla_x^T \sigma_c(x, t_f - t) \hat{W}_c(t), \end{cases} \quad (21)$$

可得近似哈密顿函数为

$$\begin{aligned} H(x, \hat{W}_c, \hat{u}, \hat{v}) = & \dot{\hat{W}}_c^T(t) \sigma_c(x, t_f - t) + \hat{W}_c^T(t) \nabla_t \sigma_c(x, t_f - t) + \\ & \hat{W}_c^T(t) \nabla_x \sigma_c(x, t_f - t) (f(x) + g(x)\hat{u} + k(x)\hat{v}) + \\ & Q(x, t) + \hat{u}^T R_1 \hat{u} - \hat{v}^T R_2 \hat{v} = e_c. \end{aligned} \quad (22)$$

为了解决有限时间微分对策问题, 同时需要妥当处理时变代价函数和末端约束, 基于神经网络的全局逼近性质和末端估计约束, 定义末端约束估计误差为

$$\begin{aligned} e_{t_f} = & \psi(x(t_f), t_f) - \hat{W}_c^T(t_f) \sigma_c(\hat{x}(t_f), 0) = \\ & W_c^T(t_f) \sigma_c(x(t_f), 0) - \\ & \hat{W}_c^T(t_f) \sigma_c(\hat{x}(t_f), 0) + \varepsilon(x, t_f) = \\ & W_c^T(t_f) \tilde{\sigma}_c(x(t_f), 0) + \\ & \tilde{W}_c^T(t_f) \sigma_c(\hat{x}(t_f), 0) + \varepsilon(x, t_f), \end{aligned} \quad (23)$$

其中 $\tilde{\sigma}_c(x(t_f), 0) = \sigma_c(x(t_f), 0) - \sigma_c(\hat{x}(t_f), 0)$ .

为在系统轨迹上最小化哈密顿函数和末端约束误差, 可以在满足末端约束条件下实现最优性, 定义如下能量函数:

$$E = \frac{1}{2} e_c^T e_c + \frac{1}{2} e_{t_f}^T e_{t_f}. \quad (24)$$

设计神经网络权值更新律为

$$\begin{aligned} \dot{W}_c(t) = & \\ -\sigma \hat{W}_c(t) - \frac{\hat{\omega}}{(1 + \hat{\omega}^T \hat{\omega})^2} e_c - \frac{\hat{\xi}}{(1 + \hat{\xi}^T \hat{\xi})^2} e_{t_f} = & \\ -\sigma \hat{W}_c(t) - \frac{\bar{\beta}_1}{m_s} e_c - \frac{\bar{\beta}_2}{m_t} e_{t_f}, & \quad (25) \end{aligned}$$

式中:

$$\begin{aligned} \sigma = \text{diag}\{\sigma_1, \dots, \sigma_n\}, \\ \hat{\omega} = \nabla_t \sigma_c(x, t_f - t) + \nabla_x \sigma_c(x, t_f - t)(f(x) + \\ g(x)\hat{u} + k(x)\hat{v}), \\ \hat{\xi} = -\sigma_c(\hat{x}(t_f), 0), \bar{\beta}_1 = \frac{\hat{\omega}}{1 + \hat{\omega}^T \hat{\omega}}, \\ m_s = 1 + \hat{\omega}^T \hat{\omega}, \\ \bar{\beta}_2 = \frac{\hat{\xi}}{1 + \hat{\xi}^T \hat{\xi}}, m_t = 1 + \hat{\xi}^T \hat{\xi}. \end{aligned}$$

**注3** 在设计更新律过程中考虑3部分, 第1项包含带系数的时变权值, 有利于权值更新律更快更准确地收敛到稳定值, 第2项和第3项可使近似哈密顿函数和末端约束误差最小化.

### 4 稳定性分析

**定理1** 考虑式(1)描述的非线性连续时间微分对策问题以及式(2)对应的代价函数, 基于假设1-3采取式(21)中的微分对策策略, 给出式(25)的评价权值更新律. 那么, 闭环系统状态 $x(t)$ 以及评价网络权值近似误差 $\tilde{W}_c$ 均为最终一致有界.

**证** 考虑如下李雅普诺夫函数:

$$L(t) = V(x, t) + \frac{1}{2} \tilde{W}_c^T(t) \tilde{W}_c(t). \quad (26)$$

对时间求偏导得

$$\begin{aligned} \dot{L}(t) = \dot{V}(x, t) + \tilde{W}_c^T(t) \dot{\tilde{W}}_c(t) = & \\ -Q(x, t) - \frac{1}{4} W_c^T(G - K)W_c + \varepsilon_0(x, t) - & \\ \varepsilon_{\text{HJI}} + \tilde{W}_c^T \frac{\bar{\beta}_2}{m_t} \varepsilon(x, t_f) + \tilde{W}_c^T \dot{W}_c + \sigma \tilde{W}_c^T W_c + & \\ \frac{1}{2} \tilde{W}_c^T(G - K)W_c + \tilde{W}_c^T \tilde{\sigma}_c(x(t_f), 0) \frac{\bar{\beta}_2^T}{m_t} W_c - & \\ \sigma \tilde{W}_c^T \tilde{W}_c + \tilde{W}_c^T \sigma_c(\hat{x}(t_f), 0) \frac{\bar{\beta}_2^T}{m_t} \tilde{W}_c + & \\ \frac{1}{2} \tilde{W}_c^T \frac{\bar{\beta}_1}{m_s} W_c^T(G - K)\tilde{W}_c - & \\ \frac{1}{4} \tilde{W}_c^T \frac{\bar{\beta}_1}{m_s} \tilde{W}_c^T(G - K)\tilde{W}_c, & \quad (27) \end{aligned}$$

式中:

$$\begin{aligned} \varepsilon_0(x, t) = & \\ \nabla_t^T \varepsilon(x, t) + \nabla_x^T \varepsilon(x, t)(f(x) + g(x)\hat{u} + k(x)\hat{v}), & \end{aligned}$$

且令

$$\begin{aligned} A = G - K, B = \tilde{\sigma}_c(x(t_f), 0) \frac{\bar{\beta}_2^T}{m_t}, \\ C = \sigma_c(\hat{x}(t_f), 0) \frac{\bar{\beta}_2^T}{m_t}, D = \frac{\bar{\beta}_1}{m_s} W_c^T(G - K), \\ M = \frac{\bar{\beta}_1^T}{m_s}(G - K), N = \frac{\bar{\beta}_2^T}{m_t} \varepsilon(x, t_f). \end{aligned}$$

由于 $Q(x, t)$ 为对称的正半定函数, 那么则存在一个正常数 $q$ 使得对任意 $x$ 都有 $x^T q x < Q(x, t)$ .

通过运用杨氏不等式, 存在非零常数 $\mu_1, \phi_1, \phi_2, \phi_3, \phi_4, \eta_1, \eta_2, \eta_3, \eta_4$ 使得如下不等式成立:

$$\begin{aligned} \dot{L} \leq & \\ -\frac{1}{8} \mu_1^2 \|\tilde{W}_c\|^4 + \frac{1}{4} [(2\phi_1^2 + 2\sigma\phi_2^2 + \phi_3^2 + 2\phi_4^2) + & \\ (-2\sigma\eta_1^2 + 2\eta_2^2 + \eta_3^2 + 2\eta_4^2) + (-\frac{2\sigma}{\eta_1^2} + \frac{2}{\eta_2^2} \lambda_{3M} + & \\ \frac{1}{\eta_3^2} \lambda_{4M}) - \frac{1}{2\mu_1^2} \lambda_{5m}] \|\tilde{W}_c\|^2 + \frac{1}{4} (\frac{2\sigma}{\phi_2^2} + \frac{1}{\phi_3^2} \lambda_{1M} + & \\ \frac{2}{\phi_4^2} \lambda_{2M}) \|W_c\|^2 - x^T q x - \frac{1}{4} W_c^T(G - K)W_c + & \\ \varepsilon_0(x, t) - \varepsilon_{\text{HJI}} + \frac{1}{2\phi_1^2} \|\dot{W}_c\|^2 + \frac{1}{2\eta_4^2} \lambda_{6M}. & \quad (28) \end{aligned}$$

**注4** 根据假设1-3易知, 存在正常数 $\lambda_{1m}, \lambda_{1M}, \lambda_{2m}, \lambda_{2M}, \lambda_{3m}, \lambda_{3M}, \lambda_{4m}, \lambda_{4M}, \lambda_{5m}, \lambda_{5M}, \lambda_{6m}, \lambda_{6M}$ 分别满足

$$\begin{aligned} \lambda_{1m} \leq \|A\| \leq \lambda_{1M}, \lambda_{2m} \leq \|B\| \leq \lambda_{2M}, \\ \lambda_{3m} \leq \|C\| \leq \lambda_{3M}, \lambda_{4m} \leq \|D\| \leq \lambda_{4M}, \\ \lambda_{5m} \leq \|M\| \leq \lambda_{5M}, \lambda_{6m} \leq \|N\| \leq \lambda_{6M}. \end{aligned}$$

经过整理化简, 可得

$$\dot{L} \leq -\frac{1}{8} \mu_1^2 \|\tilde{W}_c\|^4 + \mu_2 \|\tilde{W}_c\|^2 + \mu_3, \quad (29)$$

式中:

$$\begin{aligned} \mu_2 = \frac{1}{4} [(2\phi_1^2 + 2\sigma\phi_2^2 + \phi_3^2 + 2\phi_4^2) + (-2\sigma\eta_1^2 + & \\ 2\eta_2^2 + \eta_3^2 + 2\eta_4^2) + (-\frac{2\sigma}{\eta_1^2} + \frac{2}{\eta_2^2} \lambda_{3M} + & \\ \frac{1}{\eta_3^2} \lambda_{4M}) - \frac{1}{2\mu_1^2} \lambda_{5m}], \\ \mu_3 = \frac{1}{4} (\frac{2\sigma}{\phi_2^2} + \frac{1}{\phi_3^2} \lambda_{1M} + \frac{2}{\phi_4^2} \lambda_{2M}) \|W_c\|^2 + & \\ \frac{1}{2\phi_1^2} \|\dot{W}_c\|^2 - x^T q x - \frac{1}{4} W_c^T(G - K)W_c + & \\ \varepsilon_0(x, t) - \varepsilon_{\text{HJI}} + \frac{1}{2\eta_4^2} \lambda_{6M}. & \end{aligned}$$

当 $\tilde{W}_c$ 满足式(30)时,

$$\|\tilde{W}_c\| \geq \sqrt{\frac{4\mu_2 + \sqrt{8\mu_1^2 \mu_3 + 16\mu_2^2}}{\mu_1^2}}, \quad (30)$$

$\dot{L} < 0$ 成立.

根据Lyapunov理论可得权值估计误差 $\tilde{W}_c$ 能达到最终一致有界, 该算法具有良好的稳定性. 证毕.

**注5** 在式(30)中,  $\mu_1$ 及 $\mu_2, \mu_3$ 中包含的参数均为可调参数, 通过适当调参, 例如选择尽量大的 $\mu_1$ , 足够小的 $\phi_1, \phi_2, \phi_3, \phi_4$ 及足够大的 $\eta_1$ 等等, 即可使 $\mu_2, \mu_3$ 尽可能的小, 因此 $\|\tilde{W}_c\|$ 即可尽可能的小. 在满足式(30)后, 即有 $\dot{L} < 0$ 成立, 此时存在一个常数 $\epsilon$ ,  $\|\tilde{W}_c\|$ 将无限逼近它, 达到最终一致有界.

## 5 制导律设计

弹目运动关系如图1所示:  $V_M$ 和 $V_T$ 分别为导弹、目标的速度,  $\alpha$ 和 $\beta$ 分别为导弹、目标的航迹角,  $u$ 和 $v$ 分别为导弹、目标垂直于速度向量的控制输入. 假设追逃双方速度恒定,  $\theta$ 为视线角,  $\sigma$ 为视线角速率 $\dot{\theta}$ ;  $r$ 为弹目双方沿着视线角间的距离,  $V_r$ 为弹目相对速率.

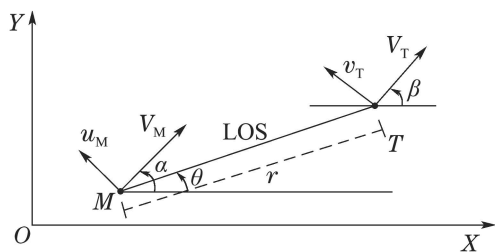


图1 拦截轨迹图

Fig. 1 Engagement trajectories

弹目相对运动方程为

$$\begin{cases} V_r = \dot{r} = V_T \cos(\beta - \theta) - V_M \cos(\alpha - \theta), \\ \sigma = \dot{\theta} = (V_T \sin(\beta - \theta) - V_M \sin(\alpha - \theta))/r, \\ \dot{\alpha} = u/V_M, \\ \dot{\beta} = v/V_T. \end{cases} \quad (31)$$

导弹运动学方程为

$$\begin{cases} \dot{x}_M = V_M \cos \alpha, \\ \dot{y}_M = V_M \sin \alpha, \\ \dot{\alpha} = a_M/V_M, \\ \dot{a}_M = (u - a_M)/\tau_M, \end{cases} \quad (32)$$

其中:  $(x_M, y_M)$ 为导弹在惯性参考坐标系下的位置坐标,  $a_M$ 为其横向加速度,  $\tau_M$ 为其自动驾驶仪的时间常数.

拦截目标运动学方程为

$$\begin{cases} \dot{x}_T = V_T \cos \beta, \\ \dot{y}_T = V_T \sin \beta, \\ \dot{\beta} = a_T/V_T, \\ \dot{a}_T = (v - a_T)/\tau_T, \end{cases} \quad (33)$$

其中:  $(x_T, y_T)$ 为目标在惯性参考坐标系下的位置坐

标,  $a_T$ 为其横向加速度,  $\tau_T$ 为其自动驾驶仪的时间常数.

为了获得更好的拦截效果, 导弹需要选择其控制策略 $u$ 使得终端零效脱靶量最小.

### 定义1 零效脱靶量.

导弹和目标从给定的时间 $t$ 起不施加任何控制, 以该瞬时参数飞行至命中, 所产生的脱靶量. 根据文献[19], 得零效脱靶量计算公式为

$$r_{\text{miss}}(t) = \frac{r^2 \sigma}{\sqrt{V_r^2 + r^2 \sigma^2}}. \quad (34)$$

从上式可以看出, 当 $\sigma$ 趋于零时 $r_{\text{miss}}(t)$ 趋于零值, 且满足 $V_r < 0$ 即可成功拦截. 因此, 在末端制导过程中, 视线角速率 $\sigma$ 和接近速率 $V_r$ 需要满足下列条件:

$$\sigma \rightarrow 0, V_r < 0. \quad (35)$$

事实上, 制导过程一直持续到 $V_r$ 变为正, 在其符号转变的瞬间, 制导过程结束, 最终脱靶距离即确定.

在现有的制导问题中, 基于制导律的考虑选择 $\sigma$ 作为系统状态变量, 通过 $\sigma$ 对时间求导来构建非线性动力学框架

$$\dot{\sigma} = -2 \frac{V_r}{r} \sigma - \frac{\cos(\alpha - \theta)}{r} u + \frac{\cos(\beta - \theta)}{r} v. \quad (36)$$

**注6** 上述系统模型直接的微分对策策略并不容易获得, 相对距离 $r$ 出现在分母中, 并随时间单调递减. 此外, 式(36)本质上是一个有限时间拦截问题, 并不满足局部李普希兹条件.

### 定义2 剩余飞行时间.

记剩余时间为 $t_{\text{go}} = -r/\dot{r}$ , 新系统状态变量定义为 $x = [x_1 \ x_2]^T = [\theta \ \sigma t_{\text{go}}]^T$ , 并有如下新动力学方程:

$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = x_2 + x_2^3 + \frac{1}{\dot{r}} (\cos(\alpha - \theta) + x_2 \sin(\alpha - \theta)) \bar{u} - \frac{1}{\dot{r}} (\cos(\beta - \theta) + x_2 \sin(\beta - \theta)) \bar{v}, \end{cases} \quad (37)$$

式中 $\bar{t} = \ln(r(0)) - \ln(r(t))$ .

导弹和目标的新控制输入分别为

$$\begin{cases} \bar{u} = t_{\text{go}} u, \\ \bar{v} = t_{\text{go}} v. \end{cases} \quad (38)$$

**注7** 转化后系统满足了 $\bar{t} \rightarrow \infty$ 且 $r \rightarrow 0$ , 因此满足了局部李普希兹条件, 式(36)中的有限时间拦截问题已转化为无限时间问题.

**注8** 由式(37)可得,  $|(\alpha - \theta)| = \pi/2, |(\beta - \theta)| = \pi/2$

是不稳定点. 则制导律可应用域为

$$\Phi = \{x : |(\alpha - \theta)| \neq \pi/2, |(\beta - \theta)| \neq \pi/2, r \neq 0, V_r < 0\}.$$

由此, 制导问题就转化为了对非线性系统的控制问题. 进一步的, 控制输入增益函数 $g(x)$ 和 $k(x)$ 都有界, 满足假设2. 因此, 本文所提出的微分对策策略可达到预期的效果, 导弹可以成功拦截目标.

### 6 仿真

本节以飞行器目标拦截为例进行仿真验证. 导弹初始参数: 惯性参考坐标系中初始位置坐标 $(x_M, y_M) = (0, 0)$ , 速度 $V_M = 600 \text{ m/s}$ , 飞行航向角 $\alpha_0 = 60^\circ$ , 自动驾驶时间常数 $\tau_M = 0.1 \text{ s}$ ; 目标初始参数: 参考惯性坐标系中初始位置坐标 $(x_T, y_T) = (5000, 0)$ , 速度 $V_T = 400 \text{ m/s}$ , 飞行航向角 $\beta_0 = 70^\circ$ , 自动驾驶时间常数 $\tau_T = 0.1 \text{ s}$ . 在制导过程中, 拦截机与目标的速度视为常数. 为实施微分博弈策略, 设时间依赖项 $Q(x, t) = 20(x_1^2 + x_2^2 + \tau^2)$ , 其中:  $\tau = t_f - t$ , 末端时间 $t_f = 50 \text{ s}$ ; 控制参数 $R_1 = 5, R_2 = 9$ ; 末端约束 $\psi(x(t_f), t_f) = x^T(t_f)x(t_f)$ . 评价网络激活函数选择为

$$\sigma_c(x, t_f - t) = [x_1^2 \exp(-\tau) \quad x_2^2 \exp(-\tau) \quad x_1 x_2 \tau \quad x_1^4 \exp(-\tau) \quad x_2^2 \exp(-\tau) \quad x_1^3 x_2 \quad x_1^2 x_2^2 \quad x_1 x_2^2]^T,$$

神经元个数 $L = 8$ , 评价网络权值 $\hat{W}_c = [\hat{W}_{c1} \quad \hat{W}_{c2} \quad \dots \quad \hat{W}_{c8}]^T$ . 为保证达到良好的持续激励的效果, 在仿真的前20s为控制输入加入一个微小激励信号:  $n(t) = \sin^5 t \cos t + \sin^5(2t) \cos(0.2t)$ .

仿真结果如图2-5所示.

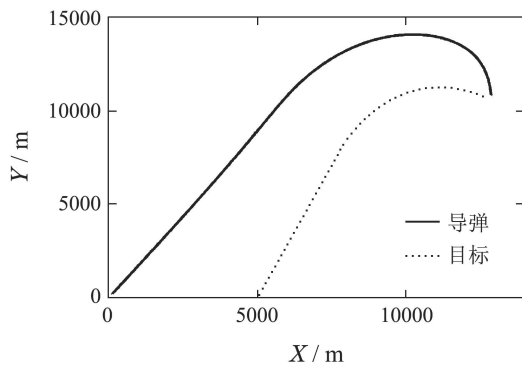


图2 拦截轨迹图

Fig. 2 Engagement trajectories

对弹目拦截非线性系统执行上述算法. 图2清晰地展示了整个拦截过程. 从图3可看出, 导弹和目标间的相对距离小于0.5 m, 导弹足以摧毁目标. 如图4所示, 视线角速率可在有限时间内(如 $t \in [0, 50 \text{ s}]$ )控制在零的领域内. 如图5所示, 整个拦截过程中, 沿视线角方

向的速度都维持在负, 即 $V_r < 0$ , 两者都满足上文所提拦截条件, 视导弹成功捕获机动目标. 同时, 零效脱靶量会减小到一个非常小的值.

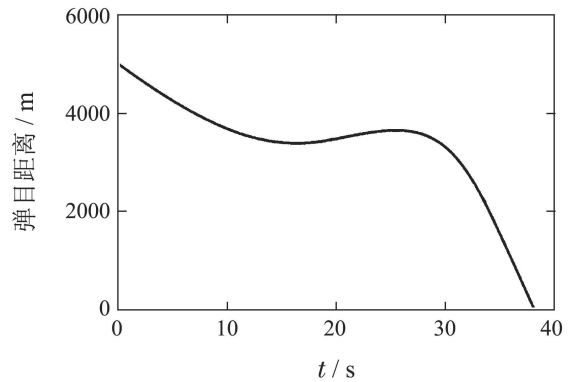


图3 弹目相对距离

Fig. 3 The relative distance

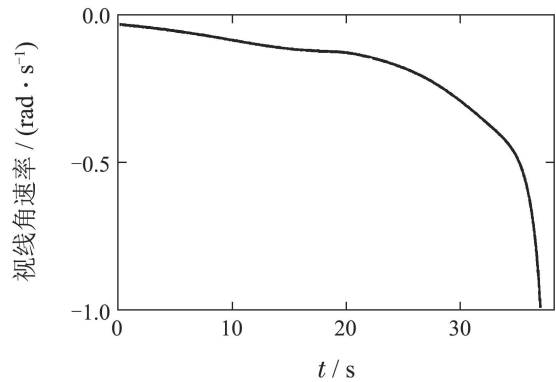


图4 视线角速率

Fig. 4 The LOS angular rate

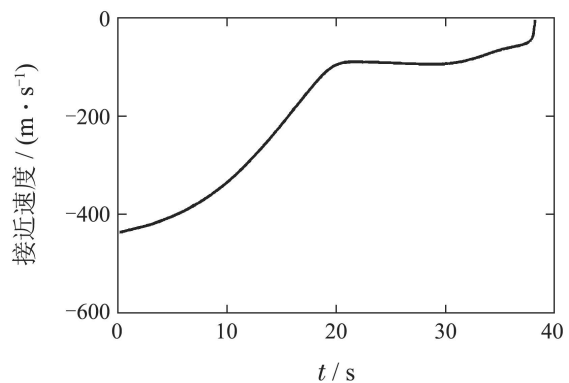


图5 弹目接近速率

Fig. 5 The range rate along the LOS

图6-7分别为导弹和目标的横向加速度. 图8为评价网络权值的变化轨迹, 证明了此评价网络的有效性. 在实际拦截末时刻处, 弹目相对距离减小到一个非常小的值, 寻的器失效, 控制器停止工作. 因此图6-7中的制导末时刻, 系统参数值发生突变. 此时, 导弹将根据惯性飞行拦截目标.

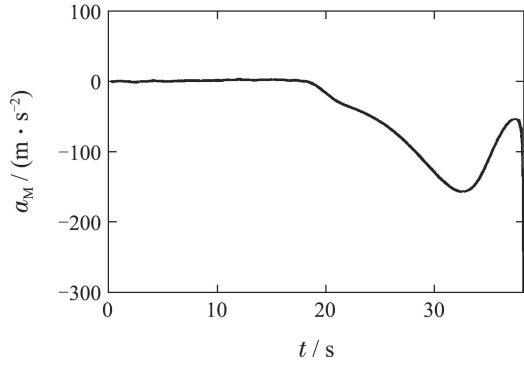


图6 导弹横向加速度

Fig. 6 The lateral acceleration of missile

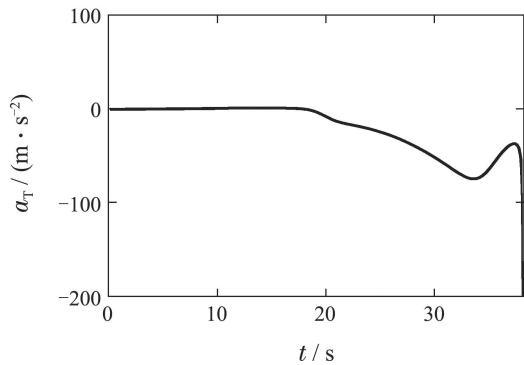


图7 目标横向加速度

Fig. 7 The lateral acceleration of target

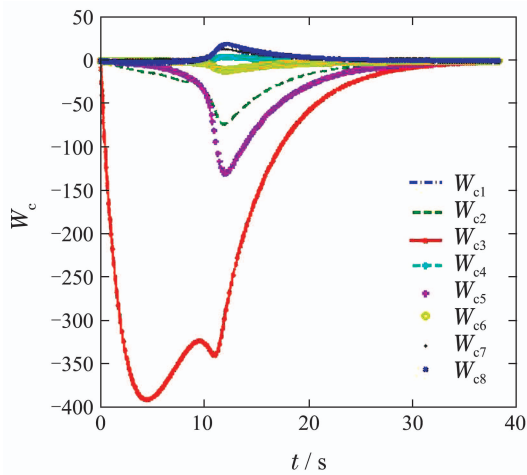


图8 评价网络权值

Fig. 8 Critic neural network weights

## 7 结论

本文提出了一种非线性系统有限时间微分对策制导律, 构建了带时变的评估权值和激活函数评价网络, 来近似相关时变HJI方程的解. 为满足末端约束, 末端估计误差和HJI近似误差都要最小. 因此, 提出了一种新型的评估权值更新律, 此外, 通过李雅普诺夫定理, 证明了闭环系统和评估权值近似误差为最终一致有界稳定. 最后, 通过对一个导弹拦截目标非线性系统的仿真, 证明了此方法的有效性.

## 参考文献:

- [1] LI Qingchun, ZHANG Wensheng, HAN Gang. Review of terminal guidance law with terminal constraints. *Control Theory & Applications*, 2016, 33(1): 1 – 12. (李庆春, 张文生, 韩刚. 终端约束条件下末端制导律研究综述. 控制理论与应用, 2016, 33(1): 1 – 12.)
- [2] HUA Wenhua, MENG Qingling, ZHANG Jinpeng. Time-to-go weighted differential game guidance with terminal constraints. *High Technology Communication*, 2014, 24(10): 1077 – 1083. (花文华, 孟庆龄, 张金鹏. 剩余飞行时间加权的终端约束微分对策制导. 高技术通讯, 2014, 24(10): 1077 – 1083.)
- [3] DING Shihong, LI Shihua. A survey for finite-time control problems. *Control and Decision*, 2011, 26(2): 161 – 169. (丁世宏, 李世华. 有限时间控制问题综述. 控制与决策, 2011, 26(2): 161 – 169.)
- [4] LI S, DING S, LI Q. Global set stabilization of the spacecraft attitude control problem based on quaternion. *International Journal of Robust and Nonlinear Control*, 2010, 20(1): 84 – 105.
- [5] LI S, DING S, LI Q. Global set stabilisation of the spacecraft attitude using finite-time control technique. *International Journal of Control*, 2009, 82(5): 822 – 836.
- [6] SHTESSEL Y, SHKOLNIKOV I. Integrated guidance and control of advanced interceptors using second order sliding modes. *Proceedings of IEEE Conference on Decision and Control*. Maui: IEEE, 2003, 5: 4587 – 4592.
- [7] LIN Xiaofeng, DING Qiang. Review of terminal guidance law with terminal constraints. *Control and Decision*, 2015, 30(3): 495 – 499. (林小峰, 丁强. 基于评价网络近似误差的自适应动态规划优化控制. 控制与决策, 2015, 30(3): 495 – 499.)
- [8] ZHANG Huaguang, ZHANG Xin, LUO Yanhong, et al. An overview of research on adaptive dynamic programming. *Acta Automatica Sinica*, 2013, 39(4): 303 – 311. (张化光, 张欣, 罗艳红, 等. 自适应动态规划综述. 自动化学报, 2013, 39(4): 303 – 311.)
- [9] SUN Jingliang, LIU Chunsheng. An overview on the adaptive dynamic programming based missile guidance law. *Acta Automatica Sinica*, 2017, 43(7): 1101 – 1113. (孙景亮, 刘春生. 基于自适应动态规划的导弹制导律研究综述. 自动化学报, 2017, 43(7): 1101 – 1113.)
- [10] ZHANG Xin, HUI Guotao, LUO Yanhong. ADP approach to solve unknown nonlinear zero-sum game. *Journal of Northeastern University (Natural Science)*, 2012, 33(12): 1673 – 1676, 1789. (张欣, 会国涛, 罗艳红. 基于ADP方法求解未知非线性零和微分对策问题. 东北大学学报(自然科学版), 2012, 33(12): 1673 – 1676, 1789.)
- [11] QU Qixia, LUO Yanhong, ZHANG Huaguang. Robust approximate optimal tracking control of time-varying trajectory for nonlinear affine systems. *Control Theory & Applications*, 2016, 33(1): 77 – 84. (屈秋霞, 罗艳红, 张化光. 针对时变轨迹的非线性仿射系统的鲁棒近似最优跟踪控制. 控制理论与应用, 2016, 33(1): 77 – 84.)
- [12] XU H. Finite-horizon near optimal design of nonlinear two-player zero-sum game in presence of completely unknown dynamics. *Journal of Control, Automation and Electrical Systems*, 2015, 26(4): 361 – 370.
- [13] CHEN S, LI X, ZHOU X Y. Stochastic linear quadratic regulators with indefinite control weight costs. *SIAM Journal on Control and Optimization*, 1998, 36(5): 1685 – 1702.
- [14] SUN Sheng, ZHOU Di. A finite time convergent variable structure guidance law. *Journal of Astronautics*, 2008, 29(4): 1258 – 1262. (孙胜, 周荻. 有限时间收敛变结构制导律. 宇航学报, 2008, 29(4): 1258 – 1262.)

- [15] BHAT S P, BERNSTEIN D S. Finite-time stability of continuous autonomous systems. *SIAM Journal on Control and Optimization*, 2000, 38(3): 751 – 766.
- [16] SONG Y, GUO J. Neuro-adaptive fault-tolerant tracking control of lagrange systems Pursuing Targets With Unknown Trajectory. *IEEE Transactions on Industrial Electronics*, 2017, 64(5): 3913 – 3920.
- [17] MENG W, YANG Q, SUN Y. Adaptive neural control of nonlinear MIMO systems with time-varying output constraints. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, 26(5): 1074 – 1085.
- [18] ZHANG T P, ZHU Q, YANG Y Q. Adaptive neural control of non-affine pure-feedback non-linear systems With input nonlinearity and perturbed uncertainties. *International Journal of Systems Science*, 2012, 43(4): 691 – 706.
- [19] REN B, SAM GE S, TEE K P, et al. Adaptive neural control for output feedback nonlinear systems using a barrier lyapunov function. *IEEE Transactions on Neural Networks*, 2010, 21(8): 1339 – 1345.
- [20] BARDHAN R, GHOSE D. An SDRE based differential game approach for maneuvering target interception. *Proceedings of AIAA Guidance, Navigation, and Control Conference*. Florida: AIAA, 2003, 5: 4587 – 4592.

#### 作者简介:

**陈燕妮** 硕士研究生, 目前研究方向为飞行制导与控制、自适应动态规划, E-mail: chenyanjenny@163.com;

**刘春生** 教授, 目前研究方向为自适应控制、最优控制、故障诊断与容错控制及其在飞行器中的应用, E-mail: liuchsh@nuaa.edu.cn;

**孙景亮** 博士研究生, 目前研究方向为最优控制、微分对策、自适应动态规划, E-mail: sunjingliangac@163.com.