第 30 卷 第 1 期
2004 年 1 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 30, No. 1
Jan., 2004

Information Rates and H_∞ Entropy in Multivariable LTI Control Systems[1)]

ZHANG Hui SUN You-Xian

(*National Laboratory of Industrial Control Technology, Institute of Modern Control Engineering, Department of Control Science & Engineering, Zhejiang University, Hangzhou* 310027)

(E-mails: zhanghui_iipc@zju.edu.cn; yxsun@iipc.zju.edu.cn)

Abstract Linear multivariable control systems disturbed by Gaussian stationary random sequences are investigated from the viewpoint of Shannon information theory. Relations between entropy rate, mutual information rate of system variables and the H_∞ entropy of closed-loop transfer functions are derived using frequency domain calculation formulae of information rates. These relations interpret the minimum entropy H_∞ control method and give time domain computing methods for information rates in terms of information theory. Our results introduce a new kind of instrument for further study of control systems in the framework of information theory.

Key words Entropy rate, mutual information rate, H_∞ entropy, linear multivariable control systems, Gaussian stationary sequence

1 Introduction

The most important objective of feedback control system is to find an admissible controller to reduce the uncertainty caused by perturbation or noise, and get acceptable performance while keeping system stable. From the viewpoint of information theory, any control system can be considered as an information or uncertainty transmission channel. Variables in systems whose uncertainty is described by extraneous random disturbance inputs are stochastic processes and can be characterized by probability measures. This makes possible the application of Shannon information theory, which is also based on probability theory, to the study of control systems [1~3]. Information theoretic method usually adopts the measures of entropy or mutual information as performance functions of control systems, such as [1,3]. The common characteristic of these functions is that they describe the uncertainty or information of a process at a certain time and thus reflect the transient performance.

In the field of robust control, as a suboptimal design method, the so-called minimum entropy H_∞ control theory has attracted a great deal of attention[4,5]. This method includes an unintuitive measure—the H_∞ entropy, which is different from the Shannon entropy—as the performance index. It was pointed out that H_∞ entropy is in fact an index measuring the tradeoff between the H_∞ optimality and the H_2 optimality[6]. However, little has been studied on the physical meaning of this measure in control systems.

With the time average property, the measures of information rates (such as entropy rate and mutual information rate) in Shannon information theory describe the characteristic of a stochastic process during its overall time evolution. In order to seek after the route of analyzing and designing control systems using information rates as performance functions, the present paper studies connections between information rates and the H_∞ entropy by investigating information or uncertain transmission in multivariable LTI systems disturbed by (Gaussian) stationary processes. Some necessary concepts and lemmas will be stated in Section 2. In Section 3, the relations between entropy rate, mutual information rate and

1) Supported by National Natural Science Foundation of P. R. China (60084001)
Received September 2, 2002; in revised form August 11, 2003
收稿日期 2002-09-02; 收修改稿件日期 2003-08-11

the H_∞ entropy will be formulated. These relations lead to the time domain computation method for information rates, and interpret the H_∞ entropy in terms of information theory. Section 4 is the conclusion.

2 Concepts and lemmas

Let $X_1^n=\{x_1,x_2,\cdots,x_n\}$, $Y_1^n=\{y_1,y_2,\cdots,y_n\}$ be sequences of discrete time stochastic processes x and y, respectively. The entropy rate[7] of x is $\overline{H}(x)=\lim_{n\to\infty}\frac{1}{n}H(X_1^n)$, describing the per unit time information or uncertainty of x, where $H(X_1^n)$ is the entropy of X_1^n. The mutual information rate[8] of x and y is $\overline{I}(x;y)=\lim_{n\to\infty}\frac{1}{n}I(X_1^n;Y_1^n)$, which measures the average transmitted information between processes x and y, where $I(X_1^n;Y_1^n)$ is the mutual information of X_1^n and Y_1^n.

Lemma 1. Let $G(z)\in RH_\infty$ be the $m\times m$ transfer function matrix of a discrete-time MIMO LTI system, where RH_∞ denotes the set of all stable and proper transfer function matrices, the stationary stochastic input $\boldsymbol{x}(k)\in R^m$ $(k=0,1,2,\cdots)$ has positive spectral density $\Phi_x(\omega)$. Then the entropy rate of system output $\boldsymbol{y}(k)\in R^m$ is

$$\overline{H}(\boldsymbol{y}) = \overline{H}(\boldsymbol{x}) + \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln |\det G(e^{i\omega})| \,d\omega \tag{1}$$

where $\overline{H}(\boldsymbol{x})$ is the entropy rate of input. When $\boldsymbol{x}$ is Gaussian,

$$\overline{H}(\boldsymbol{x}) = \frac{1}{2}\ln (2\pi e)^m + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln \det\Phi_x(\omega)\,d\omega \tag{2}$$

Proof. See Appendix.

Lemma 2[8]. Let two joint Gaussian stationary processes $\boldsymbol{x}(k)\in R^n$, $\boldsymbol{y}(k)\in R^m$, $k=1,2,\cdots$, have spectral densities Φ_x and Φ_y, $\boldsymbol{\eta}(k)=[\boldsymbol{x}^T(k)\quad \boldsymbol{y}^T(k)]^T\in R^{n+m}$ have spectral density Φ_η. Then the mutual information rate of $x(k)$, $y(k)$ is

$$\overline{I}(\boldsymbol{x};\ \boldsymbol{y}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln \frac{\det\Phi_x(\omega)\det\Phi_y(\omega)}{\det\Phi_\eta(\omega)}\,d\omega \tag{3}$$

Let any transfer function $G(z)$ satisfy $\| G(z) \|_\infty=\sup_\omega\bar{\sigma}[G(e^{i\omega})]<\lambda$, where $\bar{\sigma}$ denotes the maximum singular value; then the entropy of $G(z)$ in H_∞ control is defined as[4,5]

$$H(G,\ \lambda) = \frac{-\lambda^2}{4\pi}\int_{-\pi}^{\pi} \ln \det[I-\lambda^{-2}G^*(e^{i\omega})G(e^{i\omega})]\,d\omega \tag{4}$$

where $G^*(e^{i\omega})=G^T(e^{-i\omega})$. To distinguish it from Shannon entropy, we refer to function (4) as the H_∞ entropy.

3 Relations between information rates and the H_∞ entropy

The multivariable discrete time LTI system under study is shown in Fig. 1 where $\boldsymbol{r},\boldsymbol{d},\boldsymbol{y}\in R^n$ are the reference input, disturbance and output, respectively. $\boldsymbol{r}(k)$, $\boldsymbol{d}(k)$, $k=0,1,2,\cdots$ are mutually independent zero-mean Gaussian processes with spectral densities $\Phi_r(\omega)$ and $\Phi_d(\omega)$. $C(z)$ and $P(z)$ are $n\times n$ proper transfer function matrices. The system is well-posed and closed-loop stable.

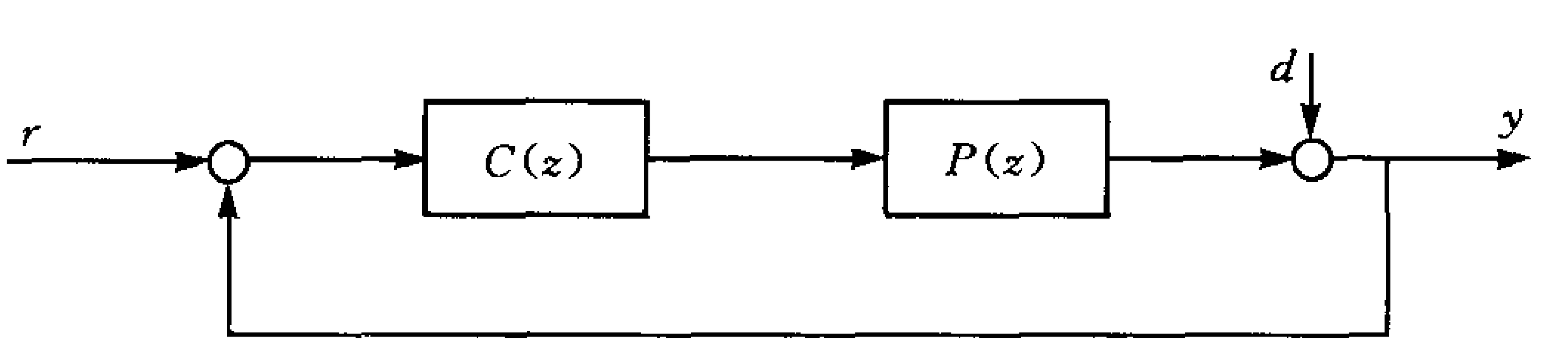

Fig. 1 LTI system with random disturbance

Let $L(z)=P(z)C(z)$ denote the system open-loop transfer function. The output closed-loop transfer functions are $S(z)=[I+L(z)]^{-1}$, $T(z)=L(z)[I+L(z)]^{-1}$, respectively. Then $y(z)=T(z)r(z)+S(z)d(z)$, where $\boldsymbol{y}(z)$, $\boldsymbol{r}(z)$ and $\boldsymbol{d}(z)$ are the z-transformations of $\boldsymbol{y}$, $\boldsymbol{r}$ and $\boldsymbol{d}$, respectively. Let $\boldsymbol{y}_r(z)=T(z)\boldsymbol{r}(z)$, $\boldsymbol{y}_d(z)=S(z)\boldsymbol{d}(z)$, and the spectral densities of $\boldsymbol{y}_r$, $\boldsymbol{y}_d$ and $\boldsymbol{y}$ be $\Phi_1(\omega)$, $\Phi_2(\omega)$ and $\Phi_y(\omega)$, respectively. Then, $\Phi_1(\omega)=T(e^{i\omega})\Phi_r(\omega)T^*(e^{i\omega})$, $\Phi_2(\omega)=S(e^{i\omega})\Phi_d(\omega)S^*(e^{i\omega})$, $\Phi_y(\omega)=\Phi_1(\omega)+\Phi_2(\omega)$. The mutual spectral densities of pairs $(\boldsymbol{r},\boldsymbol{y})$ and $(\boldsymbol{d},\boldsymbol{y})$ are $\Phi_{ry}(\omega)=\Phi_r(\omega)T^*(e^{i\omega})$, $\Phi_{dy}(\omega)=\Phi_d(\omega)S^*(e^{i\omega})$, respectively.

3.1 Entropy rate and mutual information rate

The mutual information rate of disturbance $\boldsymbol{d}$ and output $\boldsymbol{y}$ in Fig. 1 is $\bar{I}(\boldsymbol{d};\boldsymbol{y})$, which measures the information of one variable contained in another. For the purpose of disturbance rejection, the system should be designed to make $\bar{I}(\boldsymbol{d};\boldsymbol{y})$ as small as possible.

Theorem 1. The mutual information rate of disturbance $\boldsymbol{d}$ and output $\boldsymbol{y}$ in Fig. 1 is

$$\bar{I}(\boldsymbol{d};\boldsymbol{y}) = \bar{H}(\boldsymbol{y}) - \bar{H}(\boldsymbol{y}_r) \tag{5}$$

where $\bar{H}(\boldsymbol{y})$, $\bar{H}(\boldsymbol{y}_r)$ are entropy rates of $\boldsymbol{y}$ and $\boldsymbol{y}_r$, respectively.

Proof. Let $\boldsymbol{\xi}(k)=[\boldsymbol{d}^{\mathrm{T}}(k)\quad \boldsymbol{y}^{\mathrm{T}}(k)]^{\mathrm{T}}$. The spectrum of $\boldsymbol{\xi}(k)$ is $\Phi_\xi=\begin{bmatrix}\Phi_d(\omega) & \Phi_{dy}(\omega)\\ \Phi_{dy}^*(\omega) & \Phi_y(\omega)\end{bmatrix}$, then $\det\Phi_\xi(\omega)=\det\Phi_d(\omega)\det\Phi_1(\omega)$. From Lemma 2,

$$\bar{I}(\boldsymbol{d};\ \boldsymbol{y}) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\frac{\det\Phi_d(\omega)\det\Phi_y(\omega)}{\det\Phi_\xi(\omega)}\mathrm{d}\omega = \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\frac{\det\Phi_y(\omega)}{\det\Phi_1(\omega)}\mathrm{d}\omega \tag{6}$$

From Lemma 1,

$$\bar{I}(\boldsymbol{d};\boldsymbol{y}) = \frac{1}{4\pi}\int_{-\pi}^{\pi}[\ln\det\Phi_y(\omega)-\ln\det\Phi_1(\omega)]\mathrm{d}\omega = \bar{H}(\boldsymbol{y})-\bar{H}(\boldsymbol{y}_r)$$ □

Note 1. The output y consists of two parts: The signal $\boldsymbol{y}_r$ transmitted from reference, and $\boldsymbol{y}_d$ transmitted from disturbance. $\bar{H}(\boldsymbol{y})$ measures the total uncertainty of output, while $\bar{H}(\boldsymbol{y}_r)$ is the information concerning the reference obtained by output. The aim of system design is to make $\boldsymbol{y}$ track reference signal $\boldsymbol{r}$ and reject disturbance $\boldsymbol{d}$. For $\bar{H}(\boldsymbol{y})=\bar{I}(\boldsymbol{d};\boldsymbol{y})+\bar{H}(\boldsymbol{y}_r)$, it will possibly make the output 'lose' the information of reference if we adopt $\bar{H}(\boldsymbol{y})$ as the minimizing function in this case. Under this consideration, a rational selection is to make $\bar{H}(\boldsymbol{y})$ constrained by a certain bound (Based on the relation between entropy rate and variance[7], this bound corresponds to a variance bound of output.), and make $\bar{H}(\boldsymbol{y}_r)$ achieve its maximum. From (5), when $\bar{H}(\boldsymbol{y})$ is bounded, the smaller $\bar{I}(\boldsymbol{d};\boldsymbol{y})$ is, the larger $\bar{H}(\boldsymbol{y}_r)$ is. Hence, Theorem 1 demonstrates that, as a measure of disturbance rejection, $\bar{I}(\boldsymbol{d};\boldsymbol{y})$ reflects the tracking performance in the meantime.

3.2 Mutual information rate and H_∞ entropy

$\boldsymbol{y}$ and $\boldsymbol{d}$ are Gaussian stationary processes for the system is closed-loop stable. From the spectral factorization theorem of multidimensional stationary process[9], there exist $n\times n$ rational matrices $F_y(z)$ and $F_d(z)$, with zeros and poles of $\det F_y(z)$ and $\det F_d(z)$ all inside the unit circle, so that $\Phi_y(\omega)=F_y(e^{i\omega})F_y^*(e^{i\omega})$, $\Phi_d(\omega)=F_d(e^{i\omega})F_d^*(e^{i\omega})$.

Let $\boldsymbol{y}$ and $\boldsymbol{d}$ be transformed by nonsingular functions $F_y^{-1}(z)$ and $F_d^{-1}(z)$, respectively, *i.e.*, $\boldsymbol{y}'(z)=F_y^{-1}(z)\boldsymbol{y}(z)$, $\boldsymbol{d}'(z)=F_d^{-1}(z)\boldsymbol{d}(z)$. Then the spectrum of $\boldsymbol{\xi}'(z)=[\boldsymbol{d}'^{\mathrm{T}}(z)\quad \boldsymbol{y}'^{\mathrm{T}}(z)]^{\mathrm{T}}$ is $\Phi_{\xi'}(\omega)=\begin{bmatrix} I & M(e^{i\omega})\\ M^*(e^{i\omega}) & I\end{bmatrix}$, where $M(z)\in RH_\infty$, and

$$M(e^{i\omega}) = F_d^{-1}(e^{i\omega})\Phi_{dy}(\omega)[F_y^*(e^{i\omega})]^{-1} = F_d^*(e^{i\omega})S^*(e^{i\omega})[F_y^*(e^{i\omega})]^{-1} \tag{7}$$

From Lemma 2,

$$\bar{I}(\boldsymbol{d}';\boldsymbol{y}') = \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\frac{\det\Phi_{d'}(\omega)\Phi_{y'}(\omega)}{\det\Phi_{\xi'}(\omega)}\mathrm{d}\omega = \frac{-1}{4\pi}\int_{-\pi}^{\pi}\ln\det[I-M^*(e^{i\omega})M(e^{i\omega})]\mathrm{d}\omega \tag{8}$$

It is known that the mutual information is invariant under nonsingular transformation[7,10], thus

$$\bar{I}(\boldsymbol{d};\boldsymbol{y}) = \frac{-1}{4\pi}\int_{-\pi}^{\pi} \ln \det[I - M^*(e^{i\omega})M(e^{i\omega})]\mathrm{d}\omega \tag{9}$$

Because $\Phi_y(\omega) = \Phi_1(\omega) + \Phi_2(\omega) \geqslant \Phi_2(\omega)$, $\| M(z) \|_\infty \leqslant 1$. If $\| M(z) \|_\infty = 1$, it is known from equation (9) that $\bar{I}(\boldsymbol{d};\boldsymbol{y}) = +\infty$. From Theorem 1, $\bar{I}(\boldsymbol{d};\boldsymbol{y}) = \bar{H}(\boldsymbol{y}) - \bar{H}(\boldsymbol{y}_r)$. The entropy rate of Gaussian stationary process always exists, i. e., both $\bar{H}(\boldsymbol{y})$ and $\bar{H}(\boldsymbol{y}_r)$ are smaller than $+\infty$, $\bar{I}(\boldsymbol{d};\boldsymbol{y}) < +\infty$. Hence

$$\| M(z) \|_\infty < 1 \tag{10}$$

Note that the right side of (9) is exactly the H_∞ entropy of transfer matrix M under condition (10). Denote it as

$$H(M,1) = \frac{-1}{4\pi}\int_{-\pi}^{\pi} \ln \det[I - M^*(e^{i\omega})M(e^{i\omega})]\mathrm{d}\omega \tag{11}$$

Let $C^- = \{ c \in C, |c| < 1 \}$. For $A \in R^{n\times n}$, let $sp(A)$ denote the set of eigenvalues of A. Suppose the transfer matrix $M(z)$ has a state space realization, *i. e.*,

$$M(z) = C(zI - A)^{-1}B + D \text{ , } sp(A) \subset C^- \tag{12}$$

From the results of Stoorvogel and Van Schuppen[2], when $\| M(z) \|_\infty < 1$, the equation

$$\begin{aligned} &Q = A^TQA + C^TC + (A^TQB + C^TD)E^{-1}(B^TQA + D^TC), \; Q = Q^T \geqslant 0, \\ &E = I - B^TQB - D^TD > 0, \; sp(A + BE^{-1}(B^TQA + D^TC)) \subset C^- \end{aligned} \tag{13}$$

has a unique solution $Q \in R^{n\times n}$, and

$$H(M,1) = -\frac{1}{2}\ln \det(I - B^TQB - D^TD) \tag{14}$$

Based on the above analysis, the following conclusion is derived.

Theorem 2. For the system shown in Fig. 1, the mutual information rate of disturbance $\boldsymbol{d}$ and output $\boldsymbol{y}$ is equivalent to the H_∞ entropy of transfer matrix M described by (7),

$$\bar{I}(\boldsymbol{d};\ \boldsymbol{y}) = H(M,1) \tag{15}$$

Suppose M has a state space realization described by (12). Then

$$\bar{I}(\boldsymbol{d};\ \boldsymbol{y}) = -\frac{1}{2}\ln \det(I - B^TQB - D^TD) \tag{16}$$

Note 2. It is known from (6) that $\bar{I}(\boldsymbol{d};\boldsymbol{y})$ is defined by system closed-loop transfer functions and inputs $\boldsymbol{r}$ and $\boldsymbol{d}$. From (7), system M characterized by $H(M,1)$ is also defined by closed-loop transfer functions and inputs of the original system. Furthermore, (15) reflects the equivalence between these two performance functions, while (16) gives a time domain computation method of mutual information rate.

Note 3. $\bar{H}(\boldsymbol{y})$ is the total information (uncertainty) of $\boldsymbol{y}$, while $\bar{H}(\boldsymbol{y}_r)$ describes the information about $\boldsymbol{r}$ obtained by system output. Hence, (5) and (15) demonstrate that the H_∞ entropy reflects the average uncertainty caused by disturbance in the sense of information theory.

Note 4. The minimum entropy H_∞ control method is to find an admissible controller to minimize the H_∞ entropy of system when the H_∞ norm of system closed-loop transfer function is bounded. It can be seen from Theorem 2 that minimizing the H_∞ entropy of system M ($\| M(z) \|_\infty < 1$) is equivalent to making the 'pollution' on the system caused by disturbance be minimal, i. e., it is equivalent to minimizing $\bar{I}(\boldsymbol{d};\boldsymbol{y})$ when $\bar{H}(\boldsymbol{y})$ is bounded.

The relation between mutual information rate and H_∞ entropy in parameter identification problem was studied by Stoorvogel and Schuppen[2]. [2] focused on the mutual information rate between estimation error and a canonical reference signal, while we focus on the mutual information rate between variables of control system itself.

3.3 Entropy rate and H_∞ entropy

Suppose $\boldsymbol{r}$ has spectral factorization $\Phi_r(\omega) = F_r(e^{i\omega})F_r^*(e^{i\omega})$. Let

$$N(e^{i\omega}) = F_r^*(e^{i\omega})T^*(e^{i\omega})[F_y^*(e^{i\omega})]^{-1} \tag{17}$$

Then $N(z)\in RH_\infty$. Because

$$\ln \det\Phi_y(\omega)[\det\Phi_1(\omega)]^{-1}=-\ln \det\{F_y^{-1}(e^{i\omega})T(e^{i\omega})\Phi_r(\omega)T^*(e^{i\omega})[F_y^*(e^{i\omega})]^{-1}\}$$

from (6), it follows that

$$\bar{I}(\boldsymbol{d};\boldsymbol{y})=H(M,1)=\frac{-1}{4\pi}\int_{-\pi}^{\pi}\ln\det[N^*(e^{i\omega})N(e^{i\omega})]d\omega=\frac{-1}{4\pi}\int_{-\pi}^{\pi}\ln|\det N(e^{i\omega})|^2 d\omega$$

Suppose system $N(z)$ is driven by a standard Gaussian white noise (with spectrum I), and the output is $\boldsymbol{z}(k)$. Then from Lemma 1,

$$\bar{H}(\boldsymbol{z})=\frac{1}{2}\ln(2\pi e)^n+\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln|\det N(e^{i\omega})|^2 d\omega$$

Hence,

$$H(M,1)=\bar{I}(d;y)=\frac{1}{2}\ln(2\pi e)^n-\bar{H}(\boldsymbol{z}) \tag{18}$$

Suppose, when driven by a standard Gaussian white noise, the output of system $M(z)$ is $w(k)$. From (7) and (17) we get

$$M^*(z)M(z)+N^*(z)N(z)=I,\quad \|N(z)\|_\infty<1 \tag{19}$$

Then, it can also be concluded from (19) and Lemma 1 that

$$H(N,1)=\frac{1}{2}\ln(2\pi e)^n-\bar{H}(\boldsymbol{w}) \tag{20}$$

Note 5. It can be seen from (17) and Lemma 1 that $\bar{H}(\boldsymbol{z})$ is an alternative measure of the information difference between $\boldsymbol{y}$ and $\boldsymbol{y}_r$. Similarly, (7) and Lemma 1 demonstrate that $\bar{H}(\boldsymbol{w})$ reflects the information difference between $\boldsymbol{y}$ and $\boldsymbol{y}_d$. Hence, (18) and (20) also give information theoretic interpretations for the H_∞ entropy.

For general consideration, suppose an arbitrary $n\times n$ transfer function matrix $U(z)\in RH_\infty$, $\|U(z)\|_\infty<\gamma$. Then there is a spectral factorization $V(z)\in RH_\infty$, $\|V(z)\|_\infty<1$, so that[11]

$$I-\gamma^{-2}U^*(z)U(z)=V^*(z)V(z) \tag{21}$$

The following conclusion can be derived directly from Lemma 1.

Theorem 3. Suppose, when driven by a standard Gaussian white noise (with spectrum I), the outputs of systems $U(z)$ and $V(z)$ are $\boldsymbol{\mu}(k)$and $\boldsymbol{v}(k)$, respectively. Then the following relations between systems $U(z)$ and $V(z)$ hold:

$$H(U,\gamma)=\gamma^2\left[\frac{1}{2}\ln(2\pi e)^n-\bar{H}(\boldsymbol{v})\right] \tag{22}$$

$$H(V,1)=\frac{1}{2}\ln(2\pi e\gamma)^n-\bar{H}(\boldsymbol{\mu}) \tag{23}$$

where $H(U,\gamma)$ and $H(V,1)$ are the H_∞ entropies of systems $U(z)$ and $V(z)$, respectively, $\bar{H}(\boldsymbol{\mu})$ and $\bar{H}(\boldsymbol{v})$ are the entropy rates of $\boldsymbol{\mu}(k)$and $\boldsymbol{v}(k)$, respectively.

Note 6. Theorem 3 formulates relations between entropy rate and the H_∞ entropy, and a time domain computation method for entropy rate along with (14). The relations described by (18) and (20) are special cases of Theorem 3.

Note 7. Because Theorem 3 is derived based on Lemma 1, it can be extended to the case of arbitrary stationary inputs.

4 Conclusion

Information rates (such as entropy rate and mutual information rate) measure the average uncertainty or information of system variables, and thus can be considered as performance functions of control system during its overall time evolution. However, the limitation of computation in frequency domain prevents us from getting more applicable results for system analysis and design by using this theoretic method. The present results demonstrate the informational description properties of the H_∞ entropy, and make it possible to discuss the entropy rate and mutual information rate of system variables in state space, and hence provide useful instruments for further research.

References

1 Weidemann H L. Entropy analysis of feedback control systems. *In Advances in Control Systems*, 1969, **7**(2): 225～255

2 Stoorvogel A A, Van Schuppen J H. System identification with information theoretic criteria. In: Bittanti S, Picc G, Identification, Adaptation, Learning. Berlin: Springer, 1996. 289～338

3 Wang H. Minimum entropy control of non-Gaussian dynamic stochastic systems. *IEEE Transactions on Automatic Control*, 2002, **47**(2): 398～403

4 Mustafa D, Glover K. Minimum Entropy H_∞ Control, Lecture Notes in Control and Information Sciences. 146. Berlin, Heidelberg: Springer-Verlag, 1990

5 Peters M A, Iglesias P A. The relationship between minimum entropy control and risk-sensitive control for time-varying systems. *IEEE Transactions on Automatic Control*, 1999, **44**(5): 1065～1069

6 Zhou K. Essential of Robust Control. New Jersey: Prentice-Hall, 1998. 286

7 Cover T M, Thomas J A. Elements of Information Theory. Chichester: Wiley & Sons, Inc., 1991

8 Pinsker M S. Information and Information Stability of Random Variables and Processes. San Francisco: Holden-Day, Inc., 1964

9 Ljung L. System Identification: Theory for the User. 2nd edition. New Jersey: Prentice-Hall, 1999

10 Gelfand I M, Yaglom A M. Calculation of the amount of information about a random function contained in another such function. *American Mathematical Society Translations*, 1959, **2**(12): 199～246

11 Arveson W. Interpolation in nest algebras. *Journal of Functional Analysis*, 1975, **20**(3): 208～233

12 Papoulis A, Probability, Random Variables, and Stochastic Processes, 3rd edition. New York: McGraw-Hill, Inc., 1991

13 Rosenblatt M. Asymptotic distribution of eigenvalues of block Toeplitz matrices. *Bulletin of American Mathematical Society*, 1960, **66**(2): 320～321

ZHANG Hui Received his bachelor and master degrees from Zhejiang University in 1990 and 1995, respectively. Now, he is a Ph. D. candidate in the Department of Control Science and Engineering at Zhejiang University. His research interests include stochastic systems, robust control, and information theoretic methods for control systems.

SUN You-Xian Graduated from Zhejiang University. Now, he is a professor in the Department of Control Science and Engineering, the Chairman of National Centre of Industrial Automation at Zhejiang University, and the academician of the Chinese Academy of Engineering. His research interests include robust control, fault tolerance control, nonlinear systems, and industrial optimization and control.

Appendix

Proof of Lemma 1.

This lemma had been proved in the case of $m=1$ in [12]. However, so far as the authors know, it has not been proved for multivariable systems. We give the proof here.

Firstly, we suppose x is a Gaussian process. Let $x^n=\{x_1, \cdots, x_n\}$ be a sequence of x. Define a block Toeplitz matrix, $T_{x^n}=[Q(j-k)]_{j,k=-n,\cdots,-1,0,1,\cdots,n}$, where $Q(j-k)\in R^{m\times m}$ is the covariance matrix of x. Then the entropy of x^n is

$$H(x^n) = \frac{1}{2}\ln(2\pi e)^{mn} + \frac{1}{2}\ln\det T_{x^n}$$

From the Szegö theorem of Toeplitz matrix[13],

$$\lim_{n\to\infty}[\det T_{x^n}]^{\frac{1}{n}} = \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\det\Phi_x(\omega)d\omega$$

Then the entropy rate of x is

$$\overline{H}(x) = \frac{1}{2}\ln(2\pi e)^m + \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\det\Phi_x(\omega)d\omega$$

For the same reason, the entropy rate of y is

$$\overline{H}(y) = \frac{1}{2}\ln(2\pi e)^m + \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\det\Phi_y(\omega)d\omega$$

Because $\Phi_y(\omega)=G(e^{i\omega})\Phi_x(\omega)G^*(e^{i\omega})$,

$$\overline{H}(y) = \overline{H}(x) + \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|\det G(e^{i\omega})|d\omega$$

Secondly, let the input be an arbitrary stationary process. Suppose the sequence of output $\{y_1, \cdots, y_n\}$ is a linear transformation of the sequence of input $\{x_1, \cdots, x_n\}$. Let $J(x_1, \cdots, x_n)$ be the Jacobian matrix of this transformation. For $G(z)$ is linear, $J(x_1, \cdots, x_n)$ has a unique inverse, $\Delta = \det J(x_1, \cdots, x_n) \neq 0$, and Δ does not depend on $x_i (i=1,2,\cdots,n)$. Then,

$$H(y_1, \cdots, y_n) = H(x_1, \cdots, x_n) + K_0 \tag{A1}$$

where $K_0 = \ln|\Delta|$ is a constant that depends only on the coefficients of the transformation. When the transformation matrix is of infinite order, the process y depends linearly on x as

$$y_i = \sum_{k=0}^{\infty} l_k x_{i-k}, \quad i = -\infty, \cdots, \infty$$

Extending equation (A1) to the case of infinitely many variables, we conclude that

$$\overline{H}(\boldsymbol{y}) = \overline{H}(\boldsymbol{x}) + K$$

where again K is a constant that depends only on the parameters of system $G(z)$. As we have seen, when x is Gaussian, K equals the integral in (1). And since K is independent of x, it must equal that integral for arbitrary x. □

线性多变量控制系统中的信息率和 H_∞ 熵

章 辉 孙优贤

(浙江大学信息学院控制科学与工程系工业控制国家重点实验室 杭州 310027)

(E-mails: zhanghui_iipc@zju.edu.cn; yxsun@iipc.zju.edu.cn)

摘 要 运用信息论的概念和方法考察受随机干扰的线性多变量控制系统,研究了系统变量的熵率、互信息率和传递函数的 H_∞ 熵三者之间的关系,讨论了小熵 H_∞ 控制方法的信息论意义.同时得出了系统熵率和互信息率的时域计算方法,为进一步利用信息率作为性能函数研究控制系统探索了新的途径.

关键词 熵率,互信息率,H_∞ 熵,线性多变量系统,高斯平稳随机序列

中图分类号 TP13