第 22 卷 第 5 期
Vol. 22 No. 5
控 制 与 决 策
Control and Decision
2007 年 5 月
May 2007

文章编号: 1001-0920(2007)05-0530-05

近地共面轨道上两飞行器在径向连续小推力下的追逃界栅

张秋华[a], 孙 毅[a], 黄明明[a], 段广仁[b]
(1. 哈尔滨工业大学 a. 航天科学与力学系, b. 控制科学与工程系, 哈尔滨 150001)

摘 要: 针对近地共面轨道上两飞行器在轨追逃对策问题,以轨道根数为状态变量,在双方均为径向连续可变小推力的假设条件下,研究定性微分对策方法中考虑推力性能的界栅存在条件. 在此条件下推导出可变推力的最优控制量,以及以偏心率和幅角为状态变量的界栅表达式. 仿真实例在给定捕获半径的条件下,给出了偏心率的界栅曲线,并表明当对策结束时,逃逸器位于追踪器的正上方或正下方

关键词: 轨道根数; 径向推力; 追逃对策; 界栅
中图分类号: V412.4 **文献标识码**: A

Pursuit-evasion barrier of two spacecrafts under minute continuous radial thrust in coplanar orbit

ZHANG Qiu-hua[a], *SUN Yi*[a], *HUANG Ming-ming*[a], *DUAN Guang-ren*[b]
(a. Department of Astronautic Science and Mechanics, b. Department of Control Science and Engineering, Harbin Institute of Techonlogy, Harbin 150001, China. Correspondent: ZHANG Qiu-hua, E-mail: zhangq_h_@hit.edu.cn)

Abstract: The pursuit-evasion game of two coplanar near-earth orbiters is investigated based on orbital element equations. Under the assumption of radial, minute, continuous and variable thrust for both spacecrafts, the condition of barrier existence is established, which takes into account of thrust index. The optimum control parameter of the variable thrust is deduced based on the condition of the existence of the barrier, together with the expressions of the barriers in terms of eccentricity and argument. In numerical examples, the barrier curves of the eccentricity are obtained under given capture radius and it is shown that at the end of the game, the evader is right above or below the pursuer.

Key words: Orbital elements; Radial thrust; Pursuit-evasion game; Barrier

1 引 言

本文研究近地共面轨道上两飞行器的追逃对策问题. 研究工作在以下假设条件下进行:追踪器和逃逸器均在某近地圆轨道附近飞行,双方彼此知道对方的位置和机动能力信息;追踪器和逃逸器具有相当的径向机动能力,且均为连续小推力;捕获范围在以追踪器为圆心、以 r 为半径的圆形区域;与轨道周期相比,对策时间充分小. 在对策中,追踪器希望捕获逃逸器,逃逸器希望能够逃脱,双方形成追逃对策.

追踪器在追踪过程中,目标存在机动和非机动两种方式. 当目标非机动或目标机动但机动规律已知时,追踪器能按预定的控制策略实施控制. 此时,控制策略与常规卫星变轨时的控制策略类似,为单方最优. 本文考虑目标机动的情况,且机动规律未知. 这就需要对策双方假设对方的最优策略[1],以考虑对策的某种结局能否实现,故所涉及的问题为定性微分对策问题. 求解定性微分对策问题实质上是求出界栅,它是双方最优策略下的最优轨迹.

定性微分对策在许多文献中都有论述[1-5],但多是陆地、海上或空中追逃问题. 有关飞行器在近地轨道上追逃对策问题的文献则很少. 文献[6]用简化的近圆共面轨道模型,研究了逃逸器为一个脉冲机动、追踪器为两个脉冲机动的追逃对策问题. 基于轨道根数描述运动方程的对策问题还未见文献报道. 在应用定性微分对策理论求解控制量时,通过改变推力方向而推力大小不变来求得推力的控制项是容易的. 本文假设目标器和追踪器的推力方向固定

收稿日期: 2006-03-24; **修回日期**: 2006-06-06.
作者简介: 张秋华(1958—),女,上海人,教授,博士,从事空间飞行器最优控制的研究; 孙毅(1961—),男,哈尔滨人,教授,博士生导师,从事飞行力学、最优控制等研究.

而大小可变. 在构造界栅时, 如何体现推力大小变化的控制量成为关键.

界栅的确定与性能函数的形式相关[7], 定量对策形式的哈密顿函数能将与过程量有关的性能函数体现出来. 本文将定量与定性对策方法综合考虑, 在构造界栅中, 使之满足界栅存在的必要条件, 将具有过程量特征的推力性能引入定性哈密顿函数, 并按定性方法求解, 使问题得到解决. 将定量与定性微分对策方法综合集成并有效加入人的经验, 似乎是未来微分对策制导律发展的出路[7]. 本文最终得到假设条件下的界栅, 并给出了仿真实例.

2　微分对策系统状态方程的建立

用轨道根数表达轨道摄动方程在一些文献中都有论述[8,9]. 为适合小偏心率轨道, 需要对轨道偏心率 $e(0 < e < 1)$ 作无奇点变量变换[9], 即 $e_\xi = e\cos\omega$, $e_\eta = e\sin\omega$, ω 为近地点幅角. 引入纬度幅角 u, 偏心率矢量分量 e_ξ 和 e_η, 作为轨道摄动方程的独立变量, 替代偏心率 e, 近地点幅角 ω 和真近点角 $\theta(u = \theta + \omega)$. 在仅考虑径向推力 f_r 的作用时, 法向推力和周向推力为 0, 摄动方程为[10]

$$\dot{e}_\xi = \frac{p}{h}\sin u f_r,\quad \dot{e}_\eta = -\frac{p}{h}\cos u f_r,$$

$$\dot{p} = 0,\quad \dot{i} = 0,\quad \dot{\Omega} = 0,\quad \dot{u} = h/r^2.$$

其中

$$r = \frac{p}{1 + e_\xi\cos u + e_\eta\sin u},\quad h = \sqrt{\mu p}.$$

p 为轨道半通径, i 为轨道倾角, Ω 为升交点经度, μ 为地球引力常数.

在径向推力作用下, p, i 和 Ω 均为常数, 故选择 e_ξ, e_η 和 u 作为状态变量, 并用下标 p 和 e 分别表示追踪器和逃逸器, 建立微分对策状态方程组如下:

$$\begin{cases} \dot{e}_{\xi p} = \dfrac{p_p}{h_p}\sin u_p f_{rp},\quad e_{\xi p}(0) = e^0_{\xi p}; \\ \dot{e}_{\eta p} = -\dfrac{p_p}{h_p}\cos u_p f_{rp},\quad e_{\eta p}(0) = e^0_{\eta p}; \\ \dot{u}_p = h_p/r_p^2,\quad u_p(0) = u^0_p; \\ \dot{e}_{\xi e} = \dfrac{p_e}{h_e}\sin u_e f_{re},\quad e_{\xi e}(0) = e^0_{\xi e}; \\ \dot{e}_{\eta e} = -\dfrac{p_e}{h_e}\cos u_e f_{re},\quad e_{\eta e}(0) = e^0_{\eta e}; \\ \dot{u}_e = h_e/r_e^2,\quad u_e(0) = u^0_e. \end{cases} \tag{1}$$

3　最优控制量求解及界栅构造

在定性微分对策中, 界栅的存在必须满足条件

$$\min_{f_{rp}}\max_{f_{re}} H = \lambda^{\mathrm{T}}\dot{X} = 0.$$

其中

$$\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)^{\mathrm{T}}$$

为协态向量;

$$X^{\mathrm{T}} = (e_{\xi p}, e_{\eta p}, u_p, e_{\xi e}, e_{\eta e}, u_e)$$

为状态向量.

为在最优策略中体现推力大小变化的控制量, 将具有过程量特征的推力性能引入定性哈密顿函数, 有

$$H = \frac{1}{2}(f^2_{rp} - f^2_{re}) + \lambda^{\mathrm{T}}\dot{X} = H_p + H_e + \lambda_3\frac{h_p}{r_p^2} + \lambda_6\frac{h_e}{r_e^2}. \tag{2}$$

其中

$$H_p = \frac{1}{2}f^2_{rp} + \lambda_1\frac{p_p}{h_p}\sin u_p f_{rp} - \lambda_2\frac{p_p}{h_p}\cos u_p f_{rp},$$

$$H_e = -\frac{1}{2}f^2_{re} + \lambda_4\frac{p_e}{h_e}\sin u_e f_{re} - \lambda_5\frac{p_e}{h_e}\cos u_e f_{re}.$$

并满足双方最优策略的必要条件

$$\min_{f_{rp}}\max_{f_{re}} H = \min_{f_{rp}} H_p + \max_{f_{re}} H_e = 0. \tag{3}$$

由方程 $\partial H/\partial f_{ri} = 0$, 有最优控制推力表达式

$$\begin{cases} f^*_{rp} = \dfrac{p_p}{h_p}(\lambda_2\cos u_p - \lambda_1\sin u_p), \\ f^*_{re} = \dfrac{p_e}{h_e}(\lambda_4\sin u_e - \lambda_5\cos u_e). \end{cases} \tag{4}$$

其中 λ 为待求协态向量. 式(4) 显然满足式(3) 条件.

3.1　目标集和参数化方程的建立

目标集是根据捕获或逃逸的评判标准建立的. 这里假设评判标准是以追踪器的捕获半径 r 定义的圆形区域. 于是目标集边界满足如下集合:

$$\mathbf{R}_B = \{(r_p, r_e) \mid |r_p - r_e| = r\}.$$

其中 r_p 和 r_e 分别表示追踪器和逃逸器的地心矢径, 如图 1 所示.

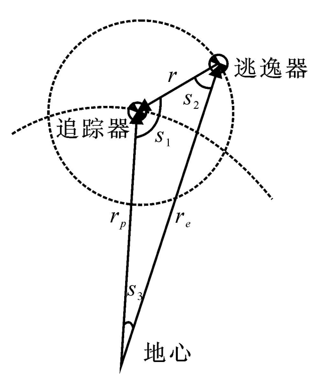

图 1　逃逸器、追踪器与地心间的矢径关系

图 1 中 s_1, s_2 和 s_3 表示对策结束时刻它们之间的夹角. 用 r, r_p 和 r_e 表示这些矢量的模, 由正弦公式有

$$r_p = \frac{r\sin s_2}{\sin s_3},\quad r_e = \frac{r\sin s_1}{\sin s_3}. \tag{5}$$

下面将目标边界集上的 $X_{\tilde{t}}$ 用参数 S 表出, $\tilde{t}$ 为终端时刻,有

$$X_{\tilde{t}}^{\mathrm{T}} = (e\xi_p, e\eta_p, u_p, e\xi_e, e\eta_e, u_e)\,|_{t=\tilde{t}},$$
$$S = (s_1, s_2, s_3, s_4, s_5)^{\mathrm{T}}.$$

其中 s_4 和 s_5 分别为终端时刻追踪器和逃逸器的幅角. 由前述偏心率变换,有

$$e\xi_i = e_i\cos\omega_i,\quad e\eta_i = e_i\sin\omega_i. \tag{6}$$

其中

$$e_i = \frac{p_i/r_i - 1}{\cos\omega_i\cos u_i + \sin\omega_i\sin u_i},\quad i = p, e.$$

将此 e_i 表达式及式(5) 代入式(6) ,可得终端时刻的参数化方程组

$$\begin{cases} e\xi_p(\tilde{t}) = \phi_1(S) = C\xi_p A_p, \\ e\eta_p(\tilde{t}) = \phi_2(S) = C\eta_p A_p, \\ u_p(\tilde{t}) = \phi_3(S) = s_4, \\ e\xi_e(\tilde{t}) = \phi_4(S) = C\xi_e A_e, \\ e\eta_e(\tilde{t}) = \phi_5(S) = C\eta_e A_e, \\ u_e(\tilde{t}) = \phi_6(S) = s_5. \end{cases} \tag{7}$$

其中

$$A_p = \frac{p_p\sin s_3}{r\sin s_2} - 1,\quad A_e = \frac{p_e\sin s_3}{r\sin s_1} - 1,$$

$$C\xi_i = \frac{\cos\omega_i}{\cos\omega_i\cos u_i + \sin\omega_i\sin u_i}\Big|_{t=\tilde{t}},$$

$$C\eta_i = \frac{\sin\omega_i}{\cos\omega_i\cos u_i + \sin\omega_i\sin u_i}\Big|_{t=\tilde{t}}.$$

3.2 协态方程和状态方程的终端条件

用下标 B 表示目标集边界, 由目标集边界上 $\min\limits_P\max\limits_E H_B = 0$,可得双方在目标边界上的各自最优策略

$$\min_P \boldsymbol{\Gamma}_p = \left(\frac{1}{2}f_{rp}^2 + \lambda_1\frac{p_p}{h_p}\sin u_p f_{rp} - \lambda_2\frac{p_p}{h_p}\cos u_p f_{rp}\right)\Big|_B, \tag{8a}$$

$$\max_E \boldsymbol{\Gamma}_e = \left(-\frac{1}{2}f_{re}^2 + \lambda_4\frac{p_e}{h_e}\sin u_e f_{re} - \lambda_5\frac{p_e}{h_e}\cos u_e f_{re}\right)\Big|_B. \tag{8b}$$

其中

$$\lambda|_B = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)^{\mathrm{T}}_{t=\tilde{t}}$$

为目标集边界上的协态向量,可由如下方程求解:

$$\begin{cases} \sum\limits_{i=1}^{6}\lambda_i\dfrac{\partial\phi_i(s_1, s_2, s_3, s_4, s_5)}{\partial s_j} = 0, \\ \lambda^{\mathrm{T}}\lambda|_B = 1,\ j = 1, 2, \cdots, 5. \end{cases} \tag{9}$$

将式(9) 求解的结果代入式(4) ,可得目标边界上的最优控制量;将此最优控制量及式(9) 求解的结果代入式(2) ,可得目标集边界上的哈密顿函数

$$H_B = \boldsymbol{\Gamma}_P + \boldsymbol{\Gamma}_E = 0. \tag{10}$$

其中

$$\boldsymbol{\Gamma}_P = -\frac{p_p^2}{2h_p^2}\frac{-1+\omega_e^2}{\omega_e^2+\omega_p^2}\cos^2(u_p - \omega_p),$$

$$\boldsymbol{\Gamma}_E = \frac{p_e^2}{2h_e^2}\frac{-1+\omega_p^2}{\omega_p^2+\omega_e^2}\cos^2(u_e - \omega_e).$$

式(10) 意味着对策结束时刻双方推力相等.

另外,由式(10) 对 u_i 和 u_p 求极值,并满足双方在目标边界上的最优策略,可得到 $0 \sim 2\pi$ 区间 u_i 与 $\omega_i\,(i = p, e)$ 的关系式

$$\begin{cases} u_p^* = \omega_p^* + \pi = s_4, \\ u_e^* = \omega_e^* + \pi = s_5. \end{cases} \tag{11}$$

将式(11) 代入式(9) 的解,可得协态方程的终端条件

$$\begin{cases} \lambda_1(\tilde{t}) = -\sin s_4 D\alpha,\ \lambda_2(\tilde{t}) = \cos s_4 D\alpha, \\ \lambda_3(\tilde{t}) = 0,\ \lambda_4(\tilde{t}) = \sin s_5 D\beta, \\ \lambda_5(\tilde{t}) = -\cos s_5 D\beta,\ \lambda_6(\tilde{t}) = 0. \end{cases} \tag{12}$$

其中

$$D\alpha = \sqrt{\frac{-1 + (s_5 - \pi)^2}{(s_5 - \pi)^2 - (s_4 - \pi)^2}},$$

$$D\beta = \sqrt{\frac{-1 + (s_4 - \pi)^2}{(s_4 - \pi)^2 - (s_5 - \pi)^2}}.$$

将式(11) 代入式(7) ,可得状态方程的终端条件

$$\begin{cases} e\xi_p(\tilde{t}) = \cos s_4 A_p,\ e\eta_p(\tilde{t}) = \sin s_4 A_p, \\ u_p(\tilde{t}) = s_4,\ e\xi_e(\tilde{t}) = \cos s_5 A_e, \\ e\eta_e(\tilde{t}) = \sin s_5 A_e,\ u_e(\tilde{t}) = s_5. \end{cases} \tag{13}$$

3.3 倒向协态和状态方程的解

倒向协态方程组可由 $\lambda^{\mathrm{T}} = \partial H/\partial X$ 得到,其中 λ 是对倒向时间 τ 求导. 将式(4) 代入倒向协态方程组,并将终端条件(12) 作为倒向协态方程组的初始条件对方程求解,整理后可得协态方程的解

$$\begin{cases} \lambda_1(\tau) = -\sin s_4 D\alpha,\ \lambda_2(\tau) = \cos s_4 D\alpha, \\ \lambda_3(\tau) = \dfrac{r_p^2 p_p^2}{4h_p^3}D_\alpha^2[\cos 2(u_p - s_4) - 1], \\ \lambda_4(\tau) = \sin s_5 D\beta,\ \lambda_5(\tau) = -\cos s_5 D\beta, \\ \lambda_6(\tau) = \dfrac{r_e^2 p_e^2}{4h_e^3}D_\beta^2[1 - \cos 2(u_e - s_5)]. \end{cases} \tag{14}$$

将式(14) 代入式(4) ,可得对策双方的最优控制量

$$\begin{cases} f_{rp}^* = \dfrac{p_p}{h_p}D\alpha\cos(u_p - s_4), \\ f_{re}^* = \dfrac{p_e}{h_e}D\beta\cos(u_e - s_5). \end{cases} \tag{15}$$

将式(15) 代入状态方程(1) ,并对时间作倒向变换,可得满足双方最优控制的倒向状态方程. 将式(13) 作为倒向状态方程的初始条件对方程求解,在

求解的具体时刻将 r_p 和 r_e 视为常值,则双方最优策略下倒向状态方程的近似解为

$$\begin{cases} e_{\xi p}^* = \cos s_4 A_p + \dfrac{r_p^2 p_p^2}{2h_p^3}[\sin(u_p - s_4)\sin u_p + (u_p - s_4)\sin s_4]D\alpha, \\ e_{\eta p}^* = \sin s_4 A_p - \dfrac{r_p^2 p_p^2}{2h_p^3}[\sin(u_p - s_4)\cos u_p + (u_p - s_4)\cos s_4]D\alpha, \\ u_p^* = \dfrac{h_p}{r_p^2}\tau + s_4, \\ e_{\xi e}^* = \cos s_5 A_e + \dfrac{r_e^2 p_e^2}{2h_e^3}[\sin(u_e - s_5)\sin u_e + (u_e - s_5)\sin s_5]D\beta, \\ e_{\eta e}^* = \sin s_5 A_e - \dfrac{r_e^2 p_e^2}{2h_e^3}[\sin(u_e - s_5)\cos u_e + (u_e - s_5)\cos s_5]D\beta, \\ u_e^* = \dfrac{h_e}{r_e^2}\tau + s_5. \end{cases} \tag{16}$$

其中 s_4 与 s_5 间的关系由终端上双方推力相等的条件确定.

根据终端几何关系,有

$$\sin s_3 = r\sin s_1 / r_e,$$
$$s_3 = | s_5 - s_4 |,$$
$$s_2 = 180° - (s_1 + s_3).$$

其中:s_1 可表为 $s_1 = s_0 + \pi/2$, s_0 为终端时刻追踪器的视线与当地水平线的夹角. 于是式(16) 中的参数 S 可由 s_0 一个参变量全部表出. 由此方程(16) 中仅剩两个参变量:捕获半径 r 和视线角 s_0,它们都是终端时刻的参数. r 在本文中为给定量, s_0 是目标边界上待确定的可用部分边界,可通过取极值来确定.

至此,本文给出了对策双方在径向连续小推力下 e_ξ, e_η 和 $u_i(i = p, e)$ 轨迹的表达式. 这些表达式是以目标边界集上可用部分边界(终端条件) 为初始点倒向积分得到的,是双方中立结局,即所对应的轨迹上既不会发生捕获也不会发生逃逸. 因此这些轨迹是划分捕获区与逃逸区的分界,亦即寻找的界栅.

4　实例仿真与分析

本文仿真实例的初始条件如下:轨道半长轴分别为: a_{0p} = 6 928.14 km, a_{0e} = 6 938.16 km;地心引力常数为 μ = 398 600.5 km^3/s^2; 捕获半径为 2 km.

追踪器控制力 f_{rp} 随倒向时间的变化如图 2 所示. 追踪器偏心率和两飞行器偏心率差随时间变化的规律以及终端视线角 s_0 对其影响,分别如图 3 和图 4 所示.

从图 3 可见,对策中偏心率的控制由大到小进行,这样有利于轨道控制,即机动后的轨道仍处于可控的小偏心率轨道范围内. 从图 3 和图 4 可以看出,追踪器偏心率及两飞行器偏心率差在 $s_0 = \pm\pi/2$ 时,存在随倒向时间变化的最大上限和最小下限.

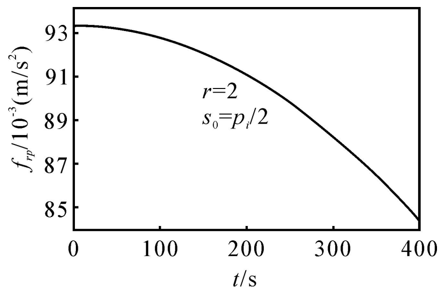

图 2　控制量 f_{rp} 随倒向时间的变化

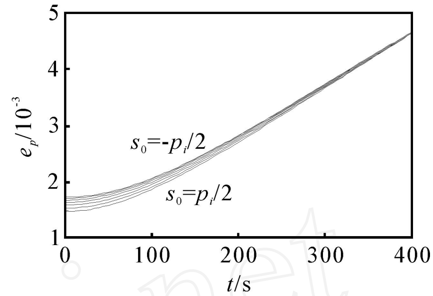

图 3　s_0 对追踪星偏心率随倒向时间变化的影响

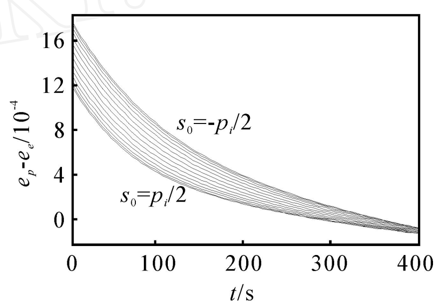

图 4　s_0 对两星偏心率差随倒向时间变化的影响

前面的推导中已说明, s_0 为终端时刻目标边界上待确定的可用部分边界. 两个极限状态点 $s_0 = \pm\pi/2$, 既是目标边界集上可用部分的两个边界点,也是协态方程和状态方程倒向积分时的起始点,因此相应的轨迹即为界栅轨迹. 这说明对于中立结局,当对策结束时逃逸器位于追踪器的正上方或正下方.

从图 3 和图 4 还可看出,极限状态的两个界栅轨迹与目标集的可用部分边界形成封闭的曲线,该闭合曲线围成的区域即为捕获区. 两飞行器的其他界栅轨迹,可用类似方法得到.

5　结　　论

本文推导得到了假设下的双方最优策略和最优策略下的状态变量表达式. 这些表达式是在满足界栅存在的必要条件下,以双方最优策略的终端条件为初始点倒向积分得到的,即与表达式对应的轨迹

上的点既不会发生捕获也不会发生逃逸,是双方中立的结局.因此这些轨迹是划分捕获区与逃逸区的分界,亦即寻找的界栅.界栅存在的结果说明,将定量与定性微分对策方法综合集成是可行的.

仿真实例给出了追踪器偏心率和两飞行器偏心率差随倒向时间变化的界栅曲线,它们是由目标集的可用部分边界与偏心率随倒向时间变化的最大上限和最小下限轨迹组成的封闭曲线.仿真结果表明,在对策结束时刻,逃逸器处于追踪器的正上方或正下方;在对策过程中,偏心率的控制由大到小进行,这对轨道控制较为有利.同时表明,推力大小对轨道偏心率的影响较大,这说明该方法仅适用于小推力的情况.

参考文献(References)

[1] Glizer V Y. Homicidal chauffeur game with target set in the shape of a circular angular sector: Conditions for existence of a closed barrier[J]. J of Optimization Theory and Applications, 1999, 101(3): 581-598.

[2] Isaacs R. Differential games[M]. New York: John Wiley, 1965.

[3] Turetsky V, Shinar J. Missile guidance laws based on pursuit-evasion game formulations[J]. Automatica, 2003, 39(4): 607-618.

[4] Guelman M, Shinar J, Green A. Qualitative study of a planar pursuit evasion game in the atmosphere[J]. J of Guidance, Control and Dynamics, 1990, 13(6): 1136-1142.

[5] 李登峰. 微分对策及其应用[M]. 北京: 国防工业出版社, 2004.
(Li D F. Differential games and applications[M]. Beijing: Defence Press, 2004.)

[6] Kelley H J, Cliff E M, Lutze F H. Pursuit/evasion in orbit[J]. J of the Astronautical Sciences, 1981, 29(3): 277-288.

[7] 周卿吉,许诚,周文松,等. 微分对策制导律研究的现状及展望[J]. 系统工程与电子技术, 1997, 19(11): 40-45.
(Zhou Q J, Xu C, Zhou W S, et al. Status and prospects of the development of DGGL[J]. Systems Engineering and Electronics, 1997, 19(11): 40-45.)

[8] 刘敦, 赵钧. 空间飞行器动力学[M]. 哈尔滨: 哈尔滨工业大学出版社, 2003.
(Liu D, Zhao J. Dynamics of spacecraft[M]. Harbin: Harbin Institute of Technology Press, 2003.)

[9] 刘林. 航天器轨道理论[M]. 北京: 国防工业出版社, 2000.
(Liu L. Orbit theory of spacecraft[M]. Beijing: Defence Press, 2000.)

[10] 朱仁章,林彦,李颐黎. 论空间交会中的径向连续推力机动与 N 次推力机动[J]. 中国空间科学技术, 2004, 24(3): 21-28.
(Zhu R Z, Lin Y, Li Y L. Analysis of constant continuous radial thrust and N-thrusts in space rendezvous[J]. Chinese Space Science and Technology, 2004, 24(3): 21-28.)

(上接第 529 页)

[2] Sutton R S, Barto A G. Reinforcement learning: An introduction[M]. Cambridge: MIT Press, 1998.

[3] 蒋国飞, 吴沧浦. 基于 Q 学习算法和 BP 神经网络的倒立摆控制[J]. 自动化学报, 1998, 24(5): 662-666.
(Jiang G F, Wu C P. Learning to control an inverted pendulum using Q-learning and neural networks[J]. Acta Automatica Sinica, 1998, 24(5): 662-666.)

[4] Claude F T. Neural reinforcement learning for behaviour synthesis[J]. Robotics and Autonomous Systems, 1997, 22(3/4): 251-281.

[5] Jouffe L. Fuzzy inference system learning by reinforcement methods[J]. IEEE Trans on Systems, Man and Cybernetics, 1998, 28(3): 338-355.

[6] Baird L C. Residual algorithms: Reinforcement learning with function approximation[C]. Proc of the 12nd Int Conf on Machine Learning. San Francisco, 1995: 9-12.

[7] 张汝波. 强化学习理论及应用[M]. 哈尔滨: 哈尔滨工程大学出版社, 2000.
(Zhang R B. Reinforement learning theory and applications[M]. Harbin: Harbin Engineering University Press, 2000.)

[8] Watkins C J, Dayan P. Q-learning[J]. Machine Learning, 1992, 8(3): 279-292.

[9] Peng J, Williams R J. Incremental multi-step Q-learning[C]. Proc of the 11th Int Conf on Machine Learning. New Brunswick: Morgan Kaufmann, 1995: 226-232.

[10] Lin C H, Wang L L. Intelligent collision avoidance by fuzzy logic control[J]. Robotics and Autonomous Systems, 1997, 20(1): 61-83.

[11] Xu W L, Tso K K. Sensor-based fuzzy reactive navigation of a mobile robot through local target switching[J]. IEEE Trans on Systems, Man and Cybernetics — Part C: Applications and Reviews, 1999, 29(3): 451-459.