离散时间二人随机微分对策问题信息模式的数学描述

范红旗,王 胜,付 强
(国防科技大学 ATR 实验室,湖南长沙 410073)

摘 要: 在离散时间二人随机微分对策问题研究中,信息模式概念尚缺乏统一而准确的描述.针对这一问题,首先将 Witsenhausen 关于信息模式的相关概念应用到该问题,从数学上严格定义了信息模式及其相关概念,然后对几种典型信息模式的性质及相应对策问题最优解的结构形式作出了严格的证明.相关概念与性质为离散时间二人随机微分对策问题的研究提供了重要的理论工具.

关键词: 随机微分对策;信息模式;状态估计;最优控制;控制律

中图分类号: TP273⁺.1 **文献标识码:** A **文章编号:** 0372-2112(2010)06-1355-07

Mathematical Description for Information Pattern of Discrete-Time Two-Person Stochastic Differential Games

FAN Hong-qi, WANG Sheng, FU Qiang
(*ATR Lab of National University of Defense Technology*, *Changsha*, *Hunan* 410073, *China*)

Abstract: In the researches of discrete-time two-person stochastic differential games, the uniform and exact descriptions on information pattern are absent. To solve this problem, the mathematical definitions of information pattern and other related concepts are addressed firstly, by introducing Witsenhausen′s concepts about information pattern to the discrete-time two-person stochastic differential games. Then the properties of several typical information patterns and the structures of optimal control law of differential game under these information patterns are strictly proved. The concepts and theorems of the paper provide an important theoretical tool for the researches of discrete-time two-person stochastic differential games.

Key words: stochastic differential games; information pattern; state estimation; optimal control; control law

1 引言

微分对策是采用微分方程来描述系统动态过程的一类对策理论,其研究最早可追溯到 Isaacs[1].20 世纪 50 年代末,出于飞行器拦截等军事上的需要,兰德公司以 Isaacs 为领导开始了这方面的研究.他们将现代控制理论中的概念和原理引入到对策论中,取得了重大进展.此后,微分对策理论迅速吸引了战术决策、寻的制导、随机系统控制、网络攻防等领域研究人员的广泛关注,并有大量的著作问世[2~4].

在微分对策研究中,二人微分对策是一类最简单的对策问题,理论成果相对较为完备.然而,目前出版的著作[2~4]大都限于完美信息下的微分对策[2].受传感器限制,这种完美信息模式假设在实际应用中会存在两方面的问题:一方面,状态观测中通常包含了一定随机噪声;另一方面,某些状态分量常常无法直接观测.比如在机动目标拦截中,雷达和图像等传感器都不能直接测量目标横向加速度.此外,由于采用计算机控制,状态观测及控制只在离散采样时刻获得[5,6],因此在研究此类系统的最优控制时,往往需要对连续时间状态方程进行离散化,并考虑对策双方信息的缺失与随机性.所以,研究离散时间随机微分对策问题具有更加实际的意义.

由于对策双方的信息结构(本文称作信息模式)决定着对策问题最优解的形式及其存在性,因此关于随机微分对策问题的讨论总是基于一定的信息模式.以寻的制导应用为例,早期的代表性成果包括:Behn 和 Ho[7]假定对策一方的观测中包含白噪声,而另一方则拥有完美信息并可估计对方的控制量,他们得到了该信息模式下对策双方的控制律;Rhodes 和 Luenberger 等人[8]研究了与文献[7]类似信息模式下的对策问题,所不同的是,拥

收稿日期:2008-06-05;修回日期:2009-02-27
基金项目:973 国家安全重大基础研究(No.51314);国防科技预研跨行业综合技术(No.51301050102)

有完美信息的一方不能估计出对方的控制量. Willman[9]研究了线性二次高斯微分对策问题,其中对策双方的状态观测都包含高斯噪声,他证明了这类对策问题满足确定性一致原理,对策问题的解可分解为确定性等价项与噪声项之和,由此得到了对策问题的极小极大解,但他并未对解的收敛性和最优性给出严格的证明.而近10年来代表性成果主要集中在Shinar等人工作中[10~12].他们考虑目标机动加速度的估计延迟,在文献[10]中将目标加速度估计等效为真实加速度的延迟,结果表明信息延迟条件下的最优制导策略为一有偏混合策略;在文献[11]中他们基于极小极大确定性等价原理与状态可达集的概念,得到了加速度估计延迟下的微分对策制导律DGL/C; Oshman等人[12]基于DGL/C,利用成像导引头观测到的目标姿态信息来缩减可达集的范围,得到了图像辅助的微分对策制导律DGL/S.最近,Swarup与Speyer[13]对不同信息模式下线性二次高斯随机微分对策的研究情况做了较为全面的总结,表明了信息模式对于微分对策问题研究的重要性.

然而目前讨论随机微分对策的各类著作中,对信息模式的概念尚缺乏统一而准确的定义.从数学上准确描述信息模式,对求解随机微分对策问题的最优解,判断最优解的结构形式,理解状态估计与最优控制间的关系等,都具有重要的意义.本文将以离散时间二人随机微分对策问题为例,从数学上严格定义信息模式的相关概念,并对几种典型信息模式的性质给出严格证明.

2 离散时间二人随机微分对策

采用追逃对策惯例,记对策双方分别为P(Pursuer,追踪者)和E(Evader,逃逸者),则 N 步离散时间二人随机微分对策可表示为:

$$\begin{aligned} &\boldsymbol{x}_{k+1}=f_k(\boldsymbol{x}_k,\boldsymbol{w}_k,\boldsymbol{u}_k^{\mathrm{P}},\boldsymbol{u}_k^{\mathrm{E}})\\ &z_k^i=h_k^i(\boldsymbol{x}_k,\boldsymbol{v}_k^i)\\ &\boldsymbol{u}_k^{\mathrm{P}}\in\boldsymbol{U}^{\mathrm{P}}\subset\mathbf{R}^{l_{\mathrm{P}}},\boldsymbol{u}_k^{\mathrm{E}}\in\boldsymbol{U}^{\mathrm{E}}\subset\mathbf{R}^{l_{\mathrm{E}}}\\ &\boldsymbol{w}_k\in\mathbf{R}^{n_w},\boldsymbol{v}_k^i\in\mathbf{R}^{n_v^i}\\ &J=\phi[\boldsymbol{x}(N),N]+\sum_{k=0}^{N-1}G(k,\boldsymbol{x}_k,\boldsymbol{u}_k^{\mathrm{P}},\boldsymbol{u}_k^{\mathrm{E}})\\ &k=0,\cdots,N-1,i=\mathrm{P},\mathrm{E}\end{aligned}\tag{1}$$

其中: $\boldsymbol{x}_k\in\mathbf{R}^n$,表示 k 时刻状态; $\boldsymbol{u}_k^i(i=\mathrm{P},\mathrm{E})$ 表示 k 时刻P或E的控制输入; $\boldsymbol{U}^i$ 为 $\mathbf{R}^{l_i}(i=\mathrm{P},\mathrm{E})$ 中的有界闭集,称为P/E的控制空间; $z_k^{\mathrm{P}}\in\mathbf{R}^{m_{\mathrm{P}}}$, $z_k^{\mathrm{E}}\in\mathbf{R}^{m_{\mathrm{E}}}$,分别表示 k 时刻 P 和 E 的观测矢量; $\boldsymbol{w}_k$、$\boldsymbol{v}_k^i$ 分别表示 k 时刻的过程噪声及P/E的观测噪声;函数 f_k 为 $\mathbf{R}^n\times\mathbf{R}^{n_w}\times\boldsymbol{U}^{\mathrm{P}}\times\boldsymbol{U}^{\mathrm{E}}\rightarrow\boldsymbol{R}^n$ 的连续(或分段连续)函数,而 h_k^i 为 $\mathbf{R}^n\times\mathbf{R}^{n_v^i}\rightarrow\mathbf{R}^{m_i}$ 的连续(或分段连续)函数,它们分别表示 k 时刻系统状态转移函数及P/E的观测函数; J 通常称作指标泛函或支付泛函,其中函数 $G(k,\boldsymbol{x}_k,\boldsymbol{u}_k^{\mathrm{P}},\boldsymbol{u}_k^{\mathrm{E}})$ 为定义在 $[0,N-1]\times\mathbf{R}^n\times\boldsymbol{U}^{\mathrm{P}}\times\boldsymbol{U}^{\mathrm{E}}$ 上的实值连续函数,而 $\phi[\boldsymbol{x}(N),N]$ 为终端指标,是定义在 $\mathbf{R}^n\times\mathbf{N}$ 上的实值函数,且在有界子集中有界.

随机微分对策的研究中,通常认为初始状态 $\boldsymbol{x}_0$、$\boldsymbol{w}_k$、$\boldsymbol{v}_k^i$ 间相互独立,它们一起构成系统的"原始随机变量".记:

$$\boldsymbol{\omega}=(\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{v}_0^{\mathrm{P}},\boldsymbol{v}_0^{\mathrm{E}},\boldsymbol{w}_1,\boldsymbol{v}_1^{\mathrm{P}},\boldsymbol{v}_1^{\mathrm{E}},\cdots,\boldsymbol{w}_{N-1},\boldsymbol{v}_{N-1}^{\mathrm{P}},\boldsymbol{v}_{N-1}^{\mathrm{E}})$$

$$\boldsymbol{\Omega}=\{\boldsymbol{\omega}\mid\boldsymbol{x}_0\in\mathbf{R}^n,\boldsymbol{w}_k\in\mathbf{R}^{n_w},\boldsymbol{v}_k^i\in\mathbf{R}^{n_v^i},\ k=0,1,\cdots,N-1,i=\mathrm{P},\mathrm{E}\}$$

上述表示下, $\boldsymbol{\omega}$ 为"原始随机变量"的一个样本; $\boldsymbol{\Omega}$ 则表示了系统"原始随机变量"的样本空间.令 $\boldsymbol{O}$ 表示拓扑空间 $\boldsymbol{\Omega}$ 上的所有开集系, $\boldsymbol{B}\triangleq\sigma(\boldsymbol{O})$ 表示 $\boldsymbol{\Omega}$ 上的Borel集合系, $\boldsymbol{B}$ 中的集合称作 $\boldsymbol{\Omega}$ 上的Borel集;那么, $(\boldsymbol{\Omega},\boldsymbol{B},P)$ 就成为系统"原始随机变量"构成的概率空间,表征了系统的不确定性,其中 P 为概率测度.

3 信息模式的数学描述

Witsenhausen在其关于分离原理的著名论文[6]中描述了具有 K 个控制单元和 M 个观测单元的离散时间随机控制系统的信息模式,并讨论了几种信息模式下状态估计与最优控制的可分离性,相关概念与结论对多入多出(MIMO)控制系统设计具有重要指导意义,但关于状态估计与最优控制可分离性的论断,Witsenhausen并未给出严格证明.本节将Witsenhausen关于信息模式的相关概念应用到式(1)的二人随机微分对策问题.

定义1(数据基及其生成的数据集) 令 Z_k、U_k 分别表示如下有序对所构成的集合:

$$Z_k=\{(\tau,i)\mid\tau=0,\cdots,k;i=\mathrm{P},\mathrm{E}\},k=0,\cdots N-1$$

$$U_k=\{(\tau,i)\mid\tau=0,\cdots,k-1;i=\mathrm{P},\mathrm{E}\},k=1,\cdots,N$$

显然有: $U_0=\varnothing$.如果 $A\subseteq Z_k$, $B\subseteq U_k$,则称有序对 $(A,B)_k$ 为 k 时刻的一个数据基,称 $(z_A,\boldsymbol{u}_B)=z_A\cup\boldsymbol{u}_B$ 为由 $(A,B)_k$ 生成的数据集.其中:

$$z_A=\{z_k^i\mid(k,i)\in A,\ z_k^i\in\mathbf{R}^{m_i}\};$$

$$\boldsymbol{u}_B=\{\boldsymbol{u}_k^i\mid(k,i)\in B,\boldsymbol{u}_k^i\in\boldsymbol{U}^i\}.$$

定义2(信息模式) 式(1)表示的离散时间对策问题中, $\forall(k,i)\in U_N$,称它对应的数据基为 k 时刻单元 i (P或E)的信息模式(information pattern),记作 $(Z_{k,i},U_{k,i})$.由此信息模式生成的数据集 $\{z_{Z_{k,i}},\boldsymbol{u}_{U_{k,i}}\}$ 表示了 k 时刻单元 i(P或E)产生控制量 $\boldsymbol{u}_k^i$ 时的所有可用信息.

定义3(控制律) 令 $\boldsymbol{\gamma}_k^i$ 表示 $\mathrm{span}(z_{Z_k}\cup\boldsymbol{u}_{U_k})\rightarrow\boldsymbol{U}^i$ 的所有Borel可测映射构成的集合. $\forall(k,i)\in U_N$, $\exists\gamma_k^i$

$\in \boldsymbol{\gamma}_k^i$,使得 $\boldsymbol{u}_k^i=\gamma_k^i(\boldsymbol{z}_{Z_{k,i}},\boldsymbol{u}_{U_{k,i}})$,则称 γ_k^i 为 k 时刻单元 i 的控制律(控制函数).对于追逃问题,γ_k^{P}、γ_k^{E} 分别为制导律和逃避策略.

定义 4(单元及系统的设计)　令 $(\gamma_k^i)_0^{N-1}=(\gamma_0^i,\gamma_1^i,\cdots,\gamma_{N-1}^i)(i=\mathrm{P},\mathrm{E})$ 表示式(1) N 步离散时间随机对策中单元 i 的控制律序列,称之为单元 i 的一个设计.将 $((\gamma_k^{\mathrm{P}})_0^{N-1},(\gamma_k^{\mathrm{E}})_0^{N-1})$ 称为式(1)系统的一个设计,记作 γ,将所有的系统设计构成的集合记作 Γ.令 $\Gamma^i=\{\gamma^i \mid \gamma^i=(\gamma_k^i)_0^{N-1},\gamma_k^i\in\boldsymbol{\gamma}_k^i\}(i=\mathrm{P},\mathrm{E})$ 表示单元 i 的所有可行的设计构成的集合,P 和 E 分别从各自的可行设计集 Γ^i 中选择控制量以使性能指标 J 满足微分对策问题的鞍点条件.

定义 5($\boldsymbol{\sigma}$ 域的基)　令 $\boldsymbol{\gamma}_L=\{\gamma_k^i \mid (k,i)\in L,L\subseteq U_N\}$ 表示设计 γ 中由集合 L 的元素检索的控制函数所构成的集合.如果 $Z\subseteq Z_k,U\subseteq U_k,L\subseteq U_N,\forall\gamma,\gamma'\in\Gamma$,下面关系成立:

$$\boldsymbol{\gamma}_L=\boldsymbol{\gamma}'_L\Rightarrow F(Z,U;\gamma)=F(Z,U;\gamma')$$

则称三元组(Z,U,L)为 k 时刻 σ 域 $F(Z,U;\gamma)$的一组基.其中,$F(Z,U;\gamma)$表示设计 γ 下由定义在概率空间$(\boldsymbol{\Omega},F,P)$上的随机变量集$(z_z,\boldsymbol{u}_U)$生成的 σ 域.

定义 6(状态变量的条件基)　如果存在函数 F,$\forall\gamma\in\Gamma$,给定 σ 域 $F(Z,U;\gamma)$后,状态变量 $\boldsymbol{y}$ 的条件概率分布几乎处处等于 $F(z_Z,\boldsymbol{u}_U,\boldsymbol{\gamma}_L)$,则称三元组$(Z,U,L)$为 $\boldsymbol{y}$ 的条件基.

定义 7(等价设计)　将控制变量 $\boldsymbol{u}$ 表示为原始随机变量 $\boldsymbol{\omega}$ 与设计 γ 的函数,即:$\boldsymbol{u}=S(\boldsymbol{\omega},\gamma)$.$\forall\gamma\in\Gamma$,如果 $\exists\gamma'\in\Gamma$,使得下式成立:

$$S(\boldsymbol{\omega},\gamma)\ \underline{\underline{\mathrm{a.e.}}}\ S(\boldsymbol{\omega},\gamma')$$

则称这两个设计等价,记作 $\gamma\Leftrightarrow\gamma'$.

定义 8(等价的信息模式)　对于信息模式$(Z_{k,i},U_{k,i})$下任意的设计 γ,如果信息模式$(Z'_{k,i},U'_{k,i})$下存在另一个设计 γ',使得 $\gamma\Leftrightarrow\gamma'$,反之亦成立,则称这两个信息模式等价,记作$(Z_{k,i},U_{k,i})\Leftrightarrow(Z'_{k,i},U'_{k,i})$.

定义 9(等价的条件基)　$\forall\gamma\in\Gamma$,如果由状态变量 $\boldsymbol{y}$ 的条件基(Z,U,L)和(Z',U',L')决定的概率分布满足 $F(z_Y,\boldsymbol{u}_U,\boldsymbol{\gamma}_L)\ \underline{\underline{\mathrm{a.e.}}}\ F(z_{Y'},\boldsymbol{u}_{U'},\boldsymbol{\gamma}_{L'})$,则称这两个条件基等价,记作$(Z,U,L)\Leftrightarrow(Z',U',L')$.

4　几类典型信息模式及其性质

定义 10(标准模式)　如果 P 和 E 的信息模式$(Z_{k,i},U_{k,i})$满足:

$$Z_{k,i}\subseteq Z_{k+1,i},U_{k,i}\subseteq U_{k+1,i};\ k=0,\cdots N-2,i=\mathrm{P},\mathrm{E},$$
$$Z_{k,\mathrm{P}}=Z_{k,\mathrm{E}},U_{k,\mathrm{P}}=U_{k,\mathrm{E}},$$

则称该信息模式为标准模式.

定义 11(j 步延迟共享模式)　如果 P 和 E 的信息模式满足:

$$Z_{k,i}=Z_{k-j}\cup\{(p,i)\mid k-j<p\leqslant k\},$$
$$U_{k,i}=U_{k-j+1}\cup\{(p,i)\mid k-j<p\leqslant k-1\},$$

则称该信息模式为 j 步延迟共享模式.

特别地,对于仅有一个控制单元的系统而言,j 步延迟共享模式意味着该控制单元拥有过去所有的观测及控制信息,称这种信息结构为严标准模式.对于二人微分对策问题,当 $j=1(N>1)$时的信息结构称为单步延迟共享模式;当 $j=N>1$ 时,所有控制单元之间无信息共享,称该信息模式为无共享模式.

由于信息观测、传输与处理过程中不可避免的会引入延迟,j 步延迟共享模式就成为实际应用中一类非常典型的信息模式.作为 j 步延迟共享模式的特例,严标准模式则是随机最优控制研究中的一类典型信息模式.下面来研究这些信息模式下微分对策问题解的结构.

定理 1　对于 j 步延迟共享模式,记:

$$L_{k,i}^j=\{(p,q)\mid(p,q)\in U_k,q\neq i,k-j<p\leqslant k-1\}.$$

$\forall(k,i)\in U_N$,三元组$(Z_{k,i},U_{k,i},L_{k,i}^j)$、$(Z_{k,P}\cap Z_{k,\mathrm{E}},U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}},\varnothing)$都是状态变量 $\boldsymbol{x}_{k-j+1}$的条件基.

证明　对于 j 步延迟共享模式,有:$Z_{k,\mathrm{P}}\cap Z_{k,\mathrm{E}}=Z_{k-j};U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}}=U_{k-j+1}$,因此 $\forall\gamma\in\Gamma$,给定由(Z_{k-j},U_{k-j+1})生成的数据集后,由式(1)知,$\boldsymbol{x}_{k-j+1}$可递归地表示为 $\boldsymbol{x}_0$ 与部分原始随机变量的函数,记之为 g_{k-j+1},即:

$$\boldsymbol{x}_{k-j+1}=g_{k-j+1}(\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{k-j},\boldsymbol{u}_{U_{k-j+1}}).$$

同理,观测数据集 $z_{Z_{k-j,i}}$ 也可表示为 $\boldsymbol{x}_0$ 与部分原始随机变量的函数,记之为 g_{k-l}^i,即:

$$\boldsymbol{z}_{k-l}^i=g_{k-l}^i(\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{k-l-1},\boldsymbol{u}_{U_{k-l}},\boldsymbol{v}_{k-l}^i);$$
$$l=j,j+1,\cdots k.$$

令 $\boldsymbol{\omega}_{k-j}=(\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{v}_0^{\mathrm{P}},\boldsymbol{v}_0^{\mathrm{E}}\cdots,\boldsymbol{w}_{k-j},\boldsymbol{v}_{k-j}^{\mathrm{P}},\boldsymbol{v}_{k-j}^{\mathrm{E}})$表示 $0\sim k-j$ 时刻所有原始随机变量构成的矢量.定义概率空间$(\boldsymbol{\Omega}_{k-j},F(\boldsymbol{\Omega}_{k-j}),P_\omega)$,由于所有原始随机变量相互独立,故联合概率测度 P_ω 存在且等于各分量概率测度的乘积.

又因函数 f_k、$h_k^i(k=0,1,\cdots N-1)$连续或分段连续,故函数 g_{k-j+1}、$g_{k-l}^i(i=\mathrm{P},\mathrm{E};l=j,j+1,\cdots,k)$为概率空间$(\boldsymbol{\Omega}_{k-j},F(\boldsymbol{\Omega}_{k-j}),P_\omega)$上的 Borel 可测映射,因此 $\boldsymbol{x}_{k-j+1}$、z_{k-l}^i为$(\boldsymbol{\Omega}_{k-j},F(\boldsymbol{\Omega}_{k-j}),P_\omega)$上的随机矢量.令$(\boldsymbol{\Omega}_Z,F(\boldsymbol{\Omega}_Z),P_z)$表示由 Z_{k-j}生成的所有观测量构成的概率空间,$(\boldsymbol{\Omega}_x,F(\boldsymbol{\Omega}_x),P_x)$表示由 $\boldsymbol{x}_{k-j+1}$确定的概率空间.

$\forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$,将关系 $x_i<y_i(i=0,1\cdots n)$记作 $\boldsymbol{x}<\boldsymbol{y}$.给定数据集 $\boldsymbol{u}_{U_{k-j+1}}$后,将 $\boldsymbol{x}_{k-j+1}$条件概率分布记作

$P_x(\boldsymbol{x}_{k-j+1}<\boldsymbol{x}\mid \boldsymbol{u}_{U_{k-j+1}})$. 由于 g_{k-j+1}为$(\boldsymbol{\Omega}_{k-j},F(\boldsymbol{\Omega}_{k-j}))$到$(\boldsymbol{\Omega}_x,F(\boldsymbol{\Omega}_x))$的 Borel 可测映射，所以存在事件 $A\in F(\boldsymbol{\Omega}_{k-j})$，使得：

$$g_{k-j+1}{}^{-1}(\boldsymbol{x}_{k-j+1}<\boldsymbol{x}\mid \boldsymbol{u}_{U_{k-j+1}})=A,$$

$$P_x(\boldsymbol{x}_{k-j+1}<\boldsymbol{x}\mid \boldsymbol{u}_{U_{k-j+1}})=P_\omega(A),$$

故存在函数 $F_1(\boldsymbol{x},\boldsymbol{u}_{U_{k-j+1}})=P_\omega(A)$.

同理，给定 $\boldsymbol{u}_{U_{k-j+1}}$后，存在事件 $B_l^i\in F(\boldsymbol{\Omega}_{k-j})$使得 $g_{k-l}^{i-1}(z_{k-l}^i)=B_l^i(i=\mathrm{P},\mathrm{E};l=j,j+1,\cdots,k)$. 记这些事件的交集为 B，则：$P_\omega(B)=P_z(\boldsymbol{Z}=z_{Z_{k-j}})$，故存在 $F_2(z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})=P_\omega(B)$.

对于事件 A 和 B 的交集，同样存在函数 $F_3(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})=P_\omega(A\cap B)$. 因此：

$$\begin{aligned}P(x_{k-j+1}<\boldsymbol{x}\mid z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})&=\frac{P(\boldsymbol{x}_{k-j+1}<\boldsymbol{x},z_{Z_{k-j}}\mid \boldsymbol{u}_{U_{k-j+1}})}{P(z_{Z_{k-j}}\mid \boldsymbol{u}_{U_{k-j+1}})}\\&=\frac{P_\omega(A\cap B)}{P_\omega(B)}=\frac{F_3(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})}{F_2(z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})}\end{aligned}$$

定义函数 $F(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})$：

$$F(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})=\begin{cases}\dfrac{F_3(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})}{F_2(z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})}, & P_\omega(B)\neq 0\\ 0, & P_\omega(B)=0\end{cases}$$

所以，在给定 σ 域 $F(Z_{k-j},U_{k-j+1},\varnothing)$后，存在 $F(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})$，使得 $\boldsymbol{x}_{k-j+1}$的条件概率分布满足：

$$P(\boldsymbol{x}_{k-j+1}<\boldsymbol{x}\mid z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}})\ \underline{\mathrm{a.e.}}\ F(\boldsymbol{x},z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}}).$$

根据定义 6，三元组$(Z_{k-j},U_{k-j+1},\varnothing)$构成状态变量 $\boldsymbol{x}_{k-j+1}$的条件基. 同理可证，$(Z_{k,i},U_{k,i},L_{k,i}^j)$为 $\boldsymbol{x}_{k-j+1}$的条件基.

定理 2 j 步延迟共享模式下随机微分对策问题的最优控制可表示为：

$$\gamma_{k,i}^*(\delta_k,\lambda_{k,i})=\varphi_{k,i}^*(\lambda_{k,i},F_k(\delta_k)),$$

其中：$\delta_k=\{z_{Z_{k-j}},\boldsymbol{u}_{U_{k-j+1}}\}$，表示 k 时刻 P 和 E 的共享信息；$\lambda_{k,i}=\boldsymbol{u}_{A_k^i}\cup z_{B_k^i}$，表示单元 i 的独有信息，$A_k^i=\{(p,i)\mid k-j<p\leqslant k-1\}$，$B_k^i=\{(p,i)\mid k-j<p\leqslant k\}$；$F_k(\delta_k)$为 $\boldsymbol{x}_{k-j+1}$的条件概率分布.

证明 根据定理 1，$(Z_{k,\mathrm{P}}\cap Z_{k,\mathrm{E}},U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}},\varnothing)$为 $\boldsymbol{x}_{k-j+1}$的条件基. 若上述 j 步延迟共享模式下的最优设计为 γ^*，根据双方极值原理及 Bellman 原理，首先考虑单元 P 最后一步的优化：

$$\begin{aligned}\hat{J}(\gamma_{N-1}^{\mathrm{P}\,*},\gamma^{\mathrm{E}*})=\inf_{\boldsymbol{u}_{N-1}^{\mathrm{P}}}\{\ &E[(G(\boldsymbol{x}_{N-1},\boldsymbol{u}_{N-1}^{\mathrm{P}},\boldsymbol{u}_{N-1}^{\mathrm{E}*})\\&+\phi\circ f_{N-1}(\boldsymbol{x}_{N-1},\boldsymbol{u}_{N-1}^{\mathrm{P}},\boldsymbol{u}_{N-1}^{\mathrm{E}*},\\&\boldsymbol{w}_{N-1}))\mid\lambda_{N-1,\mathrm{P}},\delta_{N-1}]\}\end{aligned}$$

由式(1)知 $\boldsymbol{x}_{N-1}$可递归地表示为 $\boldsymbol{x}_{N-1}$与部分原始随机变量的可测函数，记之为 g_{N-1}. 同时记：

$$\boldsymbol{\omega}_{N-1}^{\mathrm{P}}=(\boldsymbol{v}_{N-j}^{\mathrm{P}},\boldsymbol{w}_{N-j},\cdots,\boldsymbol{v}_{N-1}^{\mathrm{P}},\boldsymbol{w}_{N-1}),$$

$$\boldsymbol{V}_{N-1}^{\mathrm{E}}=(\boldsymbol{v}_{N-j}^{\mathrm{E}},\cdots,\boldsymbol{v}_{N-1}^{\mathrm{E}}),$$

则：

$$\begin{aligned}\gamma_{N-1}^{\mathrm{P}*}(\delta_{N-1},\lambda_{N-1,\mathrm{P}})=\arg\inf_{\boldsymbol{u}_{N-1}^{\mathrm{P}}}\{&E[g_{N-1}(\boldsymbol{\omega}_{N-1}^{P},\boldsymbol{x}_{N-j},\\&\boldsymbol{u}_{A_{N-1}^{\mathrm{P}}},\boldsymbol{u}_{N-1}^{\mathrm{P}},\boldsymbol{u}_{A_N^{\mathrm{E}}}^*)\mid\lambda_{N-1,\mathrm{P}},\delta_{N-1}]\}\end{aligned}\tag{2}$$

由于逃逸者 E 的控制函数为状态和观测的确定性可测函数，因此 $\boldsymbol{u}_{A_N^{\mathrm{E}}}^*$可表示为$\boldsymbol{V}_{N-1}^{\mathrm{E}}$、$\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$、$\boldsymbol{x}_{N-j}$的函数. 显然，$\boldsymbol{V}_{N-1}^{\mathrm{E}}$独立于 $\lambda_{N-1,\mathrm{P}}$、δ_{N-1}、$\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$、$\boldsymbol{x}_{N-j}$，因此式(2)中 g_{N-1}的求期望可先对 $\boldsymbol{V}_{N-1}^{\mathrm{E}}$进行，对 $\boldsymbol{V}_{N-1}^{\mathrm{E}}$求期望后的函数记为 $\bar{g}_{N-1}$. 给定 δ_{N-1}、$\boldsymbol{u}_{A_{N-1}^{\mathrm{P}}}$的条件下，$z_{B_{N-1}^{\mathrm{P}}}$中的每个元素都可表示为 $\boldsymbol{x}_{N-j}$与$\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$的 Borel 可测函数，故 $z_{B_{N-1}^{\mathrm{P}}}$的原象集为 $\boldsymbol{x}_{N-j}$与 $\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$联合概率空间中的事件. 记该事件为 S_{N-1}，则式(2)可进一步表示成式(3)：

$$\gamma_{N-1}^{\mathrm{P}*}(\delta_{N-1},\lambda_{N-1,\mathrm{P}})=\frac{\arg\inf_{\boldsymbol{u}_{N-1}^{\mathrm{P}}}\{\oiint_{S_{N-1}}\bar{g}_{N-1}(\boldsymbol{\omega}_{N-1}^{\mathrm{P}},\boldsymbol{x}_{N-j},\boldsymbol{u}_{A_{N-1}^{\mathrm{P}}},\boldsymbol{u}_{N-1}^{\mathrm{P}})\cdot pdf_{\boldsymbol{\omega}_{N-1}^{\mathrm{P}}}(\boldsymbol{\omega}_{N-1}^{\mathrm{P}})\mathrm{d}\boldsymbol{\omega}_{N-1}^{\mathrm{P}}\cdot\mathrm{d}F_{N-1}(\delta_{N-1})\}}{\oiint_{S_{N-1}}pdf_{\omega_{N-1}^{\mathrm{P}}}(\boldsymbol{\omega}_{N-1}^{\mathrm{P}})\mathrm{d}\boldsymbol{\omega}_{N-1}^{\mathrm{P}}\cdot\mathrm{d}F_{N-1}(\delta_{N-1})}\tag{3}$$

式(3)的分母表示 S_{N-1}的概率测度，它是 δ_{N-1}、$\boldsymbol{u}_{A_{N-1}^{\mathrm{P}}}$、$z_{B_{N-1}^{\mathrm{P}}}$的函数，与 $\boldsymbol{u}_{N-1}^{\mathrm{P}}$无关. 另外，给定 $\boldsymbol{x}_{N-j}$后，$\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$的积分域为$\{\boldsymbol{\omega}_{N-1}^{\mathrm{P}}\mid(\boldsymbol{\omega}_{N-1}^{\mathrm{P}},\boldsymbol{x}_{N-j})\in S_{N-1}\}$，将 $\bar{g}_{N-1}$对 $\boldsymbol{\omega}_{N-1}^{\mathrm{P}}$积分后的函数记为 $\bar{\bar{g}}_{N-1}$. 即得到：

$$\begin{aligned}\gamma_{N-1}^{\mathrm{P}*}(\delta_{N-1},\lambda_{N-1,\mathrm{P}})&=\arg\inf_{\boldsymbol{u}_{N-1}^{\mathrm{P}}}\{\int_{\boldsymbol{x}_{N-j}}\bar{\bar{g}}_{N-1}(\boldsymbol{x}_{N-j},\lambda_{N-1,\mathrm{P}},\\&\qquad\boldsymbol{u}_{N-1}^{\mathrm{P}})\mathrm{d}F_{N-1}(\delta_{N-1})\}\\&=\varphi_{N-1,\mathrm{P}}^*(\lambda_{N-1,\mathrm{P}},F_{N-1}(\delta_{N-1}))\end{aligned}\tag{4}$$

将式(4)的最优控制 $\boldsymbol{u}_{N-1}^{\mathrm{P}*}$代入 $\hat{J}(\gamma_{N-1}^{\mathrm{P}*},\gamma^{\mathrm{E}*})$中，可知 $\hat{J}(\gamma_{N-1}^{\mathrm{P}*},\gamma^{\mathrm{E}*})$亦为 $\lambda_{N-1,\mathrm{P}}$与 δ_{N-1}的函数.

依次后向归纳，便可证明对于 $k=0,\cdots,N-1$，$\gamma_{k,\mathrm{P}}^*$都可表示为 $\varphi_{k,\mathrm{P}}^*(\lambda_{k,\mathrm{P}},F_k(\delta_k))$的形式. 同理可证，$\gamma_{k,\mathrm{E}}^*$亦可表示为 $\varphi_{k,\mathrm{E}}^*(\lambda_{k,\mathrm{E}},F_k(\delta_k))$的形式.

定理 3 对于线性高斯微分对策问题，当信息模式为 j 步延迟共享模式时，给定由$(Z_{k,\mathrm{P}}\cap Z_{k,\mathrm{E}},U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}})$生成的数据集时：

(1) $\boldsymbol{x}_{k-j+1}$的条件概率分布 F_k 为高斯分布，条件均值 $\bar{F}_k(\delta_k)$为已知数据的线性变换，且线性变换的系数

及条件协方差都与已知数据无关;

(2)对策双方的最优控制可以表示为 $\gamma_{k,i}^{*}(\delta_k,\lambda_{k,i}) = \varphi_{k,i}^{*}(\lambda_{k,i},\bar{F}_k(\delta_k))$.

证明　根据定理 1,$(Z_{k,\mathrm{P}}\cap Z_{k,\mathrm{E}}, U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}}, \varnothing)$为 $\boldsymbol{x}_{k-j+1}$的条件基.对于线性系统,$\boldsymbol{x}_{k-j+1}$可递归地表示为随机矢量 $\boldsymbol{\omega}_{k-j}$(具体定义见定理 1 的证明)与 $\boldsymbol{u}_{U_{k-j+1}}$的线性函数;而观测 $\boldsymbol{z}_k^i$ 也可表示为$\boldsymbol{\omega}_{k-j}$与$\boldsymbol{u}_{U_{k-j+1}}$的线性函数.所以:

$$[\,\boldsymbol{z}_0^{\mathrm{P}} \quad \boldsymbol{z}_0^{\mathrm{E}} \quad \cdots \quad \boldsymbol{z}_{k-j}^{\mathrm{P}} \quad \boldsymbol{z}_{k-j}^{\mathrm{E}} \quad \boldsymbol{x}_{k-j+1}\,]^{\mathrm{T}}$$
$$= \boldsymbol{\Phi}\boldsymbol{\omega}_{k-j} + \boldsymbol{B}\cdot[\,\boldsymbol{u}_0^{\mathrm{P}} \quad \boldsymbol{u}_0^{\mathrm{E}} \quad \cdots \quad \boldsymbol{u}_{k-j}^{\mathrm{P}} \quad \boldsymbol{u}_{k-j}^{\mathrm{E}}\,]^{\mathrm{T}}$$

已知 $\boldsymbol{u}_{U_{k-j+1}}$的条件下,由于$\boldsymbol{\omega}_{k-j}$为各分量相互独立的高斯随机变量,所以 $\boldsymbol{x}_{k-j+1}$、$\boldsymbol{z}_{Z_{k-j}}$亦为高斯随机变量.由引理 1 可知,已知 $\boldsymbol{z}_{Z_{k-j}}$、$\boldsymbol{u}_{U_{k-j+1}}$的条件下,$\boldsymbol{x}_{k-j+1}$服从高斯分布,其条件均值 $\bar{F}_k(\delta_k)$为已知数据 δ_k 的线性变换,且线性变换的系数及条件协方差都与已知数据 δ_k 无关,结论 1 得证.

参考定理 2 的证明,单元 P 倒数第 l 步的优化可表示为:

$$\gamma_{N-l}^{\mathrm{P}*}(\delta_{N-l},\lambda_{N-l,\mathrm{P}}) = \arg\inf_{\boldsymbol{u}_{N-l}^{\mathrm{P}}}\Big\{\int_{\boldsymbol{x}_{N-l-j+1}} \bar{\bar{g}}_{N-l}(\boldsymbol{x}_{N-l-j+1}, \boldsymbol{u}_{A_{N-l}^{\mathrm{P}}}, \boldsymbol{u}_{N-l}^{\mathrm{P}})\mathrm{d}F_{N-l}(\delta_{N-l})\Big\} \quad (5)$$

已知 δ_{N-l}的条件下,$\boldsymbol{x}_{N-l-j+1}$服从高斯分布.令 $\boldsymbol{x}_{N-l-j+1} = \bar{\boldsymbol{x}}_{N-l-j+1} + \boldsymbol{x}$,则 $\boldsymbol{x}$ 服从均值为零,方差为 $\boldsymbol{\Sigma}_{N-l-j+1}$的高斯分布.式(5)即可写成:

$$\begin{aligned}\gamma_{N-l}^{\mathrm{P}*}(\delta_{N-l},\lambda_{N-l,\mathrm{P}}) &= \arg\inf_{\boldsymbol{u}_{N-l}^{\mathrm{P}}}\Big\{\int_{\boldsymbol{x}} \bar{\bar{g}}_{N-l}(\bar{\boldsymbol{x}}_{N-l-j+1},\boldsymbol{x}, \boldsymbol{u}_{A_{N-l}^{\mathrm{P}}}, \boldsymbol{u}_{N-l}^{\mathrm{P}})\cdot pdf(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\Big\} \\ &= \arg\inf_{\boldsymbol{u}_{N-l}^{\mathrm{P}}}\{\bar{\bar{\bar{g}}}_{N-l}(\bar{\boldsymbol{x}}_{N-l-j+1}, \boldsymbol{u}_{A_{N-l}^{\mathrm{P}}}, \boldsymbol{u}_{N-l}^{\mathrm{P}})\} \\ &= \varphi_{N-l,1}^{*}(\lambda_{N-l,\mathrm{P}},\bar{F}_{N-l}(\delta_{N-l}))\end{aligned}$$

结论 2 得证.

定理 4　对于单步延迟共享模式下无控制约束的线性二次高斯微分对策(LQGDG)问题,双方的最优控制 $\gamma_{k,i}^{*}$、$\varphi_{k,i}^{*}$是已知数据的线性函数.

证明　式(1)的线性二次形式可表示如下:

$$\begin{aligned}&\boldsymbol{x}_{k+1} = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k^{\mathrm{P}}\boldsymbol{u}_k^{\mathrm{P}} + \boldsymbol{B}_k^{\mathrm{E}}\boldsymbol{u}_k^{\mathrm{E}} + \boldsymbol{w}_k \\ &\boldsymbol{z}_k^{\mathrm{P}} = \boldsymbol{H}_k^{\mathrm{P}}\boldsymbol{x}_k + \boldsymbol{v}_k^{\mathrm{P}} \\ &\boldsymbol{z}_k^{\mathrm{E}} = \boldsymbol{H}_k^{\mathrm{E}}\boldsymbol{x}_k + \boldsymbol{v}_k^{\mathrm{E}} \\ &J = E[\,\boldsymbol{x}_N^{\mathrm{T}}\boldsymbol{P}_N\boldsymbol{x}_N \\ &\quad + \sum_{k=0}^{N-1}(\boldsymbol{x}_k^{\mathrm{T}}\bar{\boldsymbol{Q}}_k\boldsymbol{x}_k + (\boldsymbol{u}_k^{\mathrm{P}})^{\mathrm{T}}\bar{\boldsymbol{R}}_k^{\mathrm{P}}\boldsymbol{u}_k^{\mathrm{P}} - (\boldsymbol{u}_k^{\mathrm{E}})^{\mathrm{T}}\bar{\boldsymbol{R}}_k^{\mathrm{E}}\boldsymbol{u}_k^{\mathrm{E}})]\end{aligned} \quad (6)$$

其中:$\boldsymbol{P}_N$、$\bar{\boldsymbol{Q}}_k$、$\bar{\boldsymbol{R}}_k$ 为相应的加权矩阵,$\boldsymbol{P}_N$、$\bar{\boldsymbol{Q}}_k$ 对称半正定,$\bar{\boldsymbol{R}}_k^{\mathrm{P}}$ 对称正定、$\bar{\boldsymbol{R}}_k^{\mathrm{E}}$ 对称负定.下面只需证明定理 3 中的 $\varphi_{k,i}^{*}$为$\lambda_{k,i}$与$\bar{F}_k(\delta_k)$的线性函数.为简化表示,令:

$$\boldsymbol{u}_k = [\,\boldsymbol{u}_k^{\mathrm{P}},\boldsymbol{u}_k^{\mathrm{E}}\,]^{\mathrm{T}};\boldsymbol{B}_k = [\,\boldsymbol{B}_k^{\mathrm{P}},\boldsymbol{B}_k^{\mathrm{E}}\,];\bar{\boldsymbol{R}}_k = \begin{bmatrix}\bar{\boldsymbol{R}}_k^{\mathrm{P}} & \boldsymbol{0}\\ \boldsymbol{0} & -\bar{\boldsymbol{R}}_k^{\mathrm{E}}\end{bmatrix}$$

在单步延迟共享模式下,$\lambda_{k,i} = \boldsymbol{z}_k^i$,$\delta_k = \boldsymbol{u}_{U_k}\cup \boldsymbol{z}_{Z_{k-1}}$.考虑单元 P 最后一步的优化:

$$\begin{aligned}\boldsymbol{u}_{N-1}^{\mathrm{P}*} &= \gamma_{N-1}^{\mathrm{P}*}(\delta_{N-1},\boldsymbol{z}_{N-1}^{\mathrm{P}}) \\ &= -\underbrace{(\bar{\boldsymbol{R}}_{N-1}^{\mathrm{P}} + (\boldsymbol{B}_{N-1}^{\mathrm{P}})^{\mathrm{T}}\boldsymbol{P}_N\boldsymbol{B}_{N-1}^{\mathrm{P}})^{-1}(\boldsymbol{B}_{N-1}^{\mathrm{P}})^{\mathrm{T}}}_{\boldsymbol{\Lambda}_{N-1}^{\mathrm{P}}}\boldsymbol{P}_N \\ &\quad\cdot(E[\,\boldsymbol{A}_{N-1}\boldsymbol{x}_{N-1} + \boldsymbol{B}_{N-1}^{\mathrm{E}}\boldsymbol{u}_{N-1}^{\mathrm{E}*}\mid \boldsymbol{z}_{N-1}^{\mathrm{P}},\delta_{N-1}] + \bar{\boldsymbol{w}}_{N-1}) \\ &= -\boldsymbol{\Lambda}_{N-1}^{\mathrm{P}}\boldsymbol{P}_N(\boldsymbol{A}_{N-1}E[\,\boldsymbol{x}_{N-1}\mid \boldsymbol{z}_{N-1}^{\mathrm{P}},\delta_{N-1}]) \\ &\quad + \boldsymbol{B}_{N-1}^{\mathrm{E}}E[\,\boldsymbol{u}_{N-1}^{\mathrm{E}*}(\boldsymbol{z}_{N-1}^{\mathrm{E}},\delta_{N-1})\mid \boldsymbol{z}_{N-1}^{\mathrm{P}},\delta_{N-1}] + \bar{\boldsymbol{w}}_{N-1})\end{aligned}$$

由于 $\boldsymbol{u}_{N-1}^{\mathrm{P}*}$表达式中的时间下标全为 $N-1$(除 $\boldsymbol{P}_N$ 外),为简便起见,后面表述中省去时间下标并将身份标识下移.线性高斯条件下,根据卡尔曼滤波方程可以得到:

$$\begin{aligned}\boldsymbol{u}_{\mathrm{P}}^{*} &= -\boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}\boldsymbol{A}[\,\bar{\boldsymbol{x}} + \boldsymbol{K}_{\mathrm{P}}(\boldsymbol{z}_{\mathrm{P}} - (\boldsymbol{H}_{\mathrm{P}}\bar{\boldsymbol{x}} + \bar{\boldsymbol{v}}_{\mathrm{P}}))] \\ &\quad - \boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}\boldsymbol{B}_{\mathrm{E}}\cdot E[\,\boldsymbol{u}_{\mathrm{E}}^{*}(\boldsymbol{z}_{\mathrm{E}},\delta)\mid \boldsymbol{z}_{\mathrm{P}},\delta] - \boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}\bar{\boldsymbol{w}}\end{aligned}$$

其中:$\bar{\boldsymbol{x}} = \bar{F}_k(\delta_k)$,表示状态预测;$\boldsymbol{K}_{\mathrm{P}}$ 为相应的卡尔曼增益.同理,可得 $\boldsymbol{u}_{\mathrm{E}}^{*}$ 如下:

$$\begin{aligned}\boldsymbol{u}_{\mathrm{E}}^{*} &= -\boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P}\boldsymbol{A}[\,\bar{\boldsymbol{x}} + \boldsymbol{K}_{\mathrm{E}}(\boldsymbol{z}_{\mathrm{E}} - (\boldsymbol{H}_{\mathrm{E}}\bar{\boldsymbol{x}} + \bar{\boldsymbol{v}}_{\mathrm{E}}))] \\ &\quad - \boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P}\boldsymbol{B}_{\mathrm{P}}\cdot E[\,\boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{z}_{\mathrm{P}},\delta)\mid \boldsymbol{z}_{\mathrm{E}},\delta] - \boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P}\bar{\boldsymbol{w}}\end{aligned}$$

将 $\boldsymbol{u}_{\mathrm{E}}^{*}$ 带入 $\boldsymbol{u}_{\mathrm{P}}^{*}$,可得:

$$\begin{aligned}\boldsymbol{u}_{\mathrm{P}}^{*} &= \boldsymbol{\Lambda}_{\mathrm{P}}[\,\boldsymbol{P}\boldsymbol{B}_{\mathrm{E}}\boldsymbol{\Lambda}_{\mathrm{E}} - \boldsymbol{I}\,]\boldsymbol{P}\boldsymbol{A}\bar{\boldsymbol{x}} + (\boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}\boldsymbol{B}_{\mathrm{E}}\boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P} - \boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P})\bar{\boldsymbol{w}} \\ &\quad + \boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}[\,\boldsymbol{B}_{\mathrm{E}}\boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P}\boldsymbol{A}\boldsymbol{K}_{\mathrm{E}}\boldsymbol{H}_{\mathrm{E}} - \boldsymbol{A}\,]\boldsymbol{K}_{\mathrm{P}}[\,\boldsymbol{z}_{\mathrm{P}} - (\boldsymbol{H}_{\mathrm{P}}\bar{\boldsymbol{x}} + \bar{\boldsymbol{v}}_{\mathrm{P}})] \\ &\quad + \boldsymbol{\Lambda}_{\mathrm{P}}\boldsymbol{P}\boldsymbol{B}_{\mathrm{E}}\boldsymbol{\Lambda}_{\mathrm{E}}\boldsymbol{P}\boldsymbol{B}_{\mathrm{P}}E[\,E[\,\boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{y},\delta)\mid \boldsymbol{z}_{\mathrm{E}},\delta]\mid \boldsymbol{z}_{\mathrm{P}},\delta]\end{aligned} \quad (7)$$

式(7)中的变量 $\boldsymbol{y}$ 代表观测矢量 $\boldsymbol{z}_P$;数学期望 $E[\,E[\,\boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{y},\delta)\mid \boldsymbol{z}_{\mathrm{E}},\delta]\mid \boldsymbol{z}_{\mathrm{P}},\delta]$可表示为:

$$\begin{aligned}&E[\,E[\,\boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{y},\delta)\mid \boldsymbol{z}_{\mathrm{E}},\delta]\mid \boldsymbol{z}_{\mathrm{P}},\delta] \\ &= \int_{-\infty}^{+\infty}\underbrace{\Big(\int_{-\infty}^{+\infty} pdf_{\boldsymbol{z}_{\mathrm{P}}}(\boldsymbol{y}\mid \boldsymbol{z}_{\mathrm{E}},\delta)\cdot pdf_{\boldsymbol{z}_{\mathrm{E}}}(\boldsymbol{z}_{\mathrm{E}}\mid \boldsymbol{z}_{\mathrm{P}},\delta)\mathrm{d}\boldsymbol{z}_{\mathrm{E}}\Big)}_{K(\boldsymbol{y},\boldsymbol{z}_{\mathrm{P}})} \\ &\quad\cdot \boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{y},\delta)\cdot\mathrm{d}\boldsymbol{y}\end{aligned} \quad (8)$$

由引理 2 可知,式(8)中的积分核函数 $K(\boldsymbol{y},\boldsymbol{z}_{\mathrm{P}})$为高斯核,且均值为 $\boldsymbol{z}_{\mathrm{P}}$ 和 $\bar{\boldsymbol{x}}$ 的线性函数,故 $\boldsymbol{u}_{\mathrm{P}}^{*}$ 满足如下 Fredholm 第二类积分方程:

$$\boldsymbol{u}_{\mathrm{P}}^{*} = \boldsymbol{L}\bar{\boldsymbol{x}} + \boldsymbol{M}\boldsymbol{z}_{\mathrm{P}} + \boldsymbol{N} + \boldsymbol{O}\int_{-\infty}^{+\infty}\boldsymbol{u}_{\mathrm{P}}^{*}(\boldsymbol{y},\delta)\cdot K(\boldsymbol{y},\boldsymbol{z}_{\mathrm{P}})\mathrm{d}\boldsymbol{y} \quad (9)$$

式(9)中,$\boldsymbol{L}$、$\boldsymbol{M}$、$\boldsymbol{N}$、$\boldsymbol{O}$ 分别表示系数矩阵.方程式(9)的通解具有形如 $\boldsymbol{u}_{\mathrm{P}}^{*} = \boldsymbol{A}\bar{\boldsymbol{x}} + \boldsymbol{B}\boldsymbol{z}_{\mathrm{P}} + \boldsymbol{C}$ 的线性形式.同理可证,$\boldsymbol{u}_{\mathrm{E}}^{*}$ 也具有类似形式.所以对于最后一步的优化,$\varphi_{N-1,i}^{*}$ 为 $\lambda_{N-1,i}$、$\bar{F}_{N-1}(\delta_{N-1})$ 的线性函数.由于 $\bar{F}_{N-1}(\delta_{N-1})$为 δ_{N-1}的线性函数,故 $\gamma_{k,i}^{*}$为$\lambda_{N-1,i}$、δ_{N-1}的线性函数.

对于倒数第 k 步的优化，采用类似方法可以证明，$\varphi^*_{N-k,i}$ 为 $\lambda_{N-k,i}$、$\bar{F}_{N-k}(\delta_{N-k})$ 的线性函数，$\gamma^*_{N-k,i}$ 为 $\lambda_{N-k,i}$、δ_{N-k} 的线性函数.

下面考虑一类特殊的微分对策. 当式(1)的对策问题中仅有一个控制单元时，二人随机微分对策就退化为一个随机最优控制问题. 对于严标准模式下的随机最优控制问题，有以下定理.

定理 5 对于严标准模式下的随机控制，将 γ_k 表示为 $\gamma_k=\varphi_k\circ F_k$ 的形式不影响控制的最优性，即最优设计 γ^*_k 可表示为 $\gamma^*_k=\varphi^*_k\circ F_k$ 的形式. 这样 γ^*_k 的设计便分解为确定 φ^*_k 和 F_k 两个独立的问题.

证明 略(类似定理 2 的证明).

定理 6(确定性等价原理) 对于严标准模式下无控制约束的线性二次系统，最优设计 $\gamma^*_k=\varphi^{**}_k\circ\bar{F}_k$. 其中：$\varphi^{**}_k$ 表示原始随机变量取其均值所对应的确定性等效系统的最优控制律，为已知数据的线性函数.

证明 略(请参考文献[5]).

推论 1 对于严标准模式下的线性高斯(LG)系统，最优控制 $\gamma^*_k=\varphi^*_k\circ\bar{F}_k$. 其中：$\bar{F}_k$ 为 $\boldsymbol{x}_k$ 的条件均值，是已知数据的线性变换；φ^*_k 通常不同于定理 6 中的 φ^{**}_k.

证明 根据定理 5，严标准模式下的最优控制可表示成 $\gamma^*_k=\varphi^*_k\circ F_k$ 的形式. 考虑倒数第 l 步的优化：

$$\gamma^*_{N-l}(z_{N-l},\boldsymbol{u}_{N-l})=\arg\min_{\boldsymbol{u}_{N-l}}\left\{\int_{\boldsymbol{x}_{N-l}}\left[G(\boldsymbol{x}_{N-l},\boldsymbol{u}_{N-l})+\bar{\eta}^J_{N-l}(\boldsymbol{x}_{N-l},\boldsymbol{u}_{N-l})\right]\cdot \mathrm{d}F_{N-l}\right\}$$

对于 LG 系统，已知 δ_{N-l} 的条件下，$\boldsymbol{x}_{N-l}$ 服从 $(\bar{\boldsymbol{x}}_{N-l},\boldsymbol{\Sigma}_{N-l})$ 的高斯分布. 令 $\boldsymbol{x}_{N-l}=\bar{\boldsymbol{x}}_{N-l}+\boldsymbol{x}$，则 $\boldsymbol{x}$ 服从均值为零，方差为 $\boldsymbol{\Sigma}_{N-l}$ 的高斯分布，则：

$$\begin{aligned}\gamma^*_{N-l}(z_{N-l},\boldsymbol{u}_{N-l})&=\arg\min_{\boldsymbol{u}_{N-l}}\left\{\int_{\boldsymbol{x}}\left[G(\bar{\boldsymbol{x}}_{N-l}+\boldsymbol{x},\boldsymbol{u}_{N-l})+\bar{\eta}^J_{N-l}(\bar{\boldsymbol{x}}_{N-l}+\boldsymbol{x},\boldsymbol{u}_{N-l})\right]\cdot pdf(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right\}\\&=\arg\min_{\boldsymbol{u}_{N-l}}\left\{\bar{G}(\bar{\boldsymbol{x}}_{N-l},\boldsymbol{u}_{N-l})+\bar{\bar{\eta}}^J_{N-l}(\bar{\boldsymbol{x}}_{N-l},\boldsymbol{u}_{N-l})\right\}\\&=\varphi^*_{N-l}(\bar{\boldsymbol{x}}_{N-l})=\varphi^*_{N-l}\circ\bar{F}_{N-l}(z_{N-l},\boldsymbol{u}_{N-l})\end{aligned}$$

所以，严标准模式下 LG 系统的最优控制可表示为 $\gamma^*_k=\varphi^*_k\circ\bar{F}_k$，且 φ^*_k 通常不同于定理 6 中 φ^{**}_k.

推论 2 对于严标准模式下无控制约束的线性二次高斯(LQG)系统，最优控制 $\gamma^*_k=\varphi^{**}_k\circ\bar{F}_k$，并且 γ^*_k 为已知数据的线性函数.

证明 根据定理 6，γ^*_k 可表示为 $\gamma^*_k=\varphi^{**}_k\circ\bar{F}_k$ 的形式，其中 $\bar{F}_k$ 为 $\boldsymbol{x}_{N-1}$ 的条件均值. 对于线性高斯系统，$\bar{F}_k$ 为已知数据 z_{N-l}、$\boldsymbol{u}_{N-l}$ 的线性变换，所以最优控制 γ^*_k 为已知数据的线性函数.

5 结束语

本文在给出离散时间二人随机微分对策问题信息模式及其相关概念的数学定义后，证明了与几种典型信息模式及其相应微分对策解的结构相关的定理. 现就各定理作如下说明：

(1)定理 1 关于条件基的结论对于随机微分对策状态估计器的设计具有理论指导意义. 由于$(Z_{k,\mathrm{P}}\cap Z_{k,\mathrm{E}},U_{k,\mathrm{P}}\cap U_{k,\mathrm{E}},\varnothing)$为状态 $\boldsymbol{x}_{k-j+1}$ 的条件基，所以其估计器的设计可独立于控制律. 特别地，当 $j=1$ 时，由于 $L^j_{k,i}=\varnothing$，给定信息模式所生成的数据集$\{z_{Z_{k,i}},\boldsymbol{u}_{U_{k,i}}\}$后，状态变量 $\boldsymbol{x}_k$ 的条件分布独立于设计 γ，因此，估计器设计可完全独立于控制律. 该结论对于二人对策问题及 MIMO 随机系统最优控制非常重要.

(2)定理 2 说明了 j 步延迟共享模式下微分对策最优解的结构. 由结论可知，最优控制依赖于 $\boldsymbol{x}_{k-j+1}$ 的概率分布及对策双方的独有信息 $\lambda_{k,i}$，最优控制律不能独立于估计器. 基于此，某些文献中称定理 1 为部分分离原理[14]. 由于线性高斯系统的典型性，定理 3、4 给出了 j 步延迟共享模式下线性高斯系统最优控制的形式.

(3)定理 5、6 与推论 1、2 描述了严标准模式下最优控制的结构. 定理 6 即确定性等价原理，指出了随机最优控制与相应的确定性等效系统最优控制间的联系. 根据定理 6，最优控制律与状态估计器可相互独立设计. 目前的随机系统最优控制中广泛采用了这一结论. 特别地，推论 1 指出了线性高斯系统最优控制的形式，推论 2 关于线性二次高斯系统的论断，常被视作确定性等价原理(或分离原理[5])，它其实只是定理 6 在高斯噪声下的特殊情况. 应用中需注意每个定理的条件，定理 5 与推论 1 中的 φ^*_k 一般不等同于定理 6 与推论 2 中的 φ^{**}_k，此时的最优控制不能独立于状态估计.

在目前随机微分对策的应用中，控制律设计大都基于确定性等价原理(定理 6). 然而，实际应用中由于存在控制饱和、非二次指标、系统非线性等原因，定理 6 的条件并不成立，因而基于确定性等价原理的设计并不能获得最优的性能. 但就目前而言，非完美信息模式下最优对策问题求解仍然是一个极富挑战性的难题[15]，对策问题的最优解析解很难获得或者根本就不存在，此时更多的需要借助数值解法. 文献[15]提出了一种基于多智能体的搜索寻优方法，结论表明对于零和追－逃微分对策，该方法的性能优于传统解析方法. 因此，深刻认识理解信息模式的相关概念，研究特定信息模式下最优解的结构，可为设计快速有效的数值解法提供理论依据，这正是本文研究工作的一个重要出发点.

附录 引理 1 和 2

引理 1 若矢量 $\boldsymbol{x}$ 为 n 维联合高斯分布，已知 $\boldsymbol{x}$ 的

部分分量 $\boldsymbol{x}_1=[x_{i+1},\cdots,x_n]^{\mathrm{T}}$ 后,未知部分 $\boldsymbol{x}_0=[x_1,x_2,\cdots,x_i]^{\mathrm{T}}$ 的条件概率密度仍为联合高斯分布,$\boldsymbol{x}_0$ 的条件均值为 $\boldsymbol{x}_1$ 的线性变换,且线性变换的系数及条件协方差阵均与 $\boldsymbol{x}_1$ 的取值无关.

引理 2　若矢量 $\boldsymbol{x}\in\boldsymbol{R}^n$、$\boldsymbol{y}\in\boldsymbol{R}^m$ 服从联合高斯分布,其均值和协方差阵分别为:

$$\boldsymbol{\mu}=\begin{bmatrix}\boldsymbol{\mu}_x\\ \boldsymbol{\mu}_y\end{bmatrix};\ \boldsymbol{\Sigma}=\begin{bmatrix}\boldsymbol{C}_{xx} & \boldsymbol{C}_{xy}\\ \boldsymbol{C}_{yx} & \boldsymbol{C}_{yy}\end{bmatrix}$$

将已知 $\boldsymbol{x}$ 时 $\boldsymbol{y}$ 的条件概率密度函数记为 $pdf_{\boldsymbol{y}}$,已知 y 时 $\boldsymbol{x}$ 的条件概率密度函数计作 $pdf_{\boldsymbol{x}}$,则式($A1$)所示的积分函数 $K(\boldsymbol{x},\boldsymbol{z})$为 $\boldsymbol{x}$ 的高斯函数,且均值为参变量 $\boldsymbol{z}$ 和 $\boldsymbol{\mu}_x$ 的线性函数.

$$K(\boldsymbol{x},\boldsymbol{z})=\int_{\boldsymbol{y}} pdf_{\boldsymbol{y}}(\boldsymbol{y}\mid\boldsymbol{z})\cdot pdf_{\boldsymbol{x}}(\boldsymbol{x}\mid\boldsymbol{y})\cdot \mathrm{d}\boldsymbol{y},\ \boldsymbol{z}\in\mathbf{R}^n \quad (A1)$$

(引理的证明从略).

参考文献:

[1] Isaacs R. Difference Games[M]. New York: Willey, 1965.

[2] 张嗣瀛. 微分对策[M]. 北京:科学出版社, 1987.
Zhang Siying. Difference Games[M]. Beijing: Science Press, 1987. (in Chinese)

[3] Basar T, Olsder G J. Dynamic Noncooperative Game Theory[M]. London and San Diego: Academic Press, 1995.

[4] 李登峰. 微分对策及其应用[M]. 北京:国防工业出版社, 2000.
Li Dengfeng. Differential Games and Applications[M]. Beijing: National Defense Industry Press, 2000. (in Chinese)

[5] 张洪钺, 王青. 最优控制理论与应用[M]. 北京:高等教育出版社, 2006.
Zhang Hongyue, Wang Qing. Optimal Control Theory and Application[M]. Beijing: High Education Press, 2006. (in Chinese)

[6] Witsenhausen H S. Separation of estimation and control for discrete time systems[J]. Proceedings of the IEEE, 1971, 59(11): 1557 – 1566.

[7] Behn R D, Ho Y C. On a class of linear stochastic difference games[J]. IEEE Trans on AC, 1968, 13(3): 227 – 240.

[8] Rhodes I B, Luenberger D G. Differential games with imperfect state information[J]. IEEE Trans on AC, 1969, 14(1): 29 – 38.

[9] Willman W W. Formal solutions for a class of stochastic Pursuit-Evasion games[J]. IEEE Trans on AC, 1969, 14(5): 504 – 509.

[10] Shinar J, Shima T. A game theoretical interceptor guidance law for ballistic missile defense[A]. Proceedings of the 35th Conf on Decision and Control[C]. Kobe, Japan: IEEE Press, 1996. 2780 – 2785.

[11] Shinar J, Shima T. Robust missile guidance law against highly maneuvering targets[A]. Proceedings of the 7th Mediterranean Control and Automation[C]. Haifa, Israel: IEEE Press, 1999. 548 – 1572.

[12] Oshman Y, David A. Differential-Game-Based guidance law using target orientation observations[J]. IEEE Trans on AES, 2006, 42(1): 316 – 326.

[13] Swarup A, Speyer J L. Characterization of LQG differential games with different information patterns[A]. IEEE Conf on Decision and Control[C]. Nassau, Bahamas: IEEE Press, 2004. 3459 – 3466.

[14] Alexandre P. Multiple Model Estimation and Detection for Adaptive Guidance of Hybrid Systems[D]. Montreal: McGill University, 2004.

[15] Nguyen-Due M, Chaib-draa B. Resolution-based policy search for imperfect information differential games[A]. Proc of the IEEE/WIC/ACM International Conf on Intelligent Agent Technology[C]. Hong Kong: IEEE Press, 2006. 326 – 332.

作者简介:

范红旗　男, 1978 年 10 月生于陕西合阳. 讲师, 博士. 现于国防科大 ATR 重点实验室从事自动目标识别、目标跟踪、信号处理系统等方面的研究工作. 近年来获得国家科技进步奖二等奖和省部委级奖励各 1 项.
E-mail: fanhongqi@tsinghua.org.cn

王　胜　男, 1965 年出于吉林长春. 现为国防科大 ATR 重点实验室博士生, 主要从事雷达信号模拟与识别和电路系统设计方面的研究.　E-mail: snakerws2002@yahoo.com.cn

付　强　男, 1962 年 6 月生于湖南长沙. 教授, 博士生导师. 现于国防科大 ATR 实验室从事雷达信号处理、目标识别以及精确制导等方面的研究工作. 近年来获国家科技进步二等奖 2 项, 在国内外发表学术论文 60 余篇.　E-mail: fuqiang1962@vip.sina.com