

基于信息熵的非线性动态系统的宏观分析*

田玉楚

(浙江大学工业控制技术研究所, 工业控制技术国家重点实验室·杭州, 310027)

摘要: 本文针对一类 Itô 型随机非线性动态系统, 分别基于微观 FPK 方程和宏观信息熵方法推断系统特性, 进而分析系统的熵结构. 结果表明, 宏观方法不仅能获得与微观方法相似的结果, 而且能以统一简洁的方式处理各类系统. 通过对一类非线性动态系统的应用进一步说明宏观信息熵分析方法.

关键词: 非线性系统; 随机系统; 系统分析; 信息; 熵; 宏观方法

1 引言

在各种自然科学方法论中, 微观分析方法目前仍占主导地位. 它研究问题的起点是建立数学模型, 然后致力于模型分析和求解. 同时, 微观方法还以还原性为基础, 亦即通过系统分解, 由局部推断整体. 尽管微观方法已取得极大成功, 但它在处理越来越多的复杂系统^[1]时却遇到了巨大障碍: 难以建立数学模型, 还原性也不一定能得到保证. 因此, 除根据特定问题将微观方法向纵深发展外, 急需发展能避开上述困难并从普适和统一观点处理一般复杂系统的理论方法. 整体思想和宏观综合方法特别是信息熵方法有望为这些理论方法提供一般的框架^[1].

信息熵描述了系统的整体不定性. 本文考虑一类 Itô 型随机非线性动态系统:

$$dx(t) = f[x(t)]dt + dw(t). \quad (1)$$

其中 $x \in \mathbb{R}$, $f: \mathbb{R} \rightarrow \mathbb{R}$, w 为标准 Wiener 过程, 满足: 1) $E\{w(t)\} = 0$; 2) $E\{[w(t+s) - w(t)]^2\} = 2Qt$, $Q > 0$; 3) $E\{[w(t_2) - w(t_1)][w(s_2) - w(s_1)]\} = 0$, $t_1 < t_2 \leq s_1 < s_2$. 若该系统有解, 则状态 $x(t)$ 是具有良好 Markov 特性的扩散过程, 其状态转移 PDF (probability density function) $p(t, x|t_0, x_0)$ (简记为 P_t) 满足下列 FPK 方程^[2,3]:

$$\frac{\partial p_t}{\partial t} = -\frac{\partial [f(x)p_t]}{\partial x} + Q \frac{\partial^2 p_t}{\partial x^2}. \quad (2)$$

为描述系统从初态开始转移的最终效果, 考察上述 FPK 方程的稳态解 p_∞ . 在能量有限及能量不能突变的边界条件 $\lim_{x \rightarrow \pm\infty} p_t = 0$ 和 $\lim_{x \rightarrow \pm\infty} \frac{\partial p_t}{\partial x} = 0$ 下, 可导出^[3]:

$$p_\infty = C \exp\left[\frac{1}{Q} \int_{x_0}^x f(\theta) d\theta\right]. \quad (3)$$

其中 C 为归一化常数, 由 $\int_{-\infty}^{+\infty} p_\infty dx = 1$ 决定. 进一步, 若 $x(t)$ 具有初始 PDF $p[x(t_0)] = \delta(x - x_0)$, 则可导出 $x(t)$ 的稳态 PDF $p(x)$ 如下^[3]:

$$p(x) = \int_{-\infty}^{+\infty} p_\infty p[x(t_0)] dx(t_0) = p_\infty. \quad (4)$$

据此便能得到系统的稳态 Shannon 信息熵 $H(x) = - \int_{-\infty}^{+\infty} p(x) \ln p(x) dx$ 为

$$H(x) = -C \int_{-\infty}^{+\infty} \left[\ln C + \frac{1}{Q} \int_{x_0}^x f(\theta) d\theta \right] \exp\left[\frac{1}{Q} \int_{x_0}^x f(\theta) d\theta\right] dx. \quad (5)$$

* 国家博士后科学基金资助项目.

本文于 1995 年 5 月 12 日收到, 1996 年 8 月 23 日收到修改稿.

可见,不同的 f 一般对应于不同的 $p(x)$,故有不同的不定性即熵 $H(x)$. 改变 f 便可调整 $p(x)$,从而调整熵 $H(x)$ 之大小,使之满足期望要求. 本文避开上述微观分析方法,从宏观的最大熵原理出发推断系统的 PDF,并进而研究系统的熵特性.

2 宏观信息熵方法

设宏观上可观测到 x 的各阶矩 $E x^i, i = 1, 2, \dots, M$. 现由这些宏观量导出 x 的 PDF $p(x)$. 因对系统其他特性一无所知,故应在已知的宏观量及非负和归一化约束条件下,使所推断的 PDF 具有最大的不定性. 这可由普适的最大熵原理^[4]表示,数学上即为受约束的泛函极值问题:

$$\begin{aligned} \max_p H(x) &= - \int_{-\infty}^{+\infty} p(x) \ln p(x) dx \\ \text{s. t. } \int_{-\infty}^{+\infty} p(x) dx &= 1, \quad \int_{-\infty}^{+\infty} x^i p(x) dx = E x^i, \quad i = 1, 2, \dots, M. \end{aligned} \quad (6)$$

通过引入乘子 $\lambda_0 - 1$ 和 $\lambda_i, i = 1, 2, \dots, M$, 可将其化为无约束泛函极值问题:

$$\max_p J = H(x) - (\lambda_0 - 1) \left[\int_{-\infty}^{+\infty} p(x) dx - 1 \right] - \sum_{i=1}^M \lambda_i \left[\int_{-\infty}^{+\infty} x^i p(x) dx - E x^i \right]. \quad (7)$$

运用变分法,可得到 $p(x)$ 的一般结构形式为:

$$p(x) = \exp(-\lambda_0 - \sum_{i=1}^M \lambda_i x^i), \quad (8)$$

$$\lambda_0 = \ln Z; \quad -\frac{\partial \ln Z}{\partial \lambda_i} = E x^i, \quad i = 1, 2, \dots, M. \quad (9)$$

其中 $Z(\lambda_1, \lambda_2, \dots, \lambda_M) = \int_{-\infty}^{+\infty} \exp(-\sum_{i=1}^M \lambda_i x^i) dx$ 称为配分函数. 与式(8)的 $p(x)$ 相应的 Shannon 信息熵 $H(x)$ 可用非常简洁的方式表出:

$$H(x) = - \int_{-\infty}^{+\infty} p(x) \ln p(x) dx = \lambda_0 + \sum_{i=1}^M \lambda_i E x^i. \quad (10)$$

注意到微观方法导出的 PDF 式(4)和(3)也具有指数形式,其指数项为 $\ln C + \frac{1}{Q} \int_{x_0}^x f(\theta) d\theta$. 在一定条件下,该指数项可用多项式(或幂级数) $-\sum_{i=0}^N a_i x^i$ 逼近,于是

$$p(x) = \exp(-a_0 - \sum_{i=1}^N a_i x^i), \quad H(x) = a_0 + \sum_{i=1}^N a_i E x^i. \quad (11)$$

这些分别与宏观最大熵方法导出的结果式(8)和(10)具有完全相似的形式. 可见,只要宏观约束选择恰当,最大信息熵方法便能推断系统的 PDF 结构和熵特性.

3 实例分析

考虑式(1)中 $f(x)$ 为非线性动力学及自组织理论中广泛应用的一类形式:

$$f(x) = \alpha x - \beta x^3, \quad \beta \geq 0, \quad x_0 = 0. \quad (12)$$

这是一类结构不稳定系统,系统的平衡点 $x_1^* = 0$ 在 $\alpha < 0$ 和 $\alpha > 0$ 时分别是稳定和不稳定的, $\alpha = 0$ 是临界点. 当 $\alpha > 0$ 时,系统还存在一对稳定的平衡点 $x_{2,3}^* = \pm \sqrt{\alpha/\beta}$. 借助 FPK 方程可导出:

$$p(x) = C \exp\left(\frac{\alpha}{2Q} x^2 - \frac{\beta}{4Q} x^4\right), \quad H(x) = -\ln C - \frac{\alpha}{2Q} E x^2 + \frac{\beta}{4Q} E x^4. \quad (13)$$

其中 C 为归一化常数. 在宏观最大熵方法中,宏观约束取到四阶矩为止. 考虑到系统的对称

性,奇数阶矩必趋于零,故宏观上能观测到非零的二阶和四阶矩.于是,

$$p(x) = \exp(-\lambda_0)\exp(-\lambda_2x^2 - \lambda_4x^4), \quad H(x) = \lambda_0 + \lambda_2Ex^2 + \lambda_4Ex^4.$$

(14)

其中拉格朗日乘子 λ_0, λ_2 和 λ_4 满足关系式(9).

可见,宏观最大熵方法导出的 $p(x)$ 和 $H(x)$ 与微观方法导出的结果具有完全类似的结构形式.不同的是,宏观结果的结构参数是由宏观量表出的乘子 λ_0, λ_2 和 λ_4 ,而微观结果的结构参数却由系统本身的特性参数 α, β 和 Q 表出,其对应关系为:

$$\lambda_0 = -\ln C, \quad \lambda_2 = -\alpha/2Q, \quad \lambda_4 = \beta/4Q.$$

(15)

表 1 列出了 $\alpha \ll 0, \alpha = 0$ 和 $\alpha \gg \beta$ 三种情况下的简化结果.

宏观信息综合反映了系统的各种特性,以信息为基础的最大熵原理能从宏观上推断系统的这些特性,比如随机和结构不稳定等等.注意到本例所导出的结果与系统的平衡点分支有关,说明系统的结构不稳定性必然反映在宏观信息熵上.

表 1 部分简化结果

条件	宏观结果	微观结果
$\alpha \ll 0$	$p(x) = (1/\sqrt{2\pi Ex^2})\exp[-x^2/(2Ex^2)]$ $H(x) = (1/2)\ln(2\pi eEx^2)$	$p(x) = \sqrt{ a /2\pi Q}\exp[(\alpha/2Q)x^2]$ $H(x) = (1/2)\ln(2\pi eEx^2), Ex^2 = Q/ a $
$\alpha = 0$	$p(x) = \exp(-\lambda_0)\exp(-\lambda_4x^4)$ $\lambda_0 = \ln Z, -\partial \ln Z / \partial \lambda_4 = Ex^4, Z \text{ 为配分函数}$ $H(x) = \lambda_0 + \lambda_4Ex^4$	$p(x) = C\exp[-(\beta/4Q)x^4]$ $C \text{ 为归一化常数}$ $H(x) = -\ln C + (\beta/4Q)Ex^4$
$\alpha \gg \beta$	$p(x) = \exp(-\lambda_0)\exp(-\lambda_2x^2 - \lambda_4x^4)$ $\lambda_0 = \lambda_2^2/4\lambda_4 - \ln \sqrt{ \lambda_2 /2\pi}$ $-\partial \ln Z / \partial \lambda_i = Ex^i, i = 2, 4,$ $H(x) = \lambda_0 + \lambda_2Ex^2 + \lambda_4Ex^4$	$p(x) = C \exp[(\alpha/2Q)x^2 - (\beta/4Q)x^4]$ $C = (1/2) \sqrt{a/\pi Q} \exp(-a^2/4\beta Q)$ $H(X) = -\ln C - (\alpha/2Q)Ex^2 + (\beta/4Q)Ex^4$

4 结 论

信息熵是系统的不定性测度.本文分别基于微观 FPK 方程和宏观信息熵方法进行系统结构推断和熵分析.微观方法导出的 PDF 和信息熵由系统的微观特性参数所描述,在一定条件下信息熵也可由宏观量表出.宏观方法借助普适的最大熵原理,导出的 PDF 和信息熵均由宏观量描述.文中通过一个实例进行了详细分析.

研究表明,只要宏观约束选取适当,宏观方法便能得到与微观方法相似的结果. Boltzmann 关于宏观熵的微观解释使我们相信,熵可作为联系宏观和微观的桥梁.不过,宏观约束的选取尚无通用标准,按热力学观点一般应从系统的矩和守恒量出发进行考虑.对于不同的系统,微观方法一般需分门别类加以处理,但宏观方法却具有类似的导出方式和结果,这使得用统一方式处理各类复杂系统成为可能.

参 考 文 献

1 田玉楚,符雪桐,孙优贤等.复杂系统与宏观信息熵方法.系统工程理论与实践,1995,15(8):62-68

2 Fuller, A. T.. Analysis of nonlinear stochastic systems by means of Foker-Plank equation. Int. J. Control, 1969,9(6): 603-655

3 田玉楚,徐功仁.非线性随机系统之熵分析.华东化工学院学报,1988,14(增刊):31-38

4 Jaynes, E. T.. On the rationale of maximum entropy methods. Proc. IEEE, 1982, 70(9):939-952

Macroscopic Analysis of Nonlinear Dynamical Systems Based on Information Entropy

TIAN Yuchu

(Institute of Industrial Process Control, Zhejiang University, National Key Laboratory of Industrial Control Technology • Hangzhou, 310027, PRC)

Abstract: In this paper, the characteristics of a class of nonlinear stochastic systems described by an equation of Itô type are analyzed by means of microscopic FPK equation and the approach of macroscopic information entropy respectively. The entropy properties of the systems are then studied. It is shown that the simple and direct macroscopic approach can derive results similar to that derived by microscopic approach. An application example is discussed also.

Key words: nonlinear systems; stochastic systems; system analysis; information; entropy; macroscopic approach

本文作者简介

田玉楚 1962年生,副教授。1982年学士学位毕业于武汉化工学院,1987年在华东化工学院获硕士学位,1993年在浙江大学工业控制研究所获博士学位,之后在浙江大学电机系做博士后研究两年。曾获浙江省青年科技奖,发表论文近40篇。研究方向为有关复杂系统的动力学,模型化,信息熵,智能控制与决策等。

中国科技期刊排行表

(按被引频次和影响因子排序)

数学类:

名次	期刊名称	被引频次	名次	期刊名称	影响因子
1	数学学报	306	1	控制理论与应用	0.2348
2	应用数学和力学	184	2	计算数学	0.2184
3	应用数学学报	157	3	数值计算与计算机应用	0.1818
4	计算数学	148	4	应用概率统计	0.1724
5	数学年刊·A	134	5	数学学报	0.1591
6	系统科学与数学	117	6	系统科学与数学	0.1579
7	控制与决策	108	7	系统工程学报	0.1573
8	系统工程理论与实践	104	8	应用数学学报	0.1486
9	数学研究与评论	86	9	高等学校计算数学学报	0.1413
10	应用概率统计	85	10	J Comput Math	0.1389

说明:

1. 中国科学引文数据库在连续两年公布《被引频次最高的中国科技期刊100名排行表》的基础上,为使统计数据的排列从多种角度反映科技期刊状况,特从1996年开始按学科编制《中国科技期刊排行表(按被引频次和影响因子排序)》。

2. 被引频次是在对被中国科学引文数据库1996年582种来源期刊所引用的数千种中国出版的中英文期刊进行频次统计后编制而成。

3. 影响因子的计算方法如下:

$$1996 \text{ 年某刊的影响因子} = \frac{1996 \text{ 年引用 } 1994 \text{ 年和 } 1995 \text{ 年该刊刊载论文的总次数}}{1994 \text{ 年和 } 1995 \text{ 年该刊刊载论文的总次数}}$$

本表中1996年的影响因子是在对中国科学引文数据库1994—1995年的来源期刊作了统计后编制而成。由于计算影响因子受到期刊发量数据的限制,因此,本表中只对能在中国科学引文数据库获得发量数据的315种期刊作了统计。1996年新增加的267种来源期刊因无发量数据而未作统计。

4. 本着尊重原始数据的原则,本表对变名期刊未作任何合并处理。

①数据来源:中国科学院文献情报中心中国科学引文数据库1996年数据。

本表由中国科学引文数据库统计编制。