

现代控制论入门

徐 强 俞修海 施关澄 编



上海科学技术出版社

现代控制论入门

徐 强 俞修海 施兴澄 编

上海科学技术出版社

内 容 简 介

本书是一本现代控制论的入门读物。它主要介绍最优控制的理论基础及应用方法。

全书共分六章,内容包括:最优控制;最佳估计;随机控制及系统识别。

本书由浅入深,数学推理详细,举例典型,既可使读者对现代控制论的全貌有所了解,指导工程中经常遇到的控制技术问题的,又可为进一步深造和研究打下基础。本书可供从事自动控制的工程技术人员和大专院校师生参考。

现代控制论入门

徐强 俞修海 施关澄 编

上海科学技术出版社出版

(上海瑞金二路 450 号)

新华书店上海发行所发行 上海市印刷三厂印刷

开本 787×1092 1/32 印张 8 字数 175,000

1986 年 6 月第 1 版 1986 年 6 月第 1 次印刷

印数: 1—6,000

统一书号: 15119·2462 定价: 1.35 元

前 言

五十年代末到六十年代初,自动控制理论经历了一场重大的变革。在这以前,控制理论的主要研究对象基本上限于线性、定常、单输入-单输出系统及一些简单的非线性系统。对线性定常系统的研究主要采用频率法和根轨迹法,对非线性系统则主要采用相平面法及描述函数法。通常把这一时期的理论称为经典控制理论。

后来,由于空间技术的发展及生产过程日趋复杂,需要研究大量非线性系统及时变系统,而且它们的参数受到某些限制,控制过程也往往不再局限于单参数调节,而要求实现多参数调节,某些系统还要求考虑各种随机因素和随机干扰,并要实现最优控制。因此原来的经典控制理论已不能满足要求。随着计算机技术的迅速发展,状态空间概念被成功地应用于自动控制领域。

现代控制理论主要包括确定性最优控制、最佳估计、随机控制、系统识别及自适应控制等。在确定性最优控制方面已产生了许多理论和方法。其中解析法包括变分法、庞脱里雅金原理及动态规划等;数值分析法包括最速下降法、牛顿-康托罗维奇法及梯度法等,它们都能用来处理具有不等式约束的时变及非线性系统。最佳估计最重要的是卡尔曼滤波;在随机控制方面,主要的成果是具有高斯输入的线性系统的分离定理;在系统识别和自适应控制中,已应用了卡尔曼滤波及一

些重要的数值方法。本书将讨论确定性控制理论、估计问题及识别问题。因为它们涉及的内容相当广泛，不可能在这样一本入门小册子中作详细讨论，因此这里只是将一些主要内容作简要叙述，以便读者对现代控制理论的概貌有个初步的了解。

本书是在1979年上海自动化学会举办的现代控制论普及讲座基础上编写而成的。其目的是使工程界的技术人员能较快地掌握这方面的基本理论，并起到经典控制与现代控制之间的桥梁作用。

本书的编审和组织工作都是在上海自动化学会名誉理事长胡汝鼎教授的关怀和指导下进行的，在此谨表谢意。由于编者水平有限，书中缺点和错误难免，殷切希望广大读者批评指正。

编 者

1983 年

目 录

第一章 动力学系统的状态空间分析	1
第一节 状态、状态变量及状态方程	1
第二节 线性状态方程的解	11
第三节 离散系统的状态方程及其解	15
第四节 线性系统的可控性及可观测性	22
第二章 确定性系统的最优控制	35
第一节 最优控制问题的描述	36
第二节 无约束条件的变分法	42
第三节 等式约束条件下的动态最优化	49
第四节 波尔萨 (Bolza) 问题	58
第五节 带有不等式约束条件的变分法	61
第六节 极小值原理	63
第七节 两点边值问题的计算及其困难	71
第八节 具有二次型性能指标的线性系统的最优控制	73
第九节 状态反馈、输出反馈及状态重构	86
第三章 动态规划	98
第一节 动态规划的基本概念	99
第二节 离散系统最优控制的动态规划	103
第三节 哈密顿-雅可比-贝尔曼方程	110
第四章 最佳估计	114
第一节 维纳滤波	115
第二节 卡尔曼滤波(连续系统)	120
第三节 卡尔曼滤波(离散系统)	122

第四节	最小二乘估计	131
第五节	线性最小方差估计	136
第六节	最小方差估计	140
第七节	极大似然估计与极大验后估计	143
第八节	最优估计的贝叶斯方法	145
第五章	系统识别	160
第一节	数学初步	160
第二节	线性系统的识别	168
第三节	从冲激响应曲线求传递函数	181
第四节	自动识别	206
第六章	随机控制系统	216
第一节	确定性离散时间系统最优控制	216
第二节	离散随机控制问题	226
第三节	连续时间系统的随机控制	242
第四节	应用实例	244

第一章

动力学系统的状态空间分析

经典控制理论建立在系统输入和输出的传递函数的基础上,而现代控制理论则建立在系统的状态和状态空间分析的基础上。控制系统的状态空间分析不仅给出有关系统输入、输出关系的信息,而且还给出系统内部所有状态变量的信息。因此对系统来说,状态空间分析是一种更全面描述系统性能的方法。此外,应用状态空间分析法较容易将分析结果推广到非线性系统和时变系统。

第一节 状态、状态变量及状态方程

系统 t_0 时刻的状态是系统在 t_0 时刻的一组信息量,它与控制作用 $u(t)$ 一起,可以唯一地确定系统在所有 $t > t_0$ 时的行为。

例如在图 1-1 所示的网络中,若已知电感中的初始电流 $i_L(t_0)$ 和电容上的初始电压 $u_C(t_0)$,则网络任意时刻的状态都能唯一地被确定。故 $i_L(t_0)$ 和 $u_C(t_0)$ 可作为这一系统在 t_0 时刻的状态。

能够完全描述系统状态的最小的一组变量称为状态变量。例如图 1-1 所示的网络中,在控制电压 $u(t)$ 作用下,可

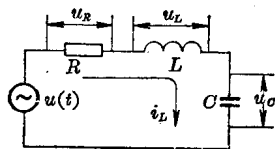


图 1-1 RLC 网络

有 i_L 、 u_C 、 u_R 、 u_L 等变量，但是我们只需要知道 i_L 和 u_C 就已经能够完全描述这一系统的行为了。故 i_L 和 u_C 这一组变量就可以选作这一系统的状态变量。当然我们也可以选 u_R 和 u_C 作为状态变量。由此可见，状态变量并非是唯一的。但是一组状态变量必为另一组状态变量的线性组合。

在选择系统状态变量时，应注意以下几点：

(1) 状态变量必须是能够完全描述系统的最小的一组变量。即这些变量之间应该是相互独立的。

(2) 状态变量一般选择系统里的储能元件上与储能有关的参量，例如电感中的电流、电容上的电压、弹性元件的位移、惯性元件的速度、加速度等。

(3) 状态变量一般具有物理意义，例如电流、电压、速度、加速度等。但有时由于数学上的需要，也可以用一些物理量的线性组合作为状态变量，此时的状态变量就不一定有明确的物理意义。

根据以上选择状态变量的原则，对于图 1-1 所示的系统，我们可以选择 i_L 和 u_C 作为状态变量。系统的微分方程为

$$LC \frac{d^2 u_C}{dt^2} + RC \frac{du_C}{dt} + u_C = u \quad (1-1)$$

令 $x_1 = i_L = C \frac{du_C}{dt}$ ， $x_2 = u_C$ ，则微分方程(1-1)可写为

$$L \dot{x}_1 + R x_1 + x_2 = u$$

整理后得

$$\dot{x}_1 = -\frac{R}{L} x_1 - \frac{1}{L} x_2 + \frac{1}{L} u \quad (1-2)$$

$$\dot{x}_2 = \frac{1}{C} x_1 \quad (1-3)$$

这样，我们就把一个二阶微分方程变成了两个一阶微分方程。

将方程(1-2)和(1-3)写成矩阵形式得

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} u$$

这一方程称为系统的状态方程。

系统的输出为

$$y = u_c = x_2$$

写成矩阵形式为

$$y = [0 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

称为系统的输出方程。

状态方程和输出方程一起,称为系统的动力学方程,它可以完全描述系统。系统的这种数学模型称为状态空间描述。

在求系统的动力学方程时,可以不必先写出系统的微分方程。在假设好状态变量后,可以直接用克希霍夫第一、第二定律写出动力学方程。

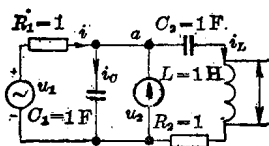


图 1-2 例 1-1 的系统图

例 1-1 设系统如图 1-2 所示。图中 u_1 为电压源, u_2 为电流源。

$$\text{令 } x_1 = u_{C1}; \quad x_2 = u_{C2}; \quad x_3 = i_L$$

对节点 a 用克希霍夫第一定律得

$$\frac{u_1 - x_1}{R_1} - C_1 \dot{x}_1 + u_2 - x_3 = 0$$

整理后得

$$\dot{x}_1 = -\frac{1}{R_1 C_1} x_1 - \frac{1}{C_1} x_3 + \frac{1}{R_1 C_1} u_1 + \frac{1}{C_1} u_2$$

$$= -x_1 - x_3 + u_1 + u_2$$

电容 C_2 上电流和电压间的关系为

$$x_3 = C_2 \frac{dx_2}{dt}$$

即

$$\dot{x}_2 = \frac{1}{C_2} x_3 = x_3$$

于是得到系统的状态方程:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

输出方程为

$$y = \dot{x}_3 = [1 \quad -1 \quad -1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

系统动力学方程可用一个简单的形式来表示:

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{U} \quad (1-4)$$

$$\mathbf{Y} = \mathbf{C}\mathbf{X} \quad (1-5)$$

这里

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

• 4 •

$$C = [1 \quad -1 \quad -1].$$

如果系统的数学模型为 n 阶线性常系数微分方程

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} \dot{y} + a_n y = u \quad (1-6)$$

时, 可选取 n 个状态变量

$$\begin{aligned} x_1 &= y \\ x_2 &= \dot{y} \\ &\vdots \\ x_n &= y^{(n)} \end{aligned}$$

于是, 方程(1-6)可改写为 n 个一阶微分方程所组成的方程组

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + u \end{cases} \quad (1-7)$$

令

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \dot{X} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix}$$

则(1-6)式所描述的系统就可用状态方程

$$\dot{X} = AX + BU$$

来描述。这里

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}_{n \times n}$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{n \times 1}$$

用 Y 表示系统的输出 y , 则有输出方程

$$Y = CX \quad (1-8)$$

其中 $C = [1 \ 0 \cdots 0]_{1 \times n}$.

更一般的情况, 当动态系统为图 1-3 所示的多输入-多输出系统时

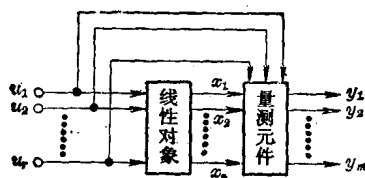


图 1-3 多输入-多输出系统

图中: $u_1, u_2, \dots, u_r$ 为输入变量; $x_1, x_2, \dots, x_n$ 为状态变量; $y_1, y_2, \dots, y_m$ 为输出变量。

系统的动特性可用下面的方程组来描述:

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}u_1 + b_{12}u_2 \\ &\quad + \cdots + b_{1r}u_r \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}u_1 + b_{22}u_2 \\ &\quad + \cdots + b_{2r}u_r : \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_{n1}u_1 + b_{n2}u_2 \\ &\quad + \cdots + b_{nr}u_r \end{aligned}$$

式中 a_{ij} 和 b_{ij} 为常数或时间 t 的函数。引进矩阵记号, 可写成

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{U} \quad (1-9)$$

式中:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ 为 } n \text{ 维状态矢量;}$$

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{bmatrix} \text{ 为 } r \text{ 维控制矢量;}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ 为 } n \times n \text{ 系统矩阵;}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ & & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{bmatrix} \text{ 为 } n \times r \text{ 控制矩阵.}$$

类似地, 对于输出有

$$y_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n + d_{11}u_1 + d_{12}u_2 + \cdots + d_{1r}u_r$$

$$y_2 = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n + d_{21}u_1 + d_{22}u_2 + \cdots + d_{2r}u_r$$

$\vdots$

$$y_m = c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n + d_{m1}u_1 + d_{m2}u_2 + \cdots + d_{mr}u_r$$

用矩阵方程表示为

$$\mathbf{Y} = \mathbf{C}\mathbf{X} + \mathbf{D}\mathbf{U} \quad (1-10)$$

式中:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \text{ 为 } m \text{ 维输出矢量;}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \vdots & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

为 $m \times n$ 矩阵, 其中的元素为常数或为时间 t 的函数;

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1r} \\ d_{21} & d_{22} & \cdots & d_{2r} \\ & & \vdots & \\ d_{m1} & d_{m2} & \cdots & d_{mr} \end{bmatrix}$$

为 $m \times r$ 矩阵, 其中的元素为常数或为时间 t 的函数。

用状态方程(1-9)和输出方程(1-10)所描述的系统方框图如图1-4所示。

用状态方程描述系统和用状态空间分析法分析设计系统给自动控制工作者带来如下许多方便:

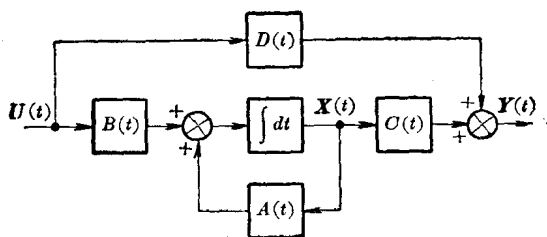


图 1-4 由(1-9)和(1-10)式定义的系统方框图

(1) 分析一个复杂的系统时, 可能要求解 n 阶的高阶微分方程, 这是一件很困难的事。但是, 应用状态空间的概念则可将 n 阶微分方程的求解化成 n 个一阶微分方程组的求解。而后者要简便得多, 而且当系统的阶数增加时, 求解的复杂程度并不增加多少。

(2) 引进微分方程组的矩阵表达式, 可大大简化系统的数学表达式, 可以应用现有的一套矩阵理论来研究系统的性能, 是研究系统的又一个有力工具。

(3) 应用状态空间处理方式的另一个显著优点是可以推广到非线性及时变系统。

(4) 控制系统的经典理论研究的是系统的输入、输出关系。但当研究一个复杂的系统, 例如一个化工过程时, 我们不仅要研究系统的输入和输出, 还必须研究系统全部有关的信号。例如, 当需要确定这些变量是否在线性范围里的时候, 就必须知道系统内部各变量的幅值因为很可能出现由于没有注意这些变量的幅值而造成系统的不稳定。用经典理论是很难满足这一要求的, 而应用状态空间分析法就自然地把对系统内部变量的研究包括进去了。

此外, 在工程实际中, 所遇到的控制系统通常都是非线性的。由于对非线性系统的研究往往很困难, 而且多数非线性系统常常可在一定的范围内用线性系统来逼近。因此, 通常我们总是希望通过模型线性化的方法, 将实际的非线性系统化为线性系统来研究。

设原来控制系统的状态方程和输出方程分别为:

$$\dot{\mathbf{X}}(t) = f[\mathbf{X}(t), \mathbf{U}(t), t] \quad (1-11)$$

$$\mathbf{Y}(t) = g[\mathbf{X}(t), t] \quad (1-12)$$

式中 f 和 g 为非线性函数。

我们只限于考察系统在 $\mathbf{X}_0(t)$ 附近的运动, $\mathbf{X}_0(t)$ 称为标称轨道或标称状态。 $\mathbf{U}_0(t)$ 是为了达到标称状态所需要的控制量, $\mathbf{Y}_0(t)$ 是根据标称状态 $\mathbf{X}_0(t)$ 算出的输出量。于是有

$$\dot{\mathbf{X}}_0(t) = f[\mathbf{X}_0(t), \mathbf{U}_0(t), t] \quad (1-13)$$

$$\mathbf{Y}_0(t) = g[\mathbf{X}_0(t), t] \quad (1-14)$$

对在标称点附近的微小变化, 定义摄动矢量

$$\delta \mathbf{X}(t) = \mathbf{X}(t) - \mathbf{X}_0(t)$$

$$\delta \mathbf{U}(t) = \mathbf{U}(t) - \mathbf{U}_0(t)$$

$$\delta \mathbf{Y}(t) = \mathbf{Y}(t) - \mathbf{Y}_0(t)$$

根据函数的泰勒展开式, 有

$$f[\mathbf{X}(t), \mathbf{U}(t), t] = f[\mathbf{X}_0(t), \mathbf{U}_0(t), t] + \left. \frac{\partial f}{\partial \mathbf{X}} \right|_0$$

$$\delta \mathbf{X}(t) + \left. \frac{\partial f}{\partial \mathbf{U}} \right|_0 \delta \mathbf{U}(t) + \alpha[\delta \mathbf{X}(t), \delta \mathbf{U}(t), t]$$

$$g[\mathbf{X}(t), t] = g[\mathbf{X}_0(t), t] + \left. \frac{\partial g}{\partial \mathbf{X}} \right|_0 \delta \mathbf{X}(t)$$

$$+ \beta[\delta \mathbf{X}(t), t]$$

其中 $\alpha[\delta \mathbf{X}(t), \delta \mathbf{U}(t), t]$ 和 $\beta[\delta \mathbf{X}(t), t]$ 为展开式的高次项。注意到 $\delta \mathbf{X}(t)$ 和 $\delta \mathbf{U}(t)$ 分别具有 n 个和 r 个分量, 因而 $\left. \frac{\partial f}{\partial \mathbf{X}} \right|_0$ 和 $\left. \frac{\partial f}{\partial \mathbf{U}} \right|_0$ 均为函数阵, 而 $\left. \frac{\partial f}{\partial \mathbf{X}} \right|_0$ 和 $\left. \frac{\partial f}{\partial \mathbf{U}} \right|_0$ 表示矩阵在

$\mathbf{X} = \mathbf{X}_0(t), \mathbf{U} = \mathbf{U}_0(t)$ 时的值。令

$$\left. \frac{\partial f}{\partial \mathbf{X}} \right|_0 = \mathbf{A}(t), \left. \frac{\partial f}{\partial \mathbf{U}} \right|_0 = \mathbf{B}(t) \quad \text{及} \quad \left. \frac{\partial g}{\partial \mathbf{X}} \right|_0 = \mathbf{C}(t),$$

略去泰勒展开式中的高次项, 则得实际系统的线性化模型

$$\delta \dot{\mathbf{X}}(t) = \mathbf{A}(t) \delta \mathbf{X}(t) + \mathbf{B}(t) \delta \mathbf{U}(t) \quad (1-15)$$

$$\delta \mathbf{Y}(t) = \mathbf{C}(t) \delta \mathbf{X}(t) \quad (1-16)$$