

系 统 辨 识

——最小二乘法

夏天昌 著

刘绍球 朱可炎 译

李连升 校

国防工业出版社

内 容 简 介

辨识理论是近代系统理论的一个重要分支,它广泛应用于工业、农业、林业、国防及医学、经济学等领域。

本书系统地论述了辨识理论,重点讨论了最小二乘辨识方法,然后将最小二乘辨识方法与其他辨识方法(如互相关法、极大似然法、卡尔曼滤波法、辅助变量法及随机逼近法)建立联系,从而提供了各种可用的系统辨识方法。

本书通俗易懂,条理清楚,适合于各行各业从事系统辨识工作的科学工作者和工程技术人员阅读,也可作为大学的教科书或参考书。

SYSTEM IDENTIFICATION —LEAST-SQUARES METHODS

T. C. Hsia

Lexington Books

D. C. Heath and Company

Lexington, Massachusetts Toronto 1977

*

系 统 辨 识

——最小二乘法

夏天昌 著

刘绍球 朱可炎 译

李连升 校

*

国防工业出版社出版

新华书店北京发行所发行 各地新华书店经售

国防工业出版社印刷厂印装

*

850×1168¹/₃₂ 印张4¹/₂ 113千字

1984年5月第一版 1984年5月第一次印刷 印数:0,001—6,000册

统一书号:15034·2685 定价:0.69元

译 者 的 话

近年来，辨识理论的应用领域日益扩大，在通信工程、航空航天工程、机械工程、地质学、经济学、生物学、医学以及工业自动测量、自动调节等方面都得到应用，它已发展成近代系统理论的一个重要分支，而且有较重要的理论和实际意义。

什么是系统辨识呢？在社会实践中，人们往往需要通过测量和计算来辨明所研究和控制的系统（对象）内在结构和参数，这种问题就是系统辨识问题。在科学研究、生产实践和社会经济活动中，存在大量这类问题需要用系统辨识的理论去研究解决，这就是辨识理论发展的重要原因之一；另一方面，也由于现代计算工具的发展，使许多问题可以通过计算机加以解决，这又推动了辨识理论的发展。

夏天昌教授所著系统辨识一书，清晰地论述了系统辨识的基本理论和方法。本书的特点是深入地讨论了系统辨识的一种方法——最小二乘法，然后再讨论它与其他估计方法（如极大似然估计、贝叶斯估计、马尔可夫估计等）之间的联系，从而沟通了各种辨识方法。

近年来，系统辨识在我国也受到重视，一些高等院校和科研单位在理论上进行研究，并在经济部门、医学部门、冶金部门、农林部门、资源系统、航空航天工程等方面的应用进行了大量工作，取得了可喜的成果。我们期望本书的翻译和出版，将进一步推动我国系统辨识的研究和更广泛的应用。

高级工程师于景元同志对本书译稿进行了最后审订，在此谨致谢意。

前 言

系统表征和系统辨识在系统工程实践中是很基本的问题。系统表征主要涉及到建立描述系统输入-输出关系的数学模型。另一方面，系统辨识则论及到从一类数学上与给定的物理系统等效的模型中选择一个特定的模型。

系统辨识技术的应用已超出工程和自然科学的范畴。许多其他研究领域，例如生物学、医学及经济学，也能利用系统辨识方法建立这些领域所出现的系统的定量模型。关于这个专题已经出版了好几本书。

过去几年已经提出了各种有关系统辨识的技术。一般来说，辨识技术是由最优理论和估计理论导出的。大多数的书，在一本中包含多种方法。本书与众不同，重点放在最小二乘法上，把它作为系统辨识问题的一个基本解法。因为最小二乘法是各个领域的科学家经常实践的经典方法，所以本书能为广大读者所喜爱。重点放在最小二乘法上的另一个因素是其他流行的辨识方法，例如互相关法、极大似然法、卡尔曼滤波法、辅助变量法及随机逼近法，都可以容易地与最小二乘算法建立关系。因此，本书在一定程度上为综合和统一各种系统辨识方法提供了基础。

本书只讨论开环结构的系统，但本书的基本结果可用于辨识闭环系统。依据系统模型，重点放在输入-输出的表征（差分方程和权序列形式）上，而不是放在用状态变量的表征上。

本书在数学处理上力求通俗，以使最广泛的科学家和工程师能够成为本书的读者。但所选的参考文献中却包含了使有兴趣的读者可进一步探求其理论发展的内容。因而，本书也可作为大学学习系统工程的研究生的教科书。

目 录

第一章 绪论	1
§1.1 系统辨识问题	1
§1.2 系统辨识问题的综述和分类	2
§1.3 参数估计方法	3
§1.4 本书组成	4
第二章 动态系统的描述	6
§2.1 引言	6
§2.2 线性差分方程	6
§2.3 权序列和卷积	8
§2.4 状态变量方程	10
§2.5 结束语	13
第三章 最小二乘原理	15
§3.1 引言	15
§3.2 最小二乘原理	15
§3.3 最小二乘估计量的统计特性	17
§3.4 序列最小二乘估计	20
§3.5 多应变量系统	24
§3.6 参数个数增加时的递推估计	25
§3.7 实时最小二乘算法	27
§3.8 结束语	29
第四章 权函数辨识	33
§4.1 引言	33
§4.2 辨识问题	34
§4.3 最小二乘估计	35
§4.4 与互相关辨识的关系	38
§4.5 最优输入信号	42
§4.6 伪随机二进制序列	44

§4.7 在线最小二乘辨识	48
§4.8 多变量系统辨识	51
§4.9 结束语	54
第五章 用最小二乘法进行线性参数模型辨识	57
§5.1 引言	57
§5.2 基本的辨识问题	58
§5.3 最小二乘解法	59
§5.4 参数估计的统计特性	61
§5.5 在线最小二乘辨识	65
§5.6 系统阶次的确定	66
§5.7 实时辨识	71
§5.8 连续系统辨识	73
§5.9 结束语	77
第六章 用广义最小二乘法的系统辨识	82
§6.1 引言	82
§6.2 噪声系统模型的叙述	82
§6.3 与相关残差有关的偏倚问题	83
§6.4 广义最小二乘问题的叙述	85
§6.5 广义最小二乘估计算法	86
§6.6 评论	90
§6.7 另一 GLS 求解技术	95
§6.8 辅助变量法	102
§6.9 结束语	105
第七章 多步最小二乘辨识技术	107
§7.1 引言	107
§7.2 MSLS 方法 I	107
§7.3 MSLS 方法 II	112
§7.4 MSLS 方法 III	113
§7.5 MSLS 法与 GLS 法的比较	117
§7.6 结束语	120
第八章 非线性系统的辨识	121
§8.1 引言	121

§8.2	伏尔特拉级数表示及辨识	121
§8.3	具有线性参数的非线性差分方程	123
§8.4	具有非线性参数的非线性差分方程	124
§8.5	哈默斯坦模型——GLS 辨识	126
§8.6	哈默斯坦模型——MSLS 辨识	129
§8.7	结束语	134

第一章 绪 论

§ 1.1 系统辨识问题

一般认为系统辨识问题就是通过观测一个系统或一个过程的输入-输出关系来确定其数学模型的问题。本书的目的是介绍系统辨识问题的基本理论和解法。本章结尾列出了关于这个问题的一般参考文献。

最近十年来,系统辨识方法论已经取得了巨大的进展。在历史上,由于设计更好的控制系统的需要,促进了系统辨识的发展。在大多数实际系统中,例如工业生产过程,常常没有足够的有关系统及其环境的先验信息可用来设计一个有效的控制方案。我们常常面临必须通过实验确定某些重要的物理参数,例如热传导系数、化学反应率、阻尼系数等等。由于最优控制和自适应控制理论的发展,对于高度精确的系统模型的需求也更加迫切,例如,在自适应系统设计中,常常需要修正设备及其环境的某些时变参数值,以便在整个过程中保持系统的最优性能。关于动态系统参数估算的其他工程应用,还包括信道检查、系统及故障检测。

在其它许多学科中,也出现需要建立模型的问题。例如,在计量经济学领域内的工作者,长期以来一直在寻求建立内生变量(输出)与外生变量(输入)之间的数学关系。最近,系统辨识技术对于生理学及生物医学问题的应用已经有了很大的进展。对于人-机器环境中人的性能、瞳孔和肌肉的控制功能、新陈代谢以及脑电波等等,已经获得了很成功的模型。因此,系统辨识这门技术正引起医学和生物科学界更大的兴趣。类似的建模应用也可以在生态学、交通运输及社会学这样的领域中找到。此外,现代估计理论及复杂算法的有效性促进了系统辨识技术的迅速发展。

1111652

§ 1.2 系统辨识问题的综述和分类

图 1.1 给出一个具有输入和输出的系统。我们要寻求的系统模型就是在整个过程中使输入与输出发生关系的数学方程。为了得到这样一个模型，我们可以用各种输入来试探该系统并观测其响应，然后将输入-输出数据进行处理来得到模型。基于对系统先验信息的了解程度，我们可以把系统辨识问题分为两类：

1. 完全辨识问题

这意味着我们不知道有关系统基本特性的任何情况，例如是线性还是非线性，是无记忆还是有记忆等等。显然这是一个很难解决的问题。通常，在尝试任何有意义的解法之前，都必须作出某种假设。这种类型的问题又被称作为“黑箱”问题。

2. 部分辨识问题

在这一类问题中，系统的某些基本特征，例如线性、带宽等等，假定是已知的。但是，我们可能不知道动态方程的特定的阶次或它的系数值。这类问题也叫“灰箱”问题。当然，这类问题比“黑箱”问题要容易处理。

幸亏我们实践中遇到的大多数工程系统和工业过程都属于后一种类型。在许多情况下，我们对系统的结构有较多的了解，以致有可能导出系统动力学的一个特定的数学模型。因而，仅仅模型方程的一组参数需待确定。这样，建立模型的问题简化成了参数辨识问题。

因为大量的系统辨识问题可以表述为或简化为参数辨识问题，所以后者的处理被认为具有极大的重要性。目前进行的研究以及在这一领域的大多数研究成果，主要针对这类问题。

从系统理论的观点来看，只要已知输入-输出数据的精确测量值，我们就能精确地定出确切的系统模型方程的未知参数。然而，实际上，输入-输出数据受到测量噪声的干扰；而且，模型方程存在不精确性；系统本身也存在随机干扰。因此系统参数的确定实质上是一个统计估值问题：我们试图确定一个拟合于带噪声

观测数据的数学模型。

实现系统辨识的过程可以分为以下几个步骤：

1. 选定并参数化一类代表被辨识系统的数学模型。
2. 在系统上加一个适当选择的试验信号，并且记录其输入-输出数据。如果系统处在连续运行中并且不允许加试验信号，则我们必须用正常运行数据来辨识。

3. 完成参数辨识，以便在给定的类型中选定模型，使其最好地拟合于统计数据。

4. 进行有效性检验以考核所选模型对于最终的辨识对象来说是否适当地代表了该系统。

5. 如果有效性检验通过，则辨识过程结束，否则必须选择另一类模型并且重复步骤 2 到 4，直到获得有效的模型为止。

第一步主要涉及到描述的问题。对于一个给定的系统，我们可以选择许多可能的描述，包括在频域或时域中表征的模型、以连续时间或离散时间表征的模型。以时间域描述时，我们要在权函数（或序列）、微分（或差分）方程及状态变量方程之间作选择。这种选择取决于辨识的对象及其有关的输入-输出数据。

§ 1.3 参数估计方法

目前已有许多著名的参数估计技术成功地用于辨识问题，它们是极大似然法、最小二乘法、互相关法、辅助变量法以及随机逼近法。也有人提出了其他各种最优方法。

由于许多重要的原因，我们主要依靠最小二乘法。其原因首

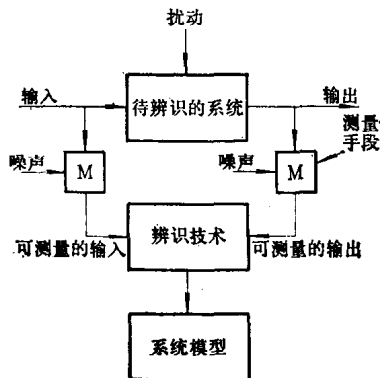


图1.1 系统辨识问题的方框图

先是最小二乘法是一个很经典的方法，许多领域的科学工作者对它很熟悉，因此它容易受到重视。其次，最小二乘法是很吸引人的，因为它提供了简单的概念，并且广泛地适用于其他统计估值理论难于应用的情况；它还表明其统计特性与用于许多实际情况的极大似然法的统计特性一样好。此外，最小二乘辨识算法可以容易地与其他许多辨识算法建立联系，从而有可能统一处理系统辨识问题。

正如所有其他估计技术那样，我们试图把某种适当定义的误差判据减到最小，作为使模型最优地拟合于系统数据的一种方法。定义误差有许多方法：如参数估计值对真值的偏差（参数误差）；对于同样的输入，系统输出与模型输出之差（输出误差）；模型方程与所测量的输入、输出之间的偏差（方程误差）。方程误差是最通常采用的。

完成辨识有两种方式。一种叫离线辨识，在这种方式中，首先要有所观测的输入-输出数据的记录，然后在整个数据记录的基础上估算模型参数。另一种叫在线辨识，参数估计是对每组数据作递归计算，并用新的数据来修正和更新现有估计值。显然，如果更新过程能很快进行，则有可能以适当的精度获得时变系统的参数估计，这种能力叫做在线实时辨识。

§ 1.4 本书组成

本书第二章研究动态系统的参数模型表示法，重点是离散时间线性系统模型，并指出了各种表达形式之间的关系。第三章介绍了最小二乘理论及有关的统计特性，并与其他估计技术之间建立了联系。第四章描述了最小二乘法对于权序列辨识的应用，介绍了可辨识性及最优试验信号设计的概念。第五章描述了最小二乘法对于简单情况下线性差分方程系统模型辨识的应用。所谓简单情况即随机方程误差是白噪声（统计独立的）；也研究了模型阶次的辨识问题。第六章研究了有色随机方程误差的情况，并且把广义最小二乘法作为基本辨识方法加以介绍。在第七章中，我

们试图用一些多步最小二乘法来求解同样的问题。这些方法在计算上比第六章讨论过的广义最小二乘技术更简单。第八章讨论了非线性系统模型的辨识问题，并且特别注重于伏尔特拉 (Volterra) 级数模型及哈默斯坦 (Hammerstein) 非线性模型。

参 考 文 献

- Åström, K. J., and Eykhoff, P., "System Identification—a Survey," *Automatica*, Vol. 7, pp. 123-162, 1971.
- Bekey, G. A., "System identification—an Introduction and a Survey," *Simulation*, pp. 151-166, Oct. 1970.
- Cuenod, M., and Sage, A. P., "Comparison of some Methods Used for Process Identification," *Automatica*, Vol. 4, pp. 235-269, 1968.
- Eykhoff, P., *System Identification, Parameter and State estimation*, Wiley, London, 1974.
- , "Some Fundamental Aspects of Process-Parameter Estimation," *IEEE Transactions on Automatic Control*, Vol. AC-8, pp. 317-357, Oct. 1963.
- Graupe, D. *Identification of Systems*, Van Nostrand Reinhold, New York 1972.
- Isermann, R., Baur, R., Bamberger, W., Kneppo, P., and Siebert, H., "Comparison of Six On-line Identification and Parameter Estimation Methods," *Automatica*, Vol. 10, pp. 81-103, Jan. 1974.
- Kagiwade, H. H., *System Identification Methods and Applications*, Addison-Wesley, Reading, Mass., 1974.
- Lee, R. C. K., *Optimal Estimation, Identification, and Control*, MIT Press, Cambridge, Mass., 1964.
- Mendel, J. M., *Discrete Techniques of Parameter Estimation*, Marcel Dekker, New York, 1973.
- Phillipson, G. A., *Identification of Distributed Systems*, American Elsevier, New York, 1971.
- Saridis, G. N., "Comparison of Six On-line Identification Algorithms," *Automatica*, Vol. 10, pp. 69-79, Jan. 1974.
- Sage, A. P., and Melsa, J. L., *System Identification*, Academic Press, New York, 1971.
- Zadeh, L. A. "On the Identification Problem," *IRE Transactions on Circuit Theory*, Vol. CT-3, Dec. 1956.

第二章 动态系统的描述

§ 2.1 引 言

动态系统模型可分为两种类型：连续时间系统和离散时间系统。两者之间的基本差别是：前者，系统的信号是连续的；而后者，系统的信号是离散的。因为连续信号的信息可以保存在以适当的采样频率所取的样本中，所以连续系统可以用离散模型来近似。关于系统理论的一些一般性参考文献列于本章结尾。

由于数字计算机使用的结果，大量的系统辨识技术向着数字化方向发展。因而离散系统模型更便于处理。本书中最小二乘辨识技术的发展也是建立在离散系统概念的基础上的。

本章，我们介绍线性离散动态系统的各种表示形式。每种形式都以不同的方式表征了系统动力学。所研究的形式有卷积求和、差分方程以及状态变量方程。我们指出这些形式相互间的关系是唯一的，因此一旦获得了特定的表示形式，则把线性系统看成是已经完全辨识了。

§ 2.2 线性差分方程

我们首先介绍单变量定常线性离散系统的差分方程表示法。参照图 2.1 的方框图，输入 $u(k)$ 与输出 $y(k)$ 之间的一般 n 阶差分方程为

$$\begin{aligned} y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) \\ = b_0 u(k) + b_1 u(k-1) + \cdots + b_n u(k-n) \end{aligned}$$

或

$$y(k) + \sum_{j=1}^n a_j y(k-j) = \sum_{j=0}^n b_j u(k-j) \quad (2.1)$$

式中 k 为整时间变数， a_j 和 b_j 为常数。

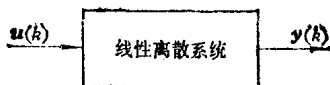


图2.1 单输入和单输出的离散系统

如果我们引入下面定义的移位算子 q

$$q^{-1}y(k) = y(k-1)$$

以及多项式

$$\begin{aligned} A(q^{-1}) &= 1 + a_1q^{-1} + a_2q^{-2} + \cdots + a_nq^{-n} \\ B(q^{-1}) &= b_0 + b_1q^{-1} + b_2q^{-2} + \cdots + b_nq^{-n} \end{aligned} \quad (2.2)$$

方程 (2.1) 可写为下列形式

$$A(q^{-1})y(k) = B(q^{-1})u(k) \quad (2.3)$$

这是我们将用于系统辨识的最基本的形式。

在这点上, 指出单变量定常系统的差分方程表示和传递函数表示之间的简单关系是方便的。对方程 (2.1) 应用 Z 变换, 假定初始条件为零, 即 $y(k) = u(k) = 0, k < 0$, 我们可以得到下式

$$\begin{aligned} (1 + a_1z^{-1} + \cdots + a_nz^{-n})Y(z) \\ = (b_0 + b_1z^{-1} + \cdots + b_nz^{-n})U(z) \end{aligned}$$

这里 z 是 Z 变换的复变量。那么传递函数可定义为

$$H(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1z^{-1} + \cdots + b_nz^{-n}}{1 + a_1z^{-1} + \cdots + a_nz^{-n}} \quad (2.4)$$

或简写为

$$H(z) = \frac{B(z^{-1})}{A(z^{-1})}$$

式中多项式 $A(z^{-1})$ 和 $B(z^{-1})$ 由方程 (2.2) 确定。可见, 系统的传递函数是直接与其差分方程 (2.3) 联系在一起的。

方程 (2.1) 的表示形式可以推广到多输入和多输出系统。我们来考虑有 m 个输入和 r 个输出的系统 (见图2.2), 并且定义矢量 $u(k)$ 和 $y(k)$ 为

$$\mathbf{u}(k) = \begin{bmatrix} u_1(k) \\ \vdots \\ u_m(k) \end{bmatrix} \quad \mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_r(k) \end{bmatrix}$$

那么, 系统可用矢量差分方程表示

$$\mathbf{y}(k) + \sum_{j=1}^n \mathbf{A}_j \mathbf{y}(k-j) = \sum_{j=0}^n \mathbf{B}_j \mathbf{u}(k-j) \quad (2.5)$$

式中 $\mathbf{A}_j$ 和 $\mathbf{B}_j$ 分别是 $r \times r$ 维、 $r \times m$ 维的常系数矩阵。然而这里要注意, 一般情况下, 方程(2.5)中的系统阶数不一定是 n , 也可能与 n 不同。

方程(2.5)也可以写成如下形式

$$\mathbf{A}(q^{-1})\mathbf{y}(k) = \mathbf{B}(q^{-1})\mathbf{u}(k) \quad (2.6)$$

式里 $\mathbf{A}(q^{-1})$ 和 $\mathbf{B}(q^{-1})$ 为下面所定义的 q^{-1} 的矩阵多项式

$$\mathbf{A}(q^{-1}) = \mathbf{I} + \mathbf{A}_1 q^{-1} + \cdots + \mathbf{A}_n q^{-n}$$

$$\mathbf{B}(q^{-1}) = \mathbf{B}_0 + \mathbf{B}_1 q^{-1} + \cdots + \mathbf{B}_n q^{-n}$$

§ 2.3 权序列和卷积

线性离散系统的另一基本表征是权序列。权序列定义为弛张系统在 $t=0$ 时对单位脉冲(δ 函数)的响应。令单变量系统的权序列表示如下

$$\{h(i)\} \quad i = 0, 1, 2, 3, \dots \quad (2.7)$$

那么建立系统输入-输出关系的卷积求和便是

$$y(k) = \sum_{i=-\infty}^k h(k-i) u(i) \quad (2.8)$$

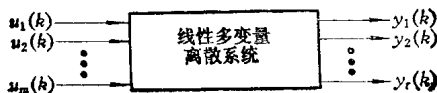


图2.2 多输入和多输出的离散系统

当在方程(2.7)中对 $i > p$, 权序列小到可忽略时, 无穷

和实际上可以简化。在这种情况下，我们有如下近似

$$y(k) = \sum_{i=k-p}^k h(k-i)u(i) \quad (2.9)$$

众所周知，权序列的 Z 变换定义为传递函数，也即

$$Z[h(k)] = H(z) \quad (2.10)$$

此方程在建立系统权序列与其差分方程表示式之间的关系时，是很有用的。我们现在来仔细研究这个关系式。

回顾方程(2.4)的传递函数 $H(z)$ 。用长除法展开 $H(z)$ 并且应用方程(2.10)，我们可以表示如下

$$\frac{b_0 + b_1 z^{-1} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots \quad (2.11)$$

方程(2.11)●再写成

$$b_0 + b_1 z^{-1} + \dots + b_n z^{-n} = (h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots)(1 + a_1 z^{-1} + \dots + a_n z^{-n})$$

然后比较等号两边 z^{-1} 的同类项的系数，则得到关系式

$$\sum_{m=0}^i a_m h(i-m) = \begin{cases} b_i & i = 0, 1, \dots, n \\ 0 & i > n \end{cases} \quad (2.12)$$

式中 $a_0=1$ 。这组方程把权序列 $h(i)$ 直接与差分方程(2.1)的系数 a_i 和 b_i 连系起来了。

对于 m 个输入和 r 个输出的多变量系统，相应的表达式变为权矩阵 $H(k)$

$$H(k) = \begin{bmatrix} h_{11}(k) & \dots & h_{1m}(k) \\ \vdots & & \vdots \\ h_{r1}(k) & \dots & h_{rm}(k) \end{bmatrix} \quad (2.13)$$

式中 $h_{ij}(k)$ 为第 j 个输入与第 i 个输出之间的权序列。图2.3描述了这样一个多变量系统，相应的卷积求和为

$$y(k) = \sum_{i=-\infty}^k H(k-i)u(i) \quad (2.14)$$

● 原书误为方程(2.5)。——译者注

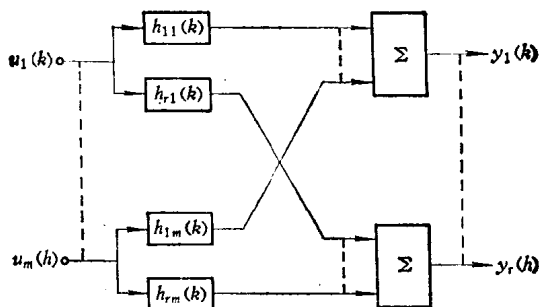


图2.3 具有 m 个输入、 r 个输出及
权序列 $h_{ij}(k)$ 的多变量系统

§ 2.4 状态变量方程

图 2.1 的单变量系统也可以用状态变量方程来描述

$$\begin{aligned} \mathbf{x}(k+1) &= \Phi \mathbf{x}(k) + \Gamma u(k) \\ \mathbf{y}(k) &= \mathbf{G} \mathbf{x}(k) + \mathbf{D} u(k) \end{aligned} \quad (2.15)$$

式中 $\mathbf{x}(k)$ 为 $n \times 1$ 状态矢量, Φ 、 Γ 、 $\mathbf{G}$ 、 $\mathbf{D}$ 分别是 $n \times n$ 、 $n \times 1$ 、 $1 \times n$ 、 1×1 维的参数矩阵。状态变量方程的方框图描述如图 2.4 所示。

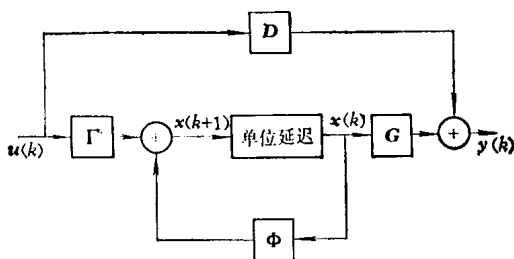


图2.4 线性离散系统的状态变量方程表示

这里我们假定状态变量系统是能控的和能观测的。(当且仅当复合矩阵 $\mathbf{M}$

$$\mathbf{M} = [\Gamma, \Phi\Gamma, \Phi^2\Gamma, \dots, \Phi^{n-1}\Gamma]$$