自 适 应 对 策

胡 恒 章

（哈尔滨工业大学）

摘 要

本文研究在最小方差意义上的动态对策，得到了最大最小对策律．当噪声统计和系统参数未知时，用虚拟噪声法得到了自适应对策律．

一、最大最小方差对策

设研究的动态系统由下述差分方程描写

$$e(k)=x(k)-y(k)=\frac{B}{A}u(k-1)-\frac{C}{A}v(k-1)+\frac{1}{A}w(k), \tag{1}$$

式中 $x(k)$ 追方位置，$u(k)$ 追方控制，$y(k)$ 逃方位置，$v(k)$ 逃方控制，$e(k)$ 追方与逃方的距离 k 采样时间，z^{-1} 一步滞后算子，

$$\begin{aligned}A&=1+a_1z^{-1}+\cdots+a_nz^{-n},\\B&=b_1+b_2z^{-1}+\cdots+b_nz^{-n+1},\\C&=c_1+c_2z^{-1}+\cdots+c_nz^{-n+1},\end{aligned} \tag{2}$$

$w(k)$ 零均值白噪声扰动．

按（1）（2）式有展开式

$$e(k)=-\sum_{i=1}^{n}a_ie(k-i)+\sum_{i=1}^{n}b_iu(k-i)-\sum_{i=1}^{n}c_iv(k-i)+w(k). \tag{3}$$

追方努力使 $e(k)$ 最小，逃方努力使 $e(k)$ 最大[4]，指标函数取为

$$J=E\{[e(k+1)]^2+[Ru(k)]^2-[Sv(k)^2\}, \tag{4}$$

式中 $R=r+r_1z^{-1}+\cdots+r_mz^{-m}$，$S=s+s_1z^{-1}+\cdots+s_mz^{-m}$，

E 表示数学期望．

设追方可测得 $e(k)$、$u(k)$．逃方可测得 $e(k)$、$v(k)$．选择 R，S 以限制 $u^2(k)$ 与 $v^2(k)$ 并保证闭环性能．由于追方不知 $v(k)$，它努力使 J 最小，只能假设 $v(k)$ 使 J 达最大的条件之下，再使 J 最小，故 $u(k)$ 的选择原则是

$$\min_{u(k)}\max_{v(k)} J\ . \tag{6}$$

本文于1984年10月30日收到，1985年2月14日收到修改稿。

同样逃方不知 $u(k)$，它努力使 J 最大，只能假设 $u(k)$ 使 J 达最小的条件下，使 J 最大，故 $v(k)$ 的选择原则是

$$\max_{v(k)}\ \min_{u(k)} J。\tag{7}$$

由（1）式

$$w(k)=Ae(k)-Bu(k-1)+Cv(k-1).\tag{8}$$

设 $\dfrac{1}{A}=1+\dfrac{G}{A}z^{-1}$， $G=g_0+g_1z^{-1}+\cdots+g_{n-1}z^{-n+1}$。 (9

式中 $g_i=-a_{i+1}$。

由（1）、（8）、（9）式有

$$e(k+1)=\frac{B}{A}u(k)-\frac{C}{A}v(k)+\frac{G}{A}w(k)+w(k+1).\tag{10}$$

在 k 时刻预报（$k+1$）时刻之 $\hat{e}(k+1|k)$，只有承认 $w(k+1)$ 为预报误差，故有

$$\hat{e}(k+1|k)=\frac{B}{A}u(k)-\frac{C}{A}v(k)+\frac{G}{A}\omega(k).\tag{11}$$

由（8）、（11）式有

$$\hat{e}(k+1|k)=Ge(k)+Bu(k)-Cv(k),\tag{12}$$

承认预报误差以 $\hat{e}(k+1|k)$ 代替 $e(k+1)$.为求取 $\max\limits_{v(k)} J$，就需

$$\frac{\partial J}{\partial v(t)}=0=-c_1[Ge(k)+Bu(k)-Cv(k)]-Ssv(k),$$

故得

$$v(k)=[Ge(k)+Bu(k)]/(C-sS/c_1).\tag{13}$$

为求取 $\min\limits_{u(k)}\max\limits_{v(k)} J$，需（13）式代入（4）式，并 $J=J_1$，有

$$J_1=E\left\{\left[Ge(k)+Bu(k)-C\frac{Ge(k)+Bu(k)}{C-sS/c_1}\right]^2-\left[S\frac{Ge(k)+Bu(k)}{C-sS/c_1}\right]^2\right\}.\tag{14}$$

再求 $\dfrac{\partial J_1}{\partial u(k)}=0$，得 $u(k)$ 为（6）式解

$$u(k)=GSe(k)/\left(rR\frac{Cc_1-Ss}{sh_1}\cdot\frac{c_1^2-S^2}{c_1^2+s^2}\right).\tag{15}$$

为求（7）式之解，需先求 $\min\limits_{u(k)} J$，同上法得

$$u(k)=[-Ge(k)+Cv(k)]/(B+Rr/b_1). \tag{16}$$

（16）式代入（4）式得 $J=J_2$，再取 $\dfrac{\partial J_2}{\partial v(k)}=0$，得

$$v(k)=GRe(k)\Bigg/\left(sS\frac{Bb_1+Rr}{rc_1}\cdot\frac{r^2+b_1^2}{r^2-b_1^2}+CR\right). \tag{17}$$

（15）、（17）即为最大最小对策律，若 $R=0$、$S=0$ 时有

$$u(k)=-\frac{G}{B}e(k),\ v(k)=\frac{G}{C}e(k). \tag{18}$$

尚需指出，中间结果（13）、（16）式作为最小方差对策律[1]，只有追、逃双方互相熟知对方策略时才能实现，就是说需要完全信息，因欲求得（13）式之 $v(k)$ 需知 $u(k)$，欲求得（16）式之 $u(k)$ 需知 $v(k)$，而这在对抗性对策中是难以实现的。完全的知己知彼是困难的，而（15）、（17）式不要求知己知彼，只要求得到 $e(k)$ 信息，是假设对方都采取了对自己最不利的策略条件下的解，这正是对策的一般解法。

二、自适应对策

当不但对方策略未知，而且动态参数和噪声统计也未知时，需在线估计这些未知量，此时对策需自适应解。

首先求追方的自适应对策律，设

$$\eta(k)=w(k)-\sum_{i=1}^{n}c_iv(k-i),$$
$$h^T(k)=[e(k-1)\cdots e(k-n)u(k-1)\cdots n(k-n)],$$
$$\theta^T=[a_1\cdots a_n,b_1\cdots b_n), \tag{19}$$

式中 T 表示转置，（3）式可改写为

$$e(k)=h^T(k)\theta+\eta(k)\ . \tag{20}$$

称 $\eta(k)$ 为虚拟噪声[7]，它的数学期望和方差都是随时间变化的

$$E[\eta(k)]=p(k),\ E[\eta(k)\eta(i)]=P(k)\delta_{ki}. \tag{21}$$

这相当于以

$$\eta(k)=p_0(k)+p_1(k)\delta(k) \tag{22}$$

逼近，$\delta(k)$ 为平稳白噪声，因对 $p_0(k)$，$p_1(k)$ 未加限制，对非零均值非平稳相关噪声以（22）式可逼近到所要求的精度。假设参数向量 $\theta(k)$ 为广义随机游动。

$$\theta(k+1)=\theta(k)+\varepsilon(k),$$

$$E[\varepsilon(k)]=q(k),\qquad E[\varepsilon(k)\varepsilon(i)]=Q(k)\delta_{ki}.\qquad (2$$

以(23)式为状态方程，以(20)式为测量方程，即可估计参数$\hat{\theta}$．用Kalr
自适应辨识[3][5]

$$\begin{aligned}
&\hat{\theta}(k+1)=\hat{\theta}(k)+K(k+1)\varepsilon(k+1)+\hat{q}(k),\\
&\varepsilon(k+1)=e(k+1)-h(k+1)[\hat{\theta}(k)+\hat{q}(k)]-\hat{p}(k),\\
&K(k+1)=M(k+1)h(k+1)[h^T(k+1)M(k+1)h(k+1)+\hat{P}(k)]^{-1},\\
&M(k+1)=B(k)+\hat{Q}(k),\\
&B(k+1)=[I-K(k+1)h^T(k+1)]M(k+1),\\
&\hat{q}(k+1)=(1-d_k)\hat{q}(k)+d_k[\hat{\theta}(k+1)-\hat{\theta}(k)],\\
&\hat{Q}(k+1)=(1-d_k)\hat{Q}(k)+d_k[K(k+1)\varepsilon^2(k+1)K^T(k+1)+B(k+1)-B(k)],\\
&\hat{p}(k+1)=(1-d_k)\hat{p}(k)+d_k[e(k+1)-h^T(k+1)(\hat{\theta}(k)+\hat{q}(k))],\\
&\hat{P}(k+1)=(1-d_k)\hat{P}(k)+d_k[\varepsilon^2(k+1)-h^T(k+1)M(k+1)h(k+1)].
\end{aligned}\qquad (24$$

用(24)式可估计$\hat{\theta}(k)$，$\hat{q}(k)$，$\hat{Q}(k)$，$\hat{p}(k)$，$\hat{P}(k)$，已知量是$e(k)$，$u(k)$，

$$d_k=\frac{1-b}{1-b^{k+1}},\qquad 0<b<1,\qquad (25$$

式中b称遗忘因子，依参数变化速度选定．此时预报$\hat{e}(k+1|k)$为

$$\hat{e}(k+1|k)=Ge(k)+Bu(k)+\hat{p}(k).\qquad (26$$

(26)式之$\hat{e}(k+1|k)$代入(4)之$e(k+1)$取$\min\limits_{u(k)} J$，可得

$$u(k)=-\frac{b_1}{Rr}\hat{e}(k+1|k)=-\frac{b_1}{Rr}\Big[Ge(k)+Bu(k)+\hat{p}(k+1)\Big].\qquad (27$$

考虑到(27)式右端含有$\frac{b_1^2}{Rr}u(k)$项，移到左端，最后得$u(k)$解，a_i b_i以$\hat{a}_i$
代之，得

$$u(k)=\Big[\sum_{i=1}^{n}\hat{a}_ie(k-i+1)-\sum_{i=2}^{n}\hat{b}_iu(k-i+1)-\hat{p}(k)\Big]\Big/\Big(\hat{b}_1+\frac{Rr}{\hat{b}_1}\Big),\qquad (28)$$

式中$\hat{a}_i$、$\hat{b}_i$为$\hat{\theta}(k)$的分量，即为自适应对策律．从这个解法中可看出，在线估计
$\hat{a}_i$、$\hat{b}_i$同时假设有虚拟噪声，其中包括了$v(k)$、$w(k)$，在线估计它的期望$\hat{p}(k)$，并
在控制律(28)中予以补偿．同时估计$\hat{P}(k)$、$\hat{Q}(k)$、$\hat{q}(k)$，在线辨识参数时可达更高
精度，亦可避免因模型误差使滤波辨识发散，对$v(k)$的求取同上方法，设

$$\eta_1(k)=w(k)+\sum_{i=1}^{n} b_i u(k-i),$$

$$h_1^T(k)=[e(k-1),\cdots e(k-n),-v(k-1),\cdots -v(k-n)],$$

$$\theta_1^T=[a_1\cdots a_n\ c_1\cdots c_n], \tag{29}$$

式中 T 表示转置.

(3)式可改写为

$$e(k)=h_1^T(k)\theta_1+\eta_1(k), \tag{30}$$

$$E[\eta_1(k)]=p_1(k),\quad E[\eta_1(k)\eta_1(i)]=P_1(k)\delta_{ki}. \tag{31}$$

假设 $\theta_1(k)$ 为广义随机游动

$$\theta_1(k+1)=\theta_1(k)+\varepsilon_1(k),$$

$$E[\varepsilon_1(k)]=q_1(k),\quad E[\varepsilon_1(k)\varepsilon_1(i)]=Q_1(k)\delta_{ki}, \tag{32}$$

仍用(24)式估计，只是变量带下标1，预报为

$$\hat{e}(k+1|k)=Ge(k)-Cv(k)+\hat{p}_1(k), \tag{33}$$

(4)式中 $e(k+1)$ 以(33)式代，求 $\max\limits_{v(k)} J$,

$$v(k)=-\frac{c_1}{Ss}\hat{e}(k+1|k)=-\frac{c_1}{Ss}[Ge(k)-Cv(k)+\hat{p}(k)], \tag{34}$$

等式(34)右端含 $\frac{c_1^2}{Ss}v(k)$ 项，移至左端后归并得

$$v(k)=\left[\sum_{i=1}^{n}\hat{a}_i e(k-i+1)+\sum_{i=2}^{n}\hat{c}_i v(k+1-i)-\hat{p}_1(k)\right]\Big/\left(\hat{c}_1+\frac{Ss}{\hat{c}_1}\right). \tag{35}$$

(35)式为逃方的自适应对策律，对 $u(k)$ 的校正包含在 $\hat{p}_1(k)$ 中，$\hat{a}_i$、$\hat{c}_i$ 为 θ_1 的元素。

(28)、(35)式组成自适应对策解，它是以自校正[2]方法求得，而对虚拟噪声统计以在线自适应递推估计求得，其中包含了对方的策略. 因此自适应对策律是通过在线辨识在最小方差的意义上做到知己知彼，与二重控制[4]对应可以称自适应对策为二

重对策，一重对策的含义为最大最小对策，二重对策的含义为虚拟噪声统计的在纟识，在对策的过程中，不断增加对对方的认识和动态参数信息．尚须指出，以上方法算量已降到最小，只有矩阵加减乘，而无矩阵求逆计算，（24）式中〔$h^T(k+1)M($ $1)h(k+1)+\hat{P}(k)]^{-1}$是标量除法，因此对计算速度要求不会过高，这一优点来源于（3）式描述动态过程，而未用状态方程表达式，从而回避了矩阵求逆，并将变量利量降到最小数目，以减少计算时间．例如只有当a_1，c_1，b，并取$R=r$，$S=s$时有

$$\theta^T=[a_1\ \ b_1],\ \theta_1^T=[a_1\ \ c_1],\ h^T(k)=[e(k-1)\ \ u(k-1)],$$

$$h_1^T=[e(k-1)\ \ v(k-1)].$$

$$u(k)=[\hat{a}_1 e(k)-\hat{p}(k)]\Big/\left(\hat{b}_1+\frac{r^2}{\hat{b}_1}\right),$$

$$v(k)=[\hat{a}_1 e(k)-\hat{p}_1(k)]\Big/\left(\hat{c}_1+\frac{s^2}{\hat{c}_1}\right),$$

式中 $\hat{a}_1$，$\hat{p}_1$，$\hat{p}$，$\hat{b}_1$，$\hat{c}_1$由（24）式求得．

参 考 文 献

〔1〕 Astrom K.J.,, Introduction to Stochastic Control, Acad. Press, New York, (1970), 159-209.

〔2〕 Clarke D. W., Gawthrop J. P., Self-tuning Controller, Proc. IEE, 122, (1975), 929.

〔3〕 Tamura H., Uenno N., Suboptimal Recursive Filter for the Distributed-Delay Model, 计测自动制御学会文集(日), **11**, 3, (1975), 296-302.

〔4〕 Ho Y. C., Decision, Control and Games, Harvard University, (1979), 112-136.

〔5〕 Sage A. P., Husa G. W., Adaptive Filtering with Unknown Prior Statistics, JACC, (1969).

〔6〕 Astrom K. J., Wittenmark B., On Self-tuning Regulators, Automatic, **9**, 2, (1973), 185-199.

〔7〕 Yoshimura T., Soeda T., A Technique for Compensating the Filter Performance by A Fictitious Noise, Tran. ASME, **100**, 2, (1978).

ADAPTIVE GAMES

Hu Hengzhang
(Harbin Institute of Technology)

Abstract

This paper discussed the dynamic games in the sense of minimal variance. The minmax game law was obtained. When the noise statistics and the system parameters are unknown, the adaptive game law can be obtained by a fictitious noise.

《多变量反馈系统的稳定性和鲁棒性》简介

郑应平
（中国科学院自动化研究所，北京）

美国MIT出版社出版的《信号处理、最优化与控制丛书》第三卷，沙弗诺夫著的《多变量反馈系统的稳定性和鲁棒性》一书的中译本，最近由我国科学出版社出版。

尽管原书出版至今已有六、七年，控制界对鲁棒性问题的关注有增无已。在1986年12月召开的第25届IEEE控制与决策会议（CDC）上，直接以鲁棒性为题的分组就有七个，散在其他分组的论文也有许多讨论这一问题，形成了该次会议的一大热门。这种情形表明了鲁棒性问题无论在理论上还是在实际上都有重要的意义，它是现代控制理论的任何实际应用都必须首先回答的问题。

作为研究一大类鲁棒性问题的工具，沙弗诺夫首先建立了一个用函数空间来描述系统，并用该函数空间中的拓扑划分来表达的稳定性判别准则（第二章）。由该准则可容易地推出许多已知的稳定性结果，这种十分普遍和有力的理论结果无论对工程师还是应用数学家都有很大的吸引力。

利用这个理论工具，作者分别对连续时间（第三章）和离散时间（第四章）情形着重研究了线性二次状态反馈及LQG这两种最常用的设计方法的鲁棒性，即当所用线性模型具有较大误差时为保证系统稳定、系统特性的偏离所应满足的条件。这种条件是对十分普遍的非线性算子提出的一些要求，从而必可适用于许多不同类型的问题，并可推广到非线性观测器的跟踪性质、广义Kalman滤波器的发散和增益调节问题，以及在保证稳定性的意义下对估计和控制分别进行设计（分离性原理）的问题。基于这些结果，书中给出了实用的线性二次状态反馈和广义Kalman滤波器的设计步骤。

可以认为，本书阐述了鲁棒性研究的一个重要方面，具有较大的参考价值。