现代控制理论

中国科学技术大学 自动化系

第二章 系统的状态空间模型

§ 2.1 状态及状态空间

§ 2.2 状态空间方程的建立

§ 2.3 状态空间的线性变换

第二章 系统的状态空间模型

§2.1 状态及状态空间

2.1.1 几个重要的概念

2.1.2 状态空间描述

2.1.3 线性系统

2.1.4 线性定常系统

2.1.5 离散时间系统

2.1.6 分布系统与非线性系统

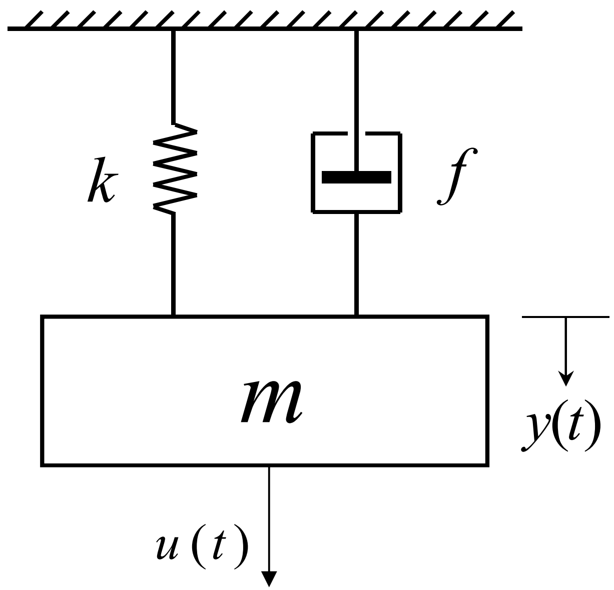
§2.2 状态空间方程的建立

§2.3 状态空间的线性变换

§2.1 状态及状态空间

【例2-1】 如图所示弹簧—阻尼器—质量系统，按经典控制理论知，其输入—输出微分方程描述是：

$$m\frac{d^2y(t)}{dt^2}+f\frac{dy(t)}{dt}+ky(t)=u(t)$$



其传递函数描述是：

$$\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{1}{ms^2+fs+k}$$

对任一时刻t_0，尽管知道端部信息$y(t_0)$、$u(t_0)$及该时刻以后的控制信号$u(t)$，$t > t_0$，我们仍不能了解的该系统以后的运动情况，简言之仍不能确知$y(t)$, $t > t_0$的变化规律：

可见，古典控制论对系统惯用的黑箱研究法，对系统行为的描述是不完整的。于是在现代控制理论中为准确地描述系统(增加研究对象，变黑箱为灰箱)，引入了状态和状态变量的概念。

2.1.1 几个重要的概念

一、输入与输出

把作用于系统的所有（外界）激励统称为系统的输入；而把可以量测的、能表征系统行为的响应统称为系统的输出。

二、状态与状态变量

动态系统的状态可定义为一组信息的集合。在已知未来外部输入的情况下，这些信息对于确定系统未来的行为是充分的。

或：

系统在 t_0 时刻的状态 $\boldsymbol{x}(t_0)$ 是其在 t_0 时刻的一组信息，它与 $t \geq t_0$ 的输入 $\boldsymbol{u}(t)$ 一起，可以唯一确定所有 $t \geq t_0$ 的输出 $\boldsymbol{y}(t)$。

The state $\boldsymbol{x}(t_0)$ *of a system at time* t_0 *is the information at* t_0 *that, together with the input* $\boldsymbol{u}(t)$, *for* $t \geq t_0$, *determines uniquely output* $\boldsymbol{y}(t)$ *for all* $t \geq t_0$.

用于确定动态系统状态的一组变量称作状态变量。

思辨：关于状态的定义

很多教科书上关于状态及状态变量的定义，往往都要强调“**最少量的**一组信息”。如我们的教材上（初版第39页，再版第41页）：

系统 t 时刻状态是系统在 t 时刻所具有的最少量的一组信息量或数据，它和输入$\boldsymbol{u}[t,\infty)$一起唯一地决定着所有 t 以后时间内系统的全部行为，$t\in(-\infty,+\infty)$。

清华大学郑大钟主编，2002年再版《线性系统理论》（第2版）第18页：

一个动力学系统的状态变量组定义为能完全表征其时间域行为的一个最小内部变量组，表为，其中t为自变量时间。

接下来，在第18页最后及第19页开头上对“最小”又作了进一步阐述：

从物理直观上看，定义中“状态变量组为最小”的含义是指，减少其中的一个变量就会破坏它们对系统行为表征的完全性，而增加一个变量将不增加行为表征的信息量，即是完全表征系统行为所不需要的。 ……

从数学的角度看，定义中“状态变量组为最小”的含义是指，它们是系统所有内部变量中线性无关的一个极大变量组，也即以外的系统内部变量都必和它们线性相关。

思辨：关于状态的定义

清华大学吴麒主编的1992年版《自动控制原理》（上册）第25页：

状态变量是这样来定义的，在描述对象运动的所有变量中，必定可以找到**数目最少**的一组变量，它们已经足以描述对象的全部运动。

哈尔滨工业大学段广仁编著，1996年版的《线性系统理论》第47页：

完全表征系统时间域行为的一个最小内部标量称为动力学系统的状态。

……

②状态变量的最小性体现在：状态变量是为完全表征系统行为所必须的系统变量的最小个数，减少变量数将破坏表征的完全性，而增加变量数将是完全表征系统行为所不需要的。

再回忆一下本课程给出的定义——

动态系统的状态可定义为一组信息的集合。在已知未来外部输入的情况下，这些信息对于确定系统未来的行为是充分的。

2.1.1 几个重要的概念

三、状态向量

n 个状态变量依次用 $x_1(t)$，$x_2(t), \ldots, x_n(t)$ 表示，再把这些状态变量看作是向量 $\boldsymbol{x}(t)$ 的分量，则$\boldsymbol{x}(t)$ 就称为状态向量，记作（见右）

$$\boldsymbol{x}(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

四、状态空间

以状态变量用$x_1(t)$，$x_2(t), \ldots, x_n(t)$为坐标轴所为基张成的n维空间，称为状态空间。

在特定的时刻t，状态向量$\boldsymbol{x}(t)$在状态空间中是一点。

五、状态轨线

随着时间的推移，$\boldsymbol{x}(t)$在状态空间中描绘出一条轨迹，称为状态轨线。

第二章 系统的状态空间模型

§2.1 状态及状态空间

2.1.1 几个重要的概念

2.1.2 状态空间描述

2.1.3 线性系统

2.1.4 线性定常系统

2.1.5 离散时间系统

2.1.6 分布系统与非线性系统

§2.2 状态空间方程的建立

§2.3 状态空间的线性变换

2.1.2 状态空间描述

（连续时间系统）

一、状态方程

$$\dot{x} = f(x,u,t) \qquad\qquad \dot{x} = Ax + Bu$$

二、输出方程

$$y = g(x,u,t) \qquad\qquad y = Cx + Du$$

三、状态空间方程（状态空间表达式、动态方程）

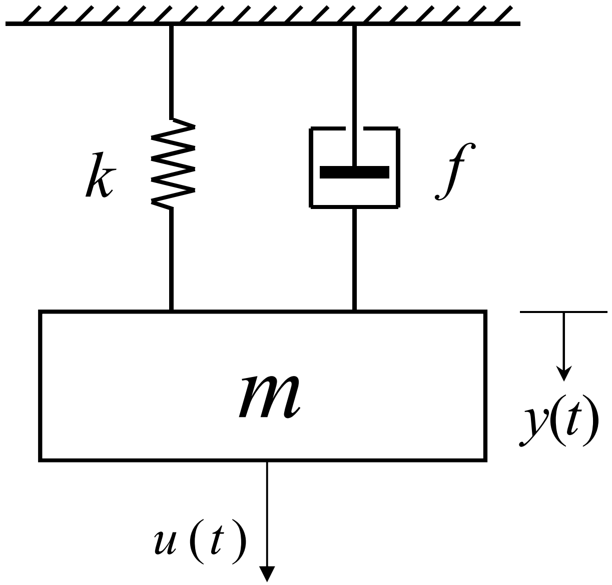
$$\begin{cases} \dot{x} = f(x,u,t) \\ y = g(x,u,t) \end{cases} \qquad\qquad \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

一般形式　　　　　　　　线性定常系统

§2.1 状态及状态空间

【例2-1】 如图所示弹簧—阻尼器—质量系统，按经典控制理论知，其输入—输出微分方程描述是：

$$m\frac{d^2y(t)}{dt^2}+f\frac{dy(t)}{dt}+ky(t)=u(t)$$



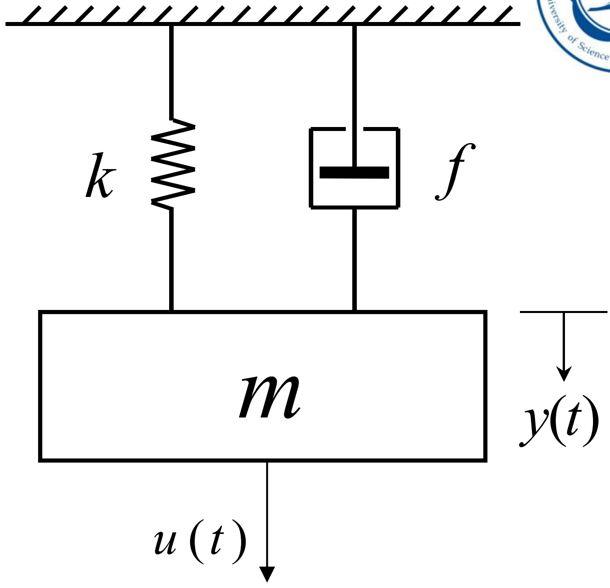
其传递函数描述是：

$$\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{1}{ms^2+fs+k}$$

对任一时刻t_0，尽管知道端部信息$y(t_0)$、$u(t_0)$及该时刻以后的控制信号$u(t)$，$t>t_0$，我们仍不能了解的该系统以后的运动情况，简言之仍不能确知$y(t), t>t_0$的变化规律：

可见，古典控制论对系统惯用的黑箱研究法，对系统行为的描述是不完整的。于是在现代控制理论中为准确地描述系统(增加研究对象，变黑箱为灰箱)，引入了状态和状态变量的概念。

$$m\frac{d^2y(t)}{dt^2}+f\frac{dy(t)}{dt}+ky(t)=u(t)$$



选取状态变量

$$\boldsymbol{x}(t)=\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}=\begin{bmatrix}y(t)\\v(t)\end{bmatrix}$$

则

$$\dot{x}_1=\dot{y}(t)=v(t)=x_2$$

$$\dot{x}_2=\dot{v}(t)=\ddot{y}(t)$$

$$=-\frac{k}{m}y(t)-\frac{f}{m}\dot{y}(t)+\frac{1}{m}u(t)$$

$$=-\frac{k}{m}x_1(t)-\frac{f}{m}x_2(t)+\frac{1}{m}u(t)$$

$$y=x_1$$

$$\begin{cases}\dot{\boldsymbol{x}}=\begin{bmatrix}0&1\\-\dfrac{k}{m}&-\dfrac{f}{m}\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}0\\\dfrac{1}{m}\end{bmatrix}u\\y=\begin{bmatrix}1&0\end{bmatrix}\boldsymbol{x}\end{cases}$$

第二章 系统的状态空间模型

§2.1 状态及状态空间

2.1.1 几个重要的概念

2.1.2 状态空间描述

2.1.3 线性系统

2.1.4 线性定常系统

2.1.5 离散时间系统

2.1.6 分布系统与非线性系统

§2.2 状态空间方程的建立

§2.3 状态空间的线性变换

2.1.3 线性系统

（连续时间系统）

在系统的状态空间方程中，若状态方程和输出方程均为线性方程，即$\dot{\boldsymbol{x}}$ 和 $\boldsymbol{y}$ 均为$\boldsymbol{x}$，$\boldsymbol{u}$的线性函数，则称系统为线性系统

一、线性系统状态空间方程的系数矩阵

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{u} \\ \boldsymbol{y} = \boldsymbol{C}(t)\boldsymbol{x} + \boldsymbol{D}(t)\boldsymbol{u} \end{cases}$$

$$\dim \boldsymbol{x} = n$$

$$\dim \boldsymbol{u} = p$$

$$\dim \boldsymbol{y} = q$$

$\boldsymbol{A} = (a_{ij})_{n\times n}$ —— 系统矩阵（状态矩阵）

$\boldsymbol{B} = (b_{ij})_{n\times p}$ —— 控制矩阵（输入矩阵）

$\boldsymbol{C} = (c_{ij})_{q\times n}$ —— 观测矩阵（输出矩阵）

$\boldsymbol{D} = (d_{ij})_{q\times p}$ —— 输入-输出矩阵

当系数矩阵确定后，系统的状态空间方程也就完全确定了，故为简便起见，有时直呼该系统为系统$\{\boldsymbol{A}(t), \boldsymbol{B}(t), \boldsymbol{C}(t), \boldsymbol{D}(t)\}$

2.1.3 线性系统

二、叠加原理

满足叠加原理是线性系统的一个基本特征，可作为线性系统的定义：

系统被称作为**线性系统**是指对每一个 t_0 及任意两个状态-输入一输出组：

$$\left.\begin{array}{l} \boldsymbol{x}_i(t_0) \\ \boldsymbol{u}_i(t), \quad t \ge t_0 \end{array}\right\} \to y_i(t), t \ge t_0 \qquad\qquad i = 1,2$$

一定有（α为任意实常数 ）

$$\left.\begin{array}{l} \boldsymbol{x}_1(t_0) + \boldsymbol{x}_2(t_0) \\ \boldsymbol{u}_1(t) + \boldsymbol{u}_2(t) \quad, t \ge t_0 \end{array}\right\} \to \boldsymbol{y}_1(t) + \boldsymbol{y}_2(t), \quad t \ge t_0$$

$$\left.\begin{array}{l} \alpha \boldsymbol{x}_1(t_0) \\ \alpha \boldsymbol{u}_1(t_0) \quad t \ge t_0 \end{array}\right\} \to \alpha \boldsymbol{y}_1(t) \qquad t \ge t_0$$

可加性 **均匀性**（齐次性）

可加性和均匀性合在一起称为叠加性。

$$\left.\begin{array}{l} \alpha_1 \boldsymbol{x}_1(t_0) + \alpha_2 \boldsymbol{x}_2(t_0) \\ \alpha_1 \boldsymbol{u}_1(t) + \alpha_2 \boldsymbol{u}_2(t), t \ge t_0 \end{array}\right\} \to \alpha_1 \boldsymbol{y}_1(t) + \alpha_2 \boldsymbol{y}_2(t), \quad t \ge t_0$$

2.1.3 线性系统

若输入恒为零，则输出仅由初始状态引起，这样的输出称作零输入响应

$$\left.\begin{array}{l} \boldsymbol{x}(t_0) \\ \boldsymbol{u}(t) \equiv 0, \quad t \ge t_0 \end{array}\right\} \to \boldsymbol{y}_{zi}(t), \quad t \ge t_0$$

若初始状态恒为零，则输出仅由输入引起，这样的输出称作零状态响应

$$\left.\begin{array}{l} \mathrm{x}(t_0) = 0 \\ \mathrm{u}(t), \quad t \ge t_0 \end{array}\right\} \to y_{zs}(t), \quad t \ge t_0$$

根据可加性，线性系统的响应可分解为零输入响应和零状态响应之和

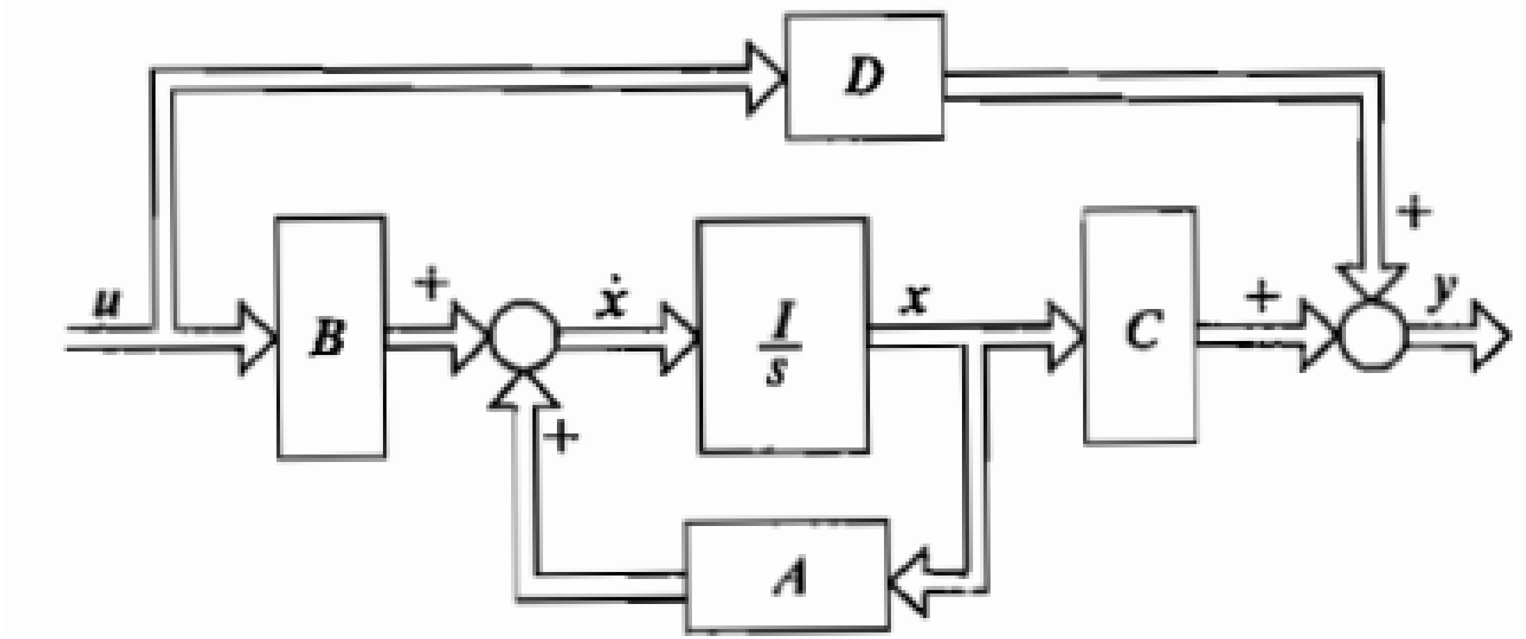
$$\left\{\begin{array}{l} \boldsymbol{x}(t_0) \\ \boldsymbol{u}(t), t \ge t_0 \end{array}\right. \text{的输出} = \left\{\begin{array}{l} \boldsymbol{x}(t_0) \\ \boldsymbol{u}(t) \equiv 0, t \ge t_0 \end{array}\right. \text{的输出} + \left\{\begin{array}{l} \boldsymbol{x}(t_0) = 0 \\ \boldsymbol{u}(t), t \ge t_0 \end{array}\right. \text{的输出}$$

响应＝零输入响应＋零状态响应

2.1.3 线性系统

（连续时间系统）

三、系统结构图

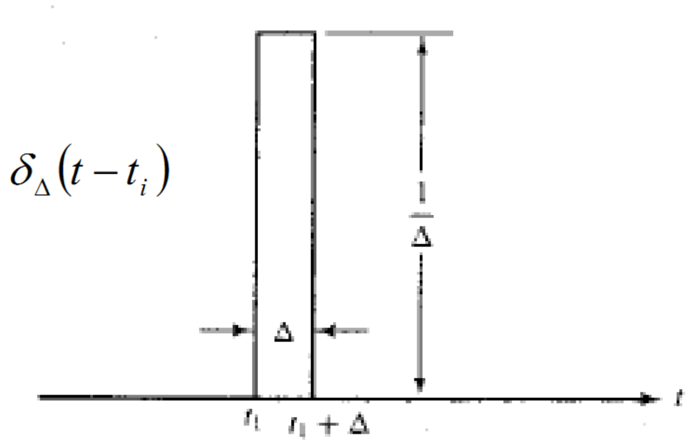


$$G(s) = C(sI - A)^{-1}B + D$$

四、线性系统的输入—输出描述

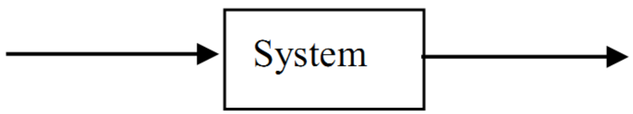
- 脉冲输入响应

$$\delta_\Delta(t-t_i) \qquad y_i(t)=g_\Delta(t,\ t_i)$$



  - (1) 齐次性:

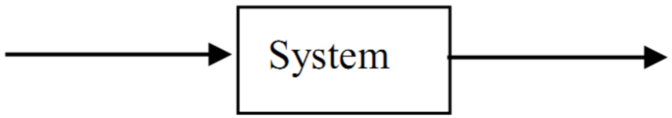
$$(u(t_i)\Delta)\times\delta_\Delta(t-t_i) \qquad (u(t_i)\Delta)\times g_\Delta(t,t_i)$$



i=1,2,3,…

- (2) 可加性:

$$\sum_i \left(u(t_i)\Delta\right)\times\delta_\Delta(t-t_i) \qquad\qquad \sum_i \left(u(t_i)\Delta\right)\times g_\Delta(t,t_i)$$



i=…, -3, -2, -1, 0,1,2,3,…

- (3) 极限

$$\lim_{\Delta\to 0^+}\delta_\Delta(t-t_i)=\delta(t-t_i),\quad \lim_{\Delta\to 0^+}\sum_i \left(u(t_i)\Delta\right)\times\delta_\Delta(t-t_i)=u(t),\quad \lim_{\Delta\to 0^+} g_\Delta(t,t_i)=g(t,t_i)$$

$$\Downarrow$$

$$\lim_{\Delta\to 0^+}\sum_i \left(u(t_i)\Delta\right)\times g_\Delta(t,t_i)=y(t)=\int_{-\infty}^{\infty}u(\tau)g(t,\tau)d\tau$$

2.1.3 线性系统

（连续时间系统）

四、线性系统的输入—输出描述

$$y(t)=\int_{-\infty}^{+\infty} g(t,\tau)\,u(\tau)\,d\tau$$

其中 $g(t,\tau)$ 是一个二元函数，值为τ时刻作用于系统的理想脉冲（输入）所引起系统在 t 时刻的响应（输出），称之为**系统的脉冲响应**。

若系统是**因果的**，则输出不可能在施加输入之前表现出来，于是有：

$$\text{因果的} \Leftrightarrow g(t,\tau)=0 \qquad \forall t<\tau$$

系统在t_0时刻是**松弛的**，是指其t_0前的输入为0。于是t_0时刻松弛的因果系统

$$y(t)=\int_{t_0}^{t} g(t,\tau)\,u(\tau)\,d\tau \qquad\qquad \boldsymbol{y}(t)=\int_{t_0}^{t} \boldsymbol{G}(t,\tau)\boldsymbol{u}(\tau)\,d\tau$$

第二章 系统的状态空间模型

§2.1 状态及状态空间

2.1.1 几个重要的概念

2.1.2 状态空间描述

2.1.3 线性系统

2.1.4 线性定常系统

2.1.5 离散时间系统

2.1.6 分布系统与非线性系统

§2.2 状态空间方程的建立

§2.3 状态空间的线性变换

2.1.4 线性定常系统

一、时间位移特性

系统被称作**定常的**（非时变的、时不变的）是指

$$\left.\begin{aligned} \boldsymbol{x}(t_0) &= \boldsymbol{x}_0 \\ \boldsymbol{u}(t), t &\ge t_0 \end{aligned}\right\} \rightarrow \boldsymbol{y}(t), \quad t \ge t_0$$

任意常数T

$$\left.\begin{aligned} \boldsymbol{x}(t_0 + T) &= \boldsymbol{x}_0 \\ \boldsymbol{u}(t-T), t &\ge t_0 + T \end{aligned}\right\} \rightarrow y(t-T), \quad t \ge t_0 + T$$

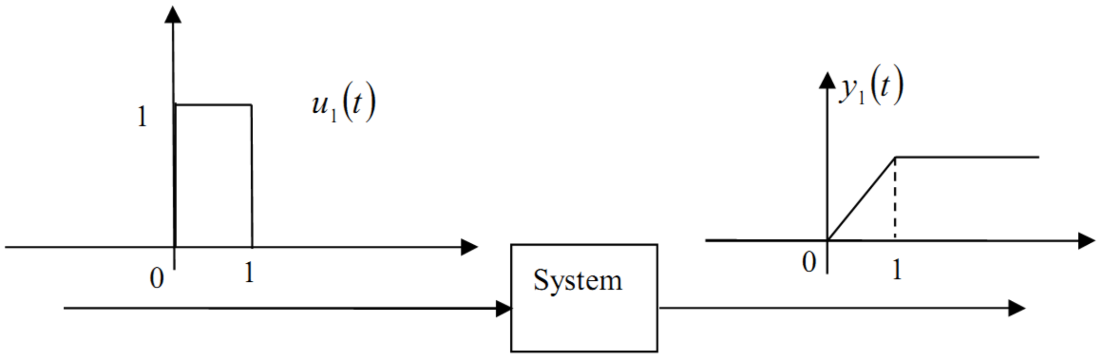
$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u} \end{cases} \qquad \boldsymbol{y}(t) = \int_{t_0}^{t} \boldsymbol{G}(t-\tau)\boldsymbol{u}(\tau)d\tau$$

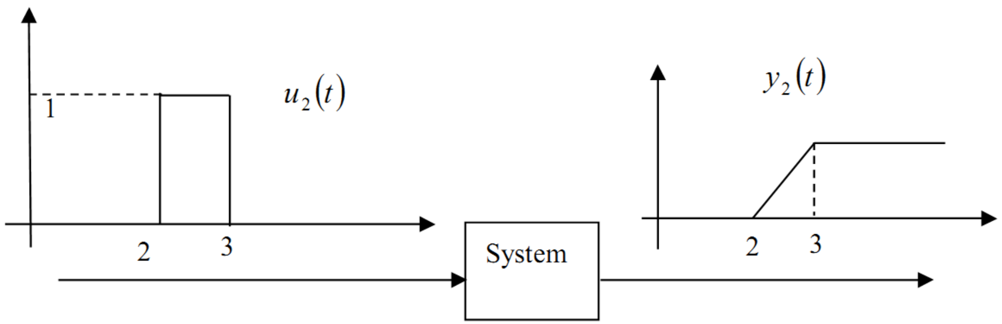
因有时间位移特性。定常系统的初始时刻通常选为零

如果系统不是定常的，则称之为**时变的**

time-invariant ***time-variant***

定常系统举例





能够看到 $u_2(t)=u_1(t-2)$ $y_2(t)=y_1(t-2)$

线性定常系统脉冲响应

- 在2个不同输入下线性定常系统的响应

$$y(t)=\int_{-\infty}^{\infty}u(\tau)g(t,\tau)d\tau$$

$$y'(t)=\int_{-\infty}^{\infty}u'(\tau)g(t,\tau)d\tau$$

- 已知 $u'(t)=u(t-T) \quad y'(t)=y(t-T)$
- (1) $y'(t)=\int_{-\infty}^{\infty}u'(\tau)g(t,\tau)d\tau=\int_{-\infty}^{\infty}u(\tau-T)g(t,\tau)d\tau=\int_{-\infty}^{\infty}u(\tau)g(t,\tau+T)d\tau$
- (2) $y'(t)=y(t-T)==\int_{-\infty}^{\infty}u(\tau)g(t-T,\tau)d\tau$
- 因此$g(t,\tau+T)=g(t-T,\tau), \quad \forall t,\forall\tau,\forall T$

$$\tau=0,T=\tau \Longrightarrow g(t,\tau)=g(t-\tau,0)=g(t-\tau)$$

$$y(t)=\int_{-\infty}^{\infty}u(\tau)g(t-\tau)d\tau$$

2.1.4 线性定常系统

二、连续时间系统的传递函数矩阵

对单输入-单输出的线性定常系统，传递函数定义为：**零初始条件下输出信号的拉氏变换与输入信号的拉氏变换之比。**

对多输入-多输出的线性定常系统，也可给出类似的定义：

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u} \end{cases} \xrightarrow{\text{零初始条件}} \hat{\boldsymbol{y}}(s) = \boldsymbol{G}(s)\hat{\boldsymbol{u}}(s)$$

$$s\,\hat{\boldsymbol{x}}(s) - \boldsymbol{x}(0) = \boldsymbol{A}\hat{\boldsymbol{x}}(s) + \boldsymbol{B}\hat{\boldsymbol{u}}(s)$$

$$\hat{\boldsymbol{y}}(s) = \boldsymbol{C}\,\hat{\boldsymbol{x}}(s) + \boldsymbol{D}\hat{\boldsymbol{u}}(s)$$

$$\hat{\boldsymbol{y}}(s) = [\boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}]\hat{\boldsymbol{u}}(s)$$

$$\boldsymbol{G}(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}$$

第二章 系统的状态空间模型

§2.1 状态及状态空间

- 2.1.1 几个重要的概念
- 2.1.2 状态空间描述
- 2.1.3 线性系统
- 2.1.4 线性定常系统
- 2.1.5 离散时间系统
- 2.1.6 分布系统与非线性系统

§2.2 状态空间方程的建立

§2.3 状态空间的线性变换

2.1.5 离散时间系统

工程系统因计算机控制的需要，时间离散化

非工程系统，根据需要建模时主动将时间离散化

其它一些本质上就是离散时间系统的情况

《计算机控制》课程中重点讨论离散时间系统

同连续时间系统一样，系统的激励$\boldsymbol{u}$和响应$\boldsymbol{y}$分别称为系统的输入和输出，$\boldsymbol{x}$为系统的状态变量；系统状态方程是一阶差分向量方程，输出方程是一阶代数方程；状态方程与输出方程合称为系统的状态空间方程，若系统还是定常的，则可写作：

$$\begin{aligned} \mathrm{x}[k+1] &= \mathrm{Fx}[k]+\mathrm{Hu}[k] \\ \mathrm{y}[k] &= \mathrm{Cx}[k]+\mathrm{Du}[k] \end{aligned}$$

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{Ax}+\boldsymbol{Bu} \\ \boldsymbol{y} = \boldsymbol{Cx}+\boldsymbol{Du} \end{cases}$$

$$\boldsymbol{u}[k] := \boldsymbol{u}(kT),\ \boldsymbol{y}[k] := \boldsymbol{y}(kT)$$

$$\boldsymbol{G}(z) = \boldsymbol{C}(z\boldsymbol{I}-\boldsymbol{F})^{-1}\boldsymbol{H}+\boldsymbol{D} \qquad \boldsymbol{G}(s) = \boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}$$

2.1.6 分布系统与非线性系统

一、集总系统与分布系统

系统被称为**集总的**是指系统状态变量的数目是有限的，或者说它的状态是有限维向量。而系统被称作**分布系统**是指它有无限多个状态变量。

传输线、柔性梁都是著名的分布系统。

【例2-2】考虑一个单位时间延迟系统，定义为： $y(t)=u(t-1)$

【例2-3】考虑例2-2 中研究的单位延迟系统，其脉冲传递函数是 当前无法显示此图像。 ，所以系统的传递函数为

$$\hat{g}(s)=L[\delta(t-1)]=\int_0^{\infty}\delta(t-1)e^{-st}dt=e^{-st}\Big|_{t=1}=e^{-s}$$

分布系统的传递函数为s的无理函数；

线性定常系统是集总的，其传递函数是s的有理函数。

2.1.6 分布系统与非线性系统

二、非线性系统

不满足叠加原理的系统就是非线性系统，集总的非线性系统的状态空间方程一般只能写作左下形式；若状态方程及输出方程中不显含时间t则又可称之为自治系统，如右下

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},t) \\ \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x},\boldsymbol{u},t) \end{cases} \qquad \begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) \\ \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x},\boldsymbol{u}) \end{cases}$$

非线性系统也有定常的、时变的，集总的、分布的，连续的、离散的

同样，离散时间系统也有定常的、时变的，集总的、分布的，线性的、非线性的，等等

第二章 系统的状态空间模型

§2.1 状态及状态空间

§2.2 状态空间方程的建立

2.2.1 从系统的机理出发

2.2.2 自经典模型转换

§2.3 状态空间的线性变换

第二章 系统的状态空间模型

§2.1 状态及状态空间

§2.2 状态空间方程的建立

2.2.1 从系统的机理出发

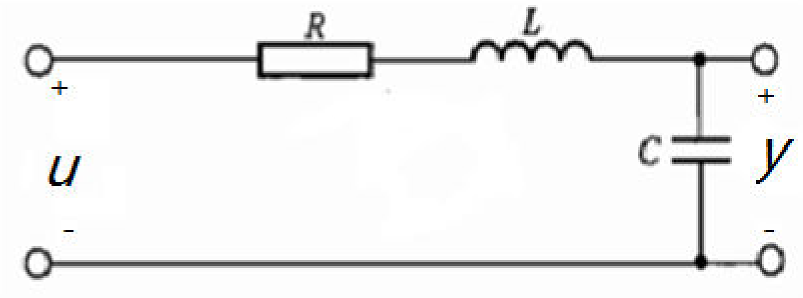
2.2.2 自经典模型转换

§2.3 状态空间的线性变换

§ 2.2 状态空间方程的建立

2.2.1 从系统的机理出发

【例2.4】如图所 RLC ，以左 口 励 动势为 入，右 口 响应 压为 出列写如图 态 方



元件 性： $v_L = L\frac{di_L}{dt}$， $i_C = C\frac{dv_C}{dt}$

KVL： $u = Ri_L + L\frac{di_L}{dt} + v_C$

取 态：

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

$$\begin{cases} \dot{\boldsymbol{x}} = \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix}\boldsymbol{x} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}u \\ y = \begin{bmatrix} 0 & 1 \end{bmatrix}\boldsymbol{x} \end{cases}$$

2.2.1 从系统的机理出发

【例2.5】试建立一个两年制中专学校学生学籍管理动态模型。设：录取新生未报到的比率为1%，在读的一年级学生的留级率为5%、退学率为2%、提前毕业率为1%，二年级学生的留级率为7%、退学率为3%。

解：以每学年开始时为建模的时间节点，记x_1为一年级学生数，x_2为二年级的学生数，输入$u[k]$为当年招生录取的新生人数，输出$y_1[k]$为下一年毕业人数、$y_2[k]$为下一年非正常离校的人数。显然

$$
\begin{aligned}
x_1[k+1] &= 0.05\times x_1[k]+0.99\times u[k]\\
x_2[k+1] &= 0.92\times x_1[k]+0.07\times x_2[k]\\
y_1[k] &= 0.01\times x_1[k]+0.90\times x_2[k]\\
y_2[k] &= 0.02\times x_1[k]+0.03\times x_2[k]+0.01\times u[k]
\end{aligned}
$$

写成矩阵形式：

$$\boldsymbol{x}[k+1]=\begin{bmatrix}0.05 & 0\\ 0.92 & 0.07\end{bmatrix}\boldsymbol{x}[k]+\begin{bmatrix}0.99\\ 0\end{bmatrix}u[k]$$

$$\boldsymbol{y}[k]=\begin{bmatrix}0.01 & 0.90\\ 0.02 & 0.03\end{bmatrix}\boldsymbol{x}[k]+\begin{bmatrix}0\\ 0.01\end{bmatrix}u[k]$$

第二章 系统的状态空间模型

§2.1 状态及状态空间

§2.2 状态空间方程的建立

2.2.1 从系统的机理出发

2.2.2 自经典模型转换

§2.3 状态空间的线性变换

2.2.2 自经典模型转换

一、由传递函数求状态空间方程

含： 入 出微分方 、 性 数学 型

【例2.6】列写如下四 性定常单 入-单 出 态 方

$$y^{(4)}(t)+\alpha_1\dddot{y}(t)+\alpha_2\ddot{y}(t)+\alpha_3\dot{y}(t)+\alpha_4 y(t)$$

$$=b_0u^{(4)}(t)+b_1\dddot{u}(t)+b_2\ddot{u}(t)+b_3\dot{u}(t)+b_4u(t)$$

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{b_0s^4+b_1s^3+b_2s^2+b_3s+b_4}{s^4+\alpha_1s^3+\alpha_2s^2+\alpha_3s+\alpha_4}=\frac{\beta_1s^3+\beta_2s^2+\beta_3s+\beta_4}{s^4+\alpha_1s^3+\alpha_2s^2+\alpha_3s+\alpha_4}+\lambda$$

一、由传递函数求状态空间方程

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}+\lambda$$

解： 引入一个新变量 $v(t)$，它的拉氏变换式定义为

$$\hat{v}(s)=\frac{1}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}\hat{u}(s)$$

即 $$(s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4)\hat{v}(s)=\hat{u}(s)$$

于是有 $$\hat{y}(s)=(\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4)\hat{v}(s)+\lambda\hat{u}(s)$$

定义状态变量为

$$\boldsymbol{x}(t):=\begin{bmatrix}x_1(t)\\x_2(t)\\x_3(t)\\x_4(t)\end{bmatrix}:=\begin{bmatrix}v(t)\\\dot{v}(t)\\\ddot{v}(t)\\\dddot{v}(t)\end{bmatrix}\qquad \text{即}\qquad \hat{\boldsymbol{x}}(s):=\begin{bmatrix}\hat{x}_1(s)\\\hat{x}_2(s)\\\hat{x}_3(s)\\\hat{x}_4(s)\end{bmatrix}=\begin{bmatrix}\hat{v}(s)\\s\hat{v}(s)\\s^2\hat{v}(s)\\s^3\hat{v}(s)\end{bmatrix}=\begin{bmatrix}1\\s\\s^2\\s^3\end{bmatrix}\hat{v}(s)$$

显然 $$\dot{x}_1=x_2,\quad \dot{x}_2=x_3,\quad \dot{x}_3=x_4$$

一、由传递函数求状态空间方程

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}+\lambda$$

$$(s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4)\hat{v}(s)=\hat{u}(s)$$

$$s\hat{x}_4(s)=-\alpha_1\hat{x}_4-\alpha_2\hat{x}_3-\alpha_3\hat{x}_2-\alpha_4\hat{x}_1+\hat{u}(s)$$

$$\hat{y}(s)=(\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4)\hat{v}(s)+\lambda\hat{u}(s)$$

$$\hat{y}(s)=\beta_1\hat{x}_4(s)+\beta_2\hat{x}_3(s)+\beta_3\hat{x}_2(s)+\beta_4\hat{x}_1(s)+\lambda\hat{u}(s)$$

$$\boldsymbol{x}(t):=\begin{bmatrix}x_1(t)\\x_2(t)\\x_3(t)\\x_4(t)\end{bmatrix}:=\begin{bmatrix}v(t)\\\dot{v}(t)\\\ddot{v}(t)\\\dddot{v}(t)\end{bmatrix}\quad \text{即}\quad \hat{\boldsymbol{x}}(s):=\begin{bmatrix}\hat{x}_1(s)\\\hat{x}_2(s)\\\hat{x}_3(s)\\\hat{x}_4(s)\end{bmatrix}=\begin{bmatrix}\hat{v}(s)\\s\hat{v}(s)\\s^2\hat{v}(s)\\s^3\hat{v}(s)\end{bmatrix}=\begin{bmatrix}1\\s\\s^2\\s^3\end{bmatrix}\hat{v}(s)$$

$$\dot{x}_1=x_2,\quad \dot{x}_2=x_3,\quad \dot{x}_3=x_4$$

$$\dot{x}_4(t)=[-\alpha_4\quad -\alpha_3\quad -\alpha_2\quad -\alpha_1]\boldsymbol{x}(t)+1\cdot u(t)$$

$$y(t)=[\beta_4\quad \beta_3\quad \beta_2\quad \beta_1]\boldsymbol{x}(t)+\lambda u(t)$$

一、由传递函数求状态空间方程

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}+\lambda$$

$$\dot{x}_1=x_2,\quad \dot{x}_2=x_3,\quad \dot{x}_3=x_4$$

$$\dot{x}_4(t)=[-\alpha_4\quad -\alpha_3\quad -\alpha_2\quad -\alpha_1]\boldsymbol{x}(t)+1\cdot u(t)$$

$$y(t)=[\beta_4\quad \beta_3\quad \beta_2\quad \beta_1]\boldsymbol{x}(t)+\lambda u(t)$$

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\-\alpha_4&-\alpha_3&-\alpha_2&-\alpha_1\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}0\\0\\0\\1\end{bmatrix}u$$

$$y=\boldsymbol{c}\boldsymbol{x}+d\,u=[\beta_4\quad \beta_3\quad \beta_2\quad \beta_1]\boldsymbol{x}(t)+\lambda\cdot u$$

一、由传递函数求状态空间方程

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}+\lambda$$

注意：在本例中，若定义状态变量为

$$\boldsymbol{x}(t):=\begin{bmatrix} x_1(t)\\ x_2(t)\\ x_3(t)\\ x_4(t)\end{bmatrix}:=\begin{bmatrix} \dddot{v}(t)\\ \ddot{v}(t)\\ \dot{v}(t)\\ v(t)\end{bmatrix} \quad \text{即} \quad \hat{\boldsymbol{x}}(s):=\begin{bmatrix} \hat{x}_1(s)\\ \hat{x}_2(s)\\ \hat{x}_3(s)\\ \hat{x}_4(s)\end{bmatrix}:=\begin{bmatrix} s^3\\ s^2\\ s\\ 1\end{bmatrix}\hat{\boldsymbol{v}}(s)$$

则易导出系统的状态空间方程是

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u=\begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix} 1\\ 0\\ 0\\ 0\end{bmatrix}u$$

$$y=\boldsymbol{c}\boldsymbol{x}+d\,u=[\,\beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4\,]\boldsymbol{x}(t)+\lambda\cdot u$$

一、由传递函数求状态空间方程

$$\hat{g}(s)=\frac{\hat{y}(s)}{\hat{u}(s)}=\frac{\beta_1 s^3+\beta_2 s^2+\beta_3 s+\beta_4}{s^4+\alpha_1 s^3+\alpha_2 s^2+\alpha_3 s+\alpha_4}+\lambda$$

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}0\\0\\0\\1\end{bmatrix}u$$

$$y=\boldsymbol{c}\boldsymbol{x}+d\,u=[\begin{matrix}\beta_4 & \beta_3 & \beta_2 & \beta_1\end{matrix}]\boldsymbol{x}(t)+\lambda\cdot u$$

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u=\begin{bmatrix}-\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\0\\0\\0\end{bmatrix}u$$

$$y=\boldsymbol{c}\boldsymbol{x}+d\,u=[\begin{matrix}\beta_1 & \beta_2 & \beta_3 & \beta_4\end{matrix}]\boldsymbol{x}(t)+\lambda\cdot u$$

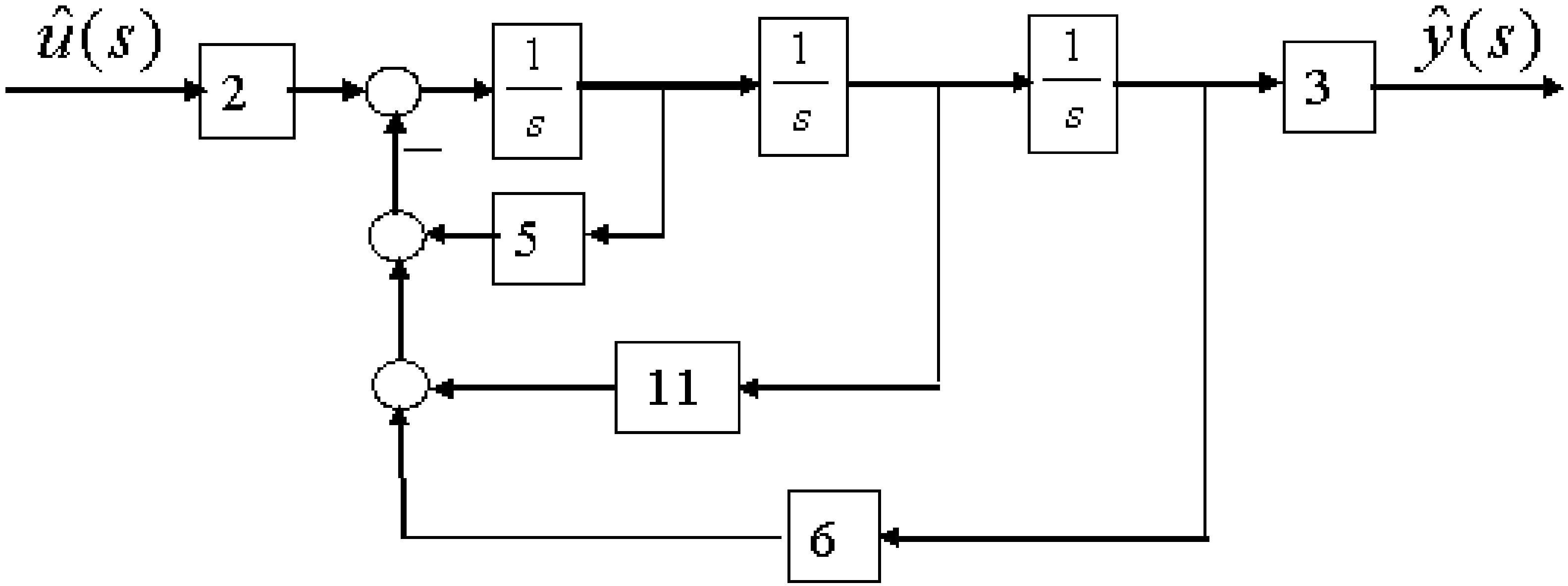
结论：同一系统因状态选择不同，状态空间方程大相径庭

2.2.2 自经典模型转换

一、由传递函数求状态空间方程

二、由动态结构图求状态空间方程

【例2.7】系统的方块图如图所示，求该系统的状态空间方程



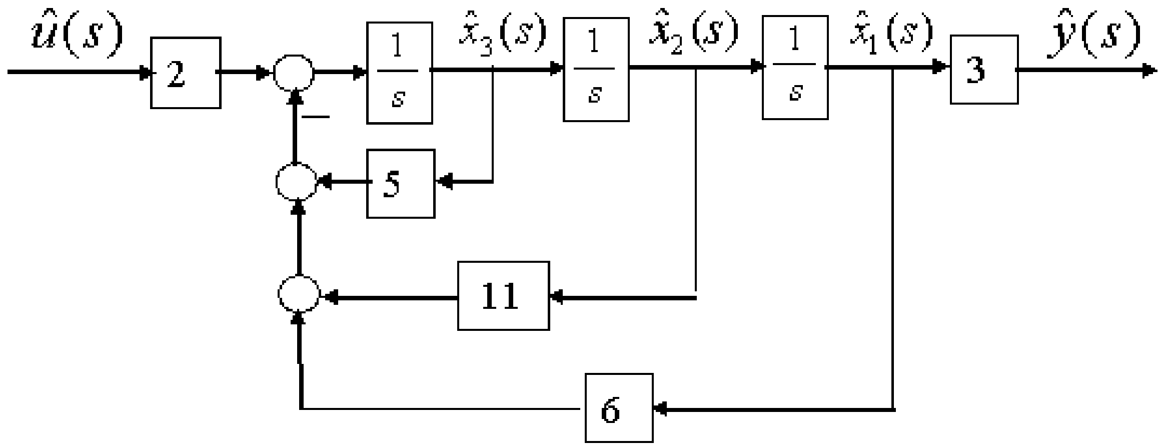
2.2.2 自经典模型转换

二、由动态结构图求状态空间方程

【例2.7】解：

按如图方式选择状态

（积分器输出）



$$s\hat{x}_1(s)=\hat{x}_2(s)$$

$$s\hat{x}_2(s)=\hat{x}_3(s)$$

$$s\hat{x}_3(s)=2\hat{u}(s)-[6\hat{x}_1(s)+11\hat{x}_2(s)+5\hat{x}_3(s)]$$

$$\hat{y}(s)=3\hat{x}_1(s)$$

$$\dot{x}_1=x_2$$

$$\dot{x}_2=x_3$$

$$\dot{x}_3=-6x_1-11x_2-5x_3+2u$$

$$y=3x_1$$

$$\dot{\boldsymbol{x}}=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -6 & -11 & -5\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}0\\ 0\\ 2\end{bmatrix}u$$

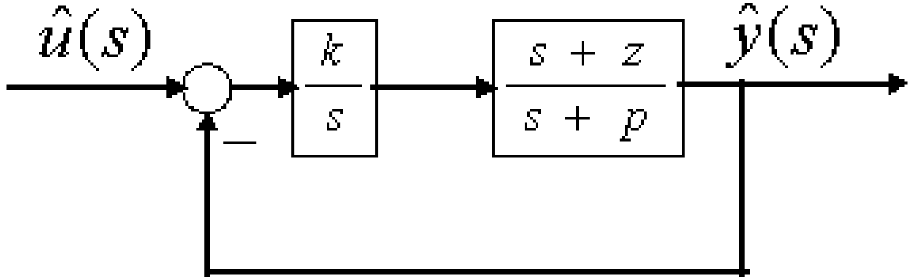
$$y=\begin{bmatrix}3 & 0 & 0\end{bmatrix}\boldsymbol{x}$$

2.2.2 自经典模型转换

二、由动态结构图求状态空间方程

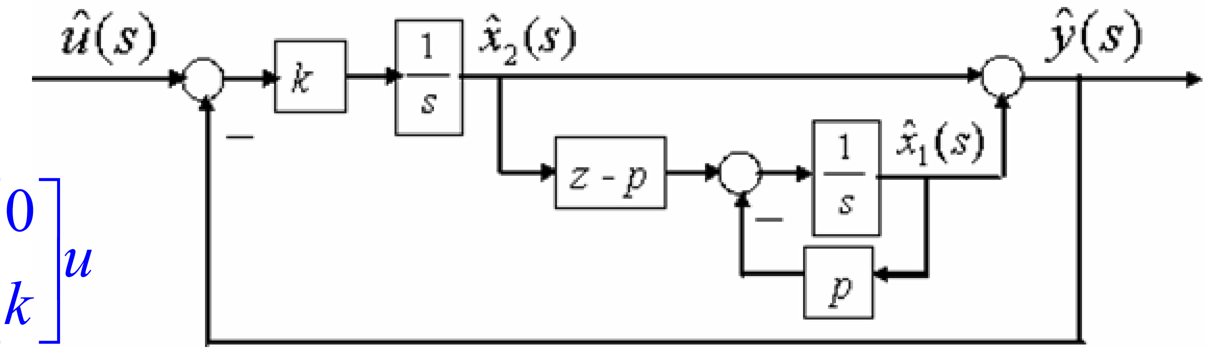
【例2.8】系统的方块图如图所示，求该系统的状态空间方程

解：注意到



$$\frac{s+z}{s+p}=1+\frac{z-p}{s+p}=1+(z-p)\frac{1/s}{1+1/s\cdot p}$$

如下图所标定义状态变量可得

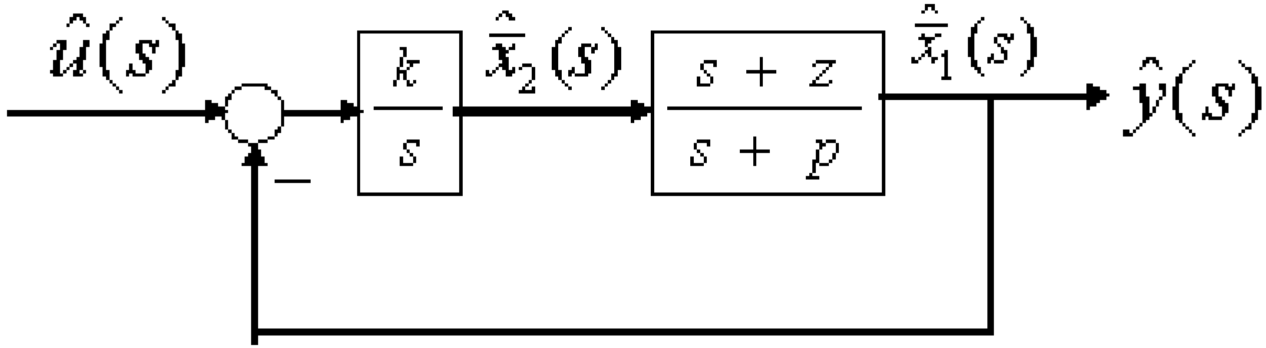


$$\dot{\boldsymbol{x}}=\begin{bmatrix}-p & z-p\\ -k & -k\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}0\\ k\end{bmatrix}u$$

$$y=\begin{bmatrix}1 & 1\end{bmatrix}\boldsymbol{x}$$

二、由动态结构图求状态空间方程

【例2.8】另一方面，按如下所标选择状态变量，可直接得到



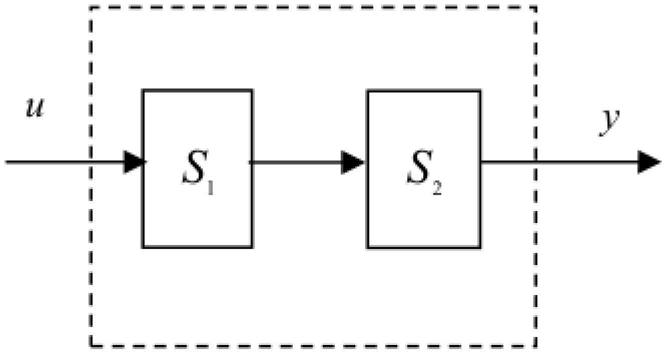
$$\hat{x}_1(s)=\frac{s+z}{s+p}\hat{x}_2(s)\,,\qquad \hat{x}_2(s)=\frac{k}{s}\left(\hat{u}(s)-\hat{x}_1(s)\right),\qquad y(s)=\hat{x}_1(s)$$

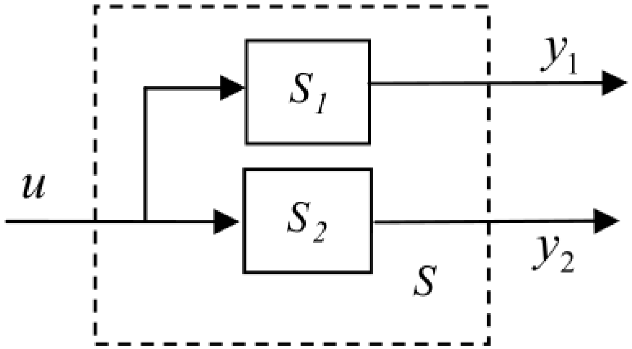
$$\dot{\boldsymbol{x}}=\begin{bmatrix}-p-k & z\\ -k & 0\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}k\\ k\end{bmatrix}u \qquad \dot{\boldsymbol{x}}=\begin{bmatrix}-p & z-p\\ -k & -k\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}0\\ k\end{bmatrix}u$$

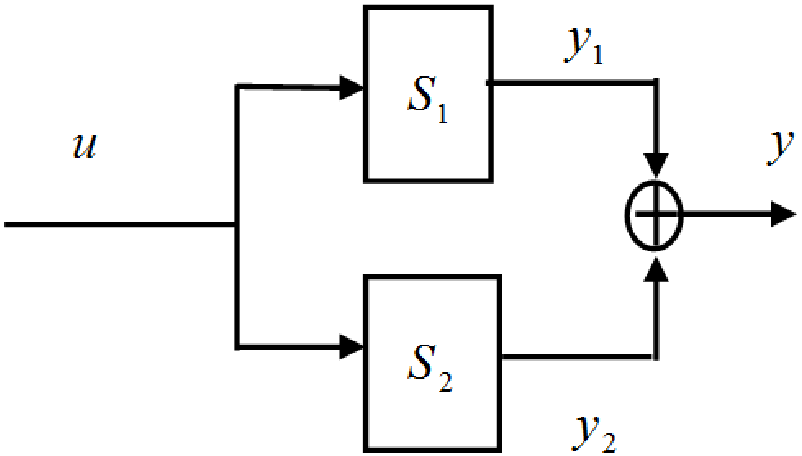
$$y=\begin{bmatrix}1 & 0\end{bmatrix}\boldsymbol{x} \qquad y=\begin{bmatrix}1 & 1\end{bmatrix}\boldsymbol{x}$$

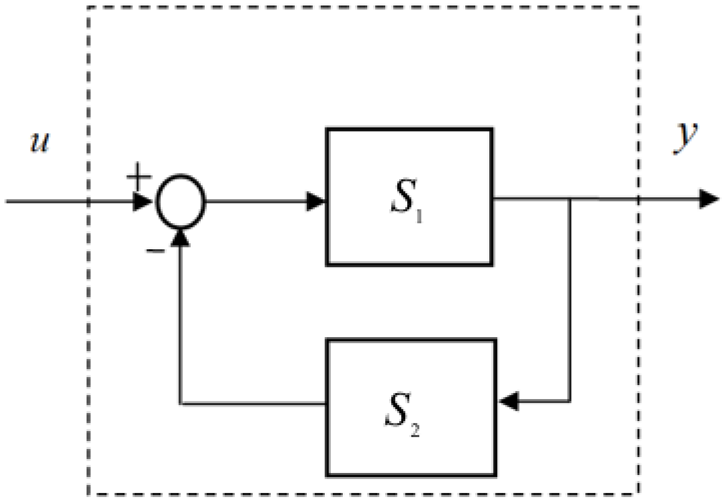
再次注意到：同一个系统从两条道路（状态变量选择不同）得到了两个不同的状态空间方程。

- 互连线性系统的总体状态方程









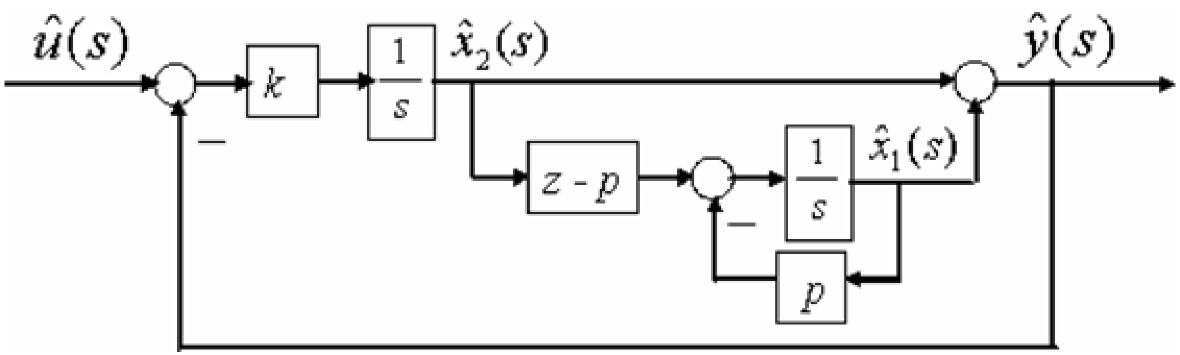
第二章 系统的状态空间模型

§2.1 状态及状态空间

§2.2 状态空间方程的建立

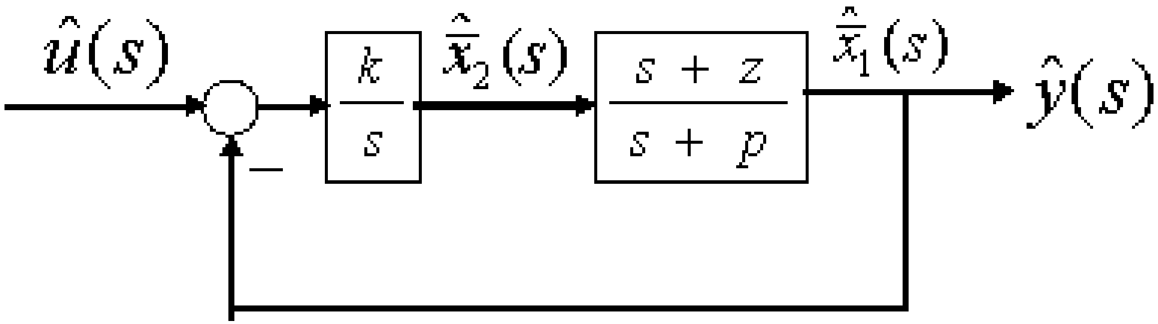
§2.3 状态空间的线性变换

§2.3 状态空间的线性变换



$$\dot{\boldsymbol{x}} = \begin{bmatrix} -p & z-p \\ -k & -k \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ k \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} \boldsymbol{x}$$



$$\dot{\bar{\boldsymbol{x}}} = \begin{bmatrix} -p-k & z \\ -k & 0 \end{bmatrix} \bar{\boldsymbol{x}} + \begin{bmatrix} k \\ k \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \bar{\boldsymbol{x}}$$

$$\begin{cases} x_1 = \bar{x}_1 - \bar{x}_2 \\ x_2 = \bar{x}_2 \end{cases} \qquad \begin{cases} \bar{x}_1 = x_1 + x_2 \\ \bar{x}_2 = x_2 \end{cases} \qquad \boldsymbol{x} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \bar{\boldsymbol{x}}, \qquad \bar{\boldsymbol{x}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}$$

2.3.1 坐标变换对状态空间方程的影响

$$\dot{x} = Ax + Bu \qquad y = Cx + Du$$

$$\xrightarrow[\bar{x} = Px = Q^{-1}x]{x = Q\bar{x}}$$

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u \qquad y = \bar{C}\bar{x} + \bar{D}u$$

$$Q\dot{\bar{x}} = AQ\bar{x} + Bu \qquad \dot{\bar{x}} = Q^{-1}AQ\,\bar{x} + Q^{-1}Bu$$

$$\bar{A} = Q^{-1}AQ \qquad \bar{B} = Q^{-1}B$$

$$\bar{C} = CQ \qquad \bar{D} = D$$

$$P = Q^{-1}$$

$$\bar{A} = PAP^{-1} \qquad \bar{B} = PB$$

$$\bar{C} = CP^{-1} \qquad \bar{D} = D$$

2.3.2 矩阵的规范化

1．特征方程、特征值及特征向量

$$\boldsymbol{A}\boldsymbol{q} = \lambda \boldsymbol{q}$$

称λ为n阶方阵$\boldsymbol{A}$的特征值，而称$\boldsymbol{q}$为$\boldsymbol{A}$对应于特征值λ的特征向量

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{q} = \boldsymbol{0} \qquad \left|\lambda \boldsymbol{I} - \boldsymbol{A}\right| = 0$$

右等式左边是一关于λ的n次首一多项式，由它可解出A的所有特征值

$$\boldsymbol{A}\boldsymbol{q}_i = \lambda_i \boldsymbol{q}_i \;,(i = 1,2,\cdots,n)$$

2．方阵的对角规范化

分四种情况讨论

（1）、单重特征值、单个特征向量

特征方程： $|sI-A| = (s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)$

n个相异特征值：$\lambda_i \neq \lambda_j$

特征向量： $\lambda_i \rightarrow Aq_i = \lambda_i q_i$

n个线性无关的特征向量： $\{q_1,\quad q_2,\quad \cdots,\quad q_n\}$

证明线性无关性：假设不是线性无关，存在不全为0的 $c_i(i=1,\cdots,n)$

$$c_1q_1 + c_2q_2 + \cdots + c_nq_n = 0$$

$$\lambda_1c_1q_1 + \lambda_2c_2q_2 + \cdots + \lambda_nc_nq_n = 0$$

$$\vdots$$

$$\lambda_1^{n-1}c_1q_1 + \lambda_2^{n-1}c_2q_2 + \cdots + \lambda_n^{n-1}c_nq_n = 0$$

$$
\begin{aligned}
c_1q_1+c_2q_2+\cdots+c_nq_n&=0\\
\lambda_1c_1q_1+\lambda_2c_2q_2+\cdots+\lambda_nc_nq_n&=0\\
&\vdots\\
\lambda_1^{\,n-1}c_1q_1+\lambda_2^{\,n-1}c_2q_2+\cdots+\lambda_n^{\,n-1}c_nq_n&=0
\end{aligned}
$$

$$\Downarrow$$

$$
\begin{bmatrix} c_1q_1 & c_2q_2 & \cdots & c_nq_n \end{bmatrix}
\begin{bmatrix}
1 & \lambda_1 & \cdots & \lambda_1^{\,n-1}\\
1 & \lambda_2 & \cdots & \lambda_2^{\,n-1}\\
\vdots & \vdots & \ddots & \vdots\\
1 & \lambda_n & \cdots & \lambda_n^{\,n-1}
\end{bmatrix} = 0
$$

$$\Downarrow$$

$$\begin{bmatrix} c_1q_1 & c_2q_2 & \cdots & c_nq_n \end{bmatrix}=0$$

$$\Downarrow$$

$$c_1=c_2=\cdots c_n=0$$

矛盾！

方阵的对角规范化

$$A[\boldsymbol{q}_1\ \ \boldsymbol{q}_2 \cdots \boldsymbol{q}_n] = [\boldsymbol{q}_1\ \ \boldsymbol{q}_2 \cdots \boldsymbol{q}_n]\begin{bmatrix} \lambda_1 & & & \boldsymbol{O} \\ & \lambda_2 & & \\ & & \ddots & \\ \boldsymbol{O} & & & \lambda_n \end{bmatrix}$$

显然， $\boldsymbol{A}$的这n个特征向量 $\boldsymbol{q}_i\ (i=1,2,\cdots n)$互不相关，则矩阵 $\boldsymbol{Q}=[\boldsymbol{q}_1\ \ \boldsymbol{q}_2 \cdots \boldsymbol{q}_n]$ 可逆，于是

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \Lambda$$

此即表明，若系统矩阵$\boldsymbol{A}$的所有特征向量互不相关，则可通过状态变换（重新定义等价的状态状态变量），使系统在新的状态空间内，其系统矩阵呈对角型。

①同一方阵对应于不同特征值的特征向量一定互不相关；
②状态变换不改变系统矩阵的特征值；
③方阵$\boldsymbol{A}$非奇异的充要条件是它没有零特征值

（2）、多重特征值、多个特征向量

特征方程： $|sI-A|=(s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2}\cdots(s-\lambda_l)^{m_l}$

m_i：λ_i的代数重数

特征向量： $\lambda_i \to Aq_{i,j}=\lambda_i q_{i,j} \quad (j=1,\cdots,m_i)$

m_i 个线性无关的特征向量： $\{q_{i,1},\ q_{i,2},\ \cdots,\ q_{i,m_i}\}$

所有 $n=m_1+m_2+\cdots+m_l$ 个特征向量线性无关，组成变换矩阵

$$Q=\begin{bmatrix} q_{1,1} & \cdots & q_{1,m_1} & \cdots & q_{l,1} & \cdots & q_{l,m_l} \end{bmatrix}$$

$$Q^{-1}AQ=\Lambda=\begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda_p & \\ & & & & & \ddots & \\ & & & & & & \lambda_p \end{bmatrix}$$

（3）、多重特征值、单个特征向量

特征方程： $|sI-A|=(s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2}\cdots(s-\lambda_l)^{m_l}$

m_i：λ_i的代数重数

仅可求到一个线性无关特征向量：$\lambda_i \rightarrow Aq_i=\lambda_i q_i$

l 个线性无关的特征向量不足以构成变换矩阵 Q

从每个特征向量触发，求取 m_i-1 个增广特征向量。

$$\left(\lambda_i, q_i=q_i^{(1)}\right) \rightarrow \boxed{q_i^{(2)},\cdots,q_i^{(m_i)}}$$

增广特征向量

（相互线性不相关）

$$(A-\lambda_i I)q_i^{(1)}=0,\ (A-\lambda_i I)^2 q_i^{(2)}=0,\cdots,(A-\lambda_i I)^{m_i} q_i^{(m_i)}=0$$

$$(A-\lambda_i I)^k q_i^{(k+m)} \neq 0, \qquad m \geq 1$$

求解次序：$q_i^{(m_i)} \to q_i^{(m_i-1)} \to \cdots \to q_i^{(1)}$

两个齐次方程的解空间：

$$S_{m_i} = \{v \mid (A-\lambda_i I)^{m_i} v = 0\},$$
$$S_{m_i-1} = \{v \mid (A-\lambda_i I)^{m_i-1} v = 0\}$$

空间包含：$S_{m_i} \supset S_{m_i-1}$

最高阶增广特征向量

$$0 \neq q_i^{(m_i)} \in S_{m_i} \setminus S_{m_i-1}$$

其余增广特征向量

$$q_i^{(m_i-1)} = (A-\lambda_i I) q_i^{(m_i)}$$

$(A-\lambda_i I)^k q_i^{(k+m)} \neq 0,\ m \geq 1$

$$q_i^{(m_i-2)} = (A-\lambda_i I) q_i^{(m_i-1)}$$

$\vdots$

$\{q_i^{(m_i)}, q_i^{(m_i-1)}, \cdots, q_i^{(1)}\}$ 线性无关

m_i-阶增广特征向量

$$(A-\lambda_i I)^m q_i^{(m_i)} = 0$$

$$(A-\lambda_i I)^{m-1} q_i^{(m_i)} \neq 0$$

(m_i-1)-阶增广特征向量

$$(A-\lambda_i I)^{m-1}\left((A-\lambda_i I)q_i^{(m_i)}\right)=0$$

$$(A-\lambda_i I)^{m-2}\left((A-\lambda_i I)q_i^{(m_i)}\right)\neq 0$$

(m_i-2)-阶增广特征向量

$$(A-\lambda_i I)^{m-2}\left((A-\lambda_i I)^2 q_i^{(m_i)}\right)=0$$

$$(A-\lambda_i I)^{m-3}\left((A-\lambda_i I)^2 q_i^{(m_i)}\right)\neq 0$$

$$q_i^{(m_i)} \to (A-\lambda_i I)q_i^{(m_i)} \to \cdots \to (A-\lambda_i I)^{m_i-1} q_i^{(m_i)}$$

$$q^{(3)} \rightarrow (A-\lambda I)q^{(3)} \rightarrow (A-\lambda I)^2 q^{(3)}$$ 线性无关

$$(A-\lambda I)^3 q^{(3)} = 0, \quad (A-\lambda I)^2 q^{(3)} \neq 0$$

$$c_3 q^{(3)} + c_2 (A-\lambda I) q^{(3)} + c_1 (A-\lambda_i I)^2 q_i^{(3)} = 0$$

$$\Downarrow$$

$$(A-\lambda I)^2 \left(c_3 q^{(3)} + c_2 (A-\lambda I) q^{(3)} + c_1 (A-\lambda_i I)^2 q_i^{(3)} \right) = 0$$

$$\Downarrow$$

$$c_3 (A-\lambda I)^2 q^{(3)} = 0 \Longrightarrow c_3 = 0$$

$$\Downarrow$$

$$c_2 (A-\lambda I) q^{(3)} + c_1 (A-\lambda_i I)^2 q_i^{(3)} = 0$$

变换矩阵：

$$Q = \left[q_1^{(1)} \quad \cdots \quad q_1^{(m_1)} \quad \cdots \quad q_l^{(1)} \quad \cdots \quad q_l^{(m_l)} \right]$$

线性无关？

$$Q^{-1}AQ = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_l \end{bmatrix}$$

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & \ddots & \\ & & & \ddots & \ddots \\ & & & & \lambda_i \end{bmatrix}_{m_i \times m_i}$$

$$QJ = AQ \leftarrow q_i^{(j-1)} + \lambda_i q_i^{(j)} = A q_i^{(j)}$$

举例：

$$A=\begin{bmatrix}1&1&0\\0&1&2\\0&0&1\end{bmatrix}\qquad |sI-A|=(s-1)^3,\quad \lambda=1, m=3$$

$$rank(A-\lambda I)=2 \quad \Rightarrow \quad (A-\lambda I)q=0$$ 仅1个特征向量

$$(A-\lambda I)^3=0,\ S_3=\{v \mid (A-\lambda I)^3 v=0\}=\{e_1,e_2,e_3\}$$

$$I=[e_1 \quad e_2 \quad e_3]$$

$$(A-\lambda I)^2=\begin{bmatrix}0&0&2\\0&0&0\\0&0&0\end{bmatrix},\ S_2=\{v \mid (A-\lambda I)^2 v=0\}=\{e_1,e_2\}$$

$$S_3 \setminus S_2=\{e_3\},\quad q^{(3)}=e_3$$

一般情况需用Schmidt 正交化计算差集。

$$q^{(2)}=\begin{bmatrix}0\\2\\0\end{bmatrix},q^{(1)}=\begin{bmatrix}2\\0\\0\end{bmatrix},Q=[q^{(1)},q^{(2)},q^{(3)}]=\begin{bmatrix}2&&\\&2&\\&&1\end{bmatrix},J=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$$

3阶约当块

$$J=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}_{3\times 3}$$

$$rank\left((A-\lambda I)^0\right)-rank(A-\lambda I)=1$$

$$rank(A-\lambda I)-rank\left((A-\lambda I)^2\right)=1$$

$$rank\left((A-\lambda I)^2\right)-rank\left((A-\lambda I)^3\right)=1$$

$$rank\left((A-\lambda I)^3\right)-rank\left((A-\lambda I)^4\right)=0$$

$$g=3$$

$$g=\min_k\left\{k \;\middle|\; rank\left((A-\lambda I)^k\right)-rank\left((A-\lambda I)^{k+1}\right)=0\right\}$$

g ： λ 的几何重数

$$rank\left((A-\lambda I)^0\right)-rank(A-\lambda I)=1 \Longrightarrow$$

仅1个特征向量

λ 对应1个约当块

（4）、多重特征值、多个特征向量

特征方程： $\left|sI-A\right|=\left(s-\lambda_1\right)^{m_1}\left(s-\lambda_2\right)^{m_2}\cdots\left(s-\lambda_l\right)^{m_l}$

m_i：λ_i的代数重数

可求到 n_i 个线性无关特征向量：$\lambda_i \rightarrow Aq_i=\lambda_i q_i$

$$n_i = rank\left(\left(A-\lambda_i I\right)^0\right)-rank\left(A-\lambda I\right)$$

有 n_i 组增广特征向量串：　$m_{i,1}+m_{i,2}+\cdots+m_{i,n_i}=m_i$

$q_{i,1}^{(m_{i,1})} \rightarrow q_{i,1}^{(m_{i,1}-1)} \rightarrow \cdots \rightarrow q_{i,1}^{(1)}$ $\Longrightarrow$ $m_{i,1}$ 阶约当块 $J_{i,1}$

$q_{i,2}^{(m_{i,2})} \rightarrow q_{i,2}^{(m_{i,2}-1)} \rightarrow \cdots \rightarrow q_{i,2}^{(1)}$ $\Longrightarrow$ $m_{i,2}$ 阶约当块 $J_{i,2}$

$\vdots$

$q_{i,n_i}^{(m_{i,n_i})} \rightarrow q_{i,n_i}^{(m_{i,n_i}-1)} \rightarrow \cdots \rightarrow q_{i,n_i}^{(1)}$ $\Longrightarrow$ m_{i,n_i} 阶约当块 J_{i,n_i}

$$Q_i = \left[q_{i,1}^{(1)} \quad \cdots \quad q_{i,1}^{(m_{i,1})} \quad \cdots \quad q_{i,n_i}^{(1)} \quad \cdots \quad q_{i,n_i}^{(m_{i,n_i})} \right]$$

变换矩阵： $Q = \left[Q_1 \quad Q_2 \quad \cdots \quad Q_l \right]$

约当矩阵： $J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_l \end{bmatrix}$, $J_i = \begin{bmatrix} J_{i,1} & & \\ & \ddots & \\ & & J_{i,n_i} \end{bmatrix}$

$$Q^{-1} A Q = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_l \end{bmatrix}$$

$$m_{i,j}=?\qquad j=1,\cdots,n_i$$

令 $c_k=rank\left((A-\lambda_i I)^k\right)-rank\left((A-\lambda_i I)^{k+1}\right)$

$$\left|\left\{j \quad | \quad m_{i,j}>k\right\}\right|=c_k$$

$|\{\cdots\}|$: 集合元素的个数

g_i：λ_i的几何重数

$$g_i=\max_j m_{i,j}=\min_k\{k \quad | \quad c_k=0\}$$

计算过程：首先定出 g_i

$$c_0\ge 1\to c_1\ge 1\to\cdots\to c_{g_i-1}\ge 1\to c_{g_i}=0$$

$$\left|\left\{j \quad | \quad m_{i,j}>k\right\}\right|=c_k \Longrightarrow m_{i,j}$$

$$k=0,1,\cdots,g_i$$

举例：

$$A=\begin{bmatrix}1 & 1 & & & \\ & 1 & 2 & & \\ & & 1 & & \\ & & & 1 & \\ & & & 2 & 1\end{bmatrix}$$

$$|sI-A|=(s-1)^5,$$

$$\lambda=1,$$

$$c_0=2\to c_1=2\to c_2=1\to c_3=0$$

$$g_i=\max_j m_{i,j}=\min_k\{k \mid c_k=0\}=3$$

$$|\{j \mid m_{i,j}>2\}|=c_2=1 \Longrightarrow$$ 有一个$m_{i,j}$ 超过2 } 有一个 $m_{i,j}=3$

$$|\{j \mid m_{i,j}>1\}|=c_1=2 \Longrightarrow$$ 有2个 $m_{i,j}$大于等于2

$$|\{j \mid m_{i,j}>0\}|=c_0=2 \Longrightarrow$$ 有2个$m_{i,j}$ 大于等于1

$$\Longrightarrow \quad m_{i,1}=3$$

$$m_{i,2}=2$$

计算增广特征向量链

$$q_{i,1}^{(3)} \rightarrow q_{i,1}^{(2)} \rightarrow q_{i,1}^{(1)} \qquad\qquad q_{i,2}^{(2)} \rightarrow q_{i,2}^{(1)}$$

$$S_{i,m} = \{v \quad | \quad (A-\lambda_i I)^m = 0\}$$

$$q_{i,1}^{(3)} \in S_{i,3} \setminus S_{i,2} \quad \Longrightarrow \quad \begin{aligned} q_{i,1}^{(2)} &= (A-\lambda_i I) q_{i,1}^{(3)} \\ q_{i,1}^{(1)} &= (A-\lambda_i I) q_{i,1}^{(2)} \end{aligned}$$

$$q_{i,2}^{(2)} \in S_{i,2} \setminus S_{i,1} \setminus span\{q_{i,1}^{(2)}\}$$

$$\Longrightarrow \quad q_{i,2}^{(1)} = (A-\lambda_i I) q_{i,2}^{(2)}$$

一般情况：

http://staff.ustc.edu.cn/~qling/download/jordan.pdf

2.3.2 矩阵的规范化

3．方阵的约当规范化

我们来讨论约当标准型的一个有用的性质，并以此来结束本节。考虑前述所述 4 阶约当块 $\boldsymbol{J}$，易见

$$(\boldsymbol{J}-\lambda\boldsymbol{I})=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}\qquad(\boldsymbol{J}-\lambda\boldsymbol{I})^2=\begin{bmatrix}0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{bmatrix}\qquad(\boldsymbol{J}-\lambda\boldsymbol{I})^3=\begin{bmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

且对所有的 $k\geqslant4$，有 $(\boldsymbol{J}-\lambda\boldsymbol{I})^k=0$，这叫做幂零。最左一个矩阵被称为 4 阶幂零矩阵。

通过选择适当的线性变换，将方阵化为约当形的方法和步骤，本课程也不再继续展开。如果忘了，请同学们找些参考书来补上。

$$J = \begin{bmatrix} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 5 & 1 \\ & & & & 5 \end{bmatrix}$$

$$\left(J - 3I\right)^3 \left(J - 5I\right)^2 = 0$$

3．方阵的模式规范化

实方阵的特征多项式为实系数多项式

实方阵可能有非实的复特征值

实方阵的非实复特征值以共轭形式成对出现，相应的特征向量也是共轭的。即若σ和ω均为实数，$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 均为实向量，则

$$(\sigma+j\omega)(\boldsymbol{u}+j\boldsymbol{v})=A(\boldsymbol{u}+j\boldsymbol{v}) \longrightarrow (\sigma-j\omega)(\boldsymbol{u}-j\boldsymbol{v})=A(\boldsymbol{u}-j\boldsymbol{v})$$

模式标准化的方法是：将对角化变换矩阵 $\boldsymbol{Q}$ 中的共轭非实复特征值 $\lambda=\sigma+j\omega$ 和 $\bar{\lambda}=\sigma-j\omega$ 对应的共轭非实复特征向量 $\boldsymbol{q}=\boldsymbol{u}+j\boldsymbol{v}$ 和 $\bar{\boldsymbol{q}}=\boldsymbol{u}-j\boldsymbol{v}$ 分别用实向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 取代；

模式标准化的结果是：变换后的目标矩阵中，对角化的二阶矩阵块（含有非实复特征值 $\lambda=\sigma+j\omega$ 和 $\bar{\lambda}=\sigma-j\omega$）变为相应的实的非对角的二阶矩阵块，即：

$$\begin{bmatrix}\sigma+j\omega & 0\\ 0 & \sigma-j\omega\end{bmatrix} \Rightarrow \begin{bmatrix}\sigma & \omega\\ -\omega & \sigma\end{bmatrix}$$

3．方阵的模式规范化

实方阵的特征多项式为实系数多项式

实方阵可能有非实的复特征值

实方阵的非实复特征值以共轭形式成对出现，相应的特征向量也是共轭的。即若σ和ω均为实数，$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 均为实向量，则

$$(\sigma + j\omega)(\boldsymbol{u} + j\boldsymbol{v}) = A(\boldsymbol{u} + j\boldsymbol{v}) \longrightarrow (\sigma - j\omega)(\boldsymbol{u} - j\boldsymbol{v}) = A(\boldsymbol{u} - j\boldsymbol{v})$$

模式标准化的方法是：将对角化变换矩阵 $\boldsymbol{Q}$ 中的共轭非实复特征值 $\lambda = \sigma + j\omega$ 和 $\bar{\lambda} = \sigma - j\omega$ 对应的共轭非实复特征向量 $\boldsymbol{q} = \boldsymbol{u} + j\boldsymbol{v}$ 和 $\bar{\boldsymbol{q}} = \boldsymbol{u} - j\boldsymbol{v}$ 分别用实向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 取代；

模式标准化的结果是：变换后的目标矩阵中，对角化的二阶矩阵块（含有非实复特征值 $\lambda = \sigma + j\omega$ 和 $\bar{\lambda} = \sigma - j\omega$）变为相应的实的非对角的二阶矩阵块，即：

$$\begin{bmatrix} \sigma + j\omega & 0 \\ 0 & \sigma - j\omega \end{bmatrix} \Rightarrow \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

第二章习题：p72-76

2.1，2.6， 2.7， 2.8(a),(c)， 2.10

补充习题：求相似变换使如下矩阵等价为约当规范型和模式规范型

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$