

2154^v

校印

数理经济学与经济控制论基础

下 册

杨 小 凯 编

武汉大学经济管理系

一九八二年六月

数理经济学与经济控制论基础

目 录

下 册 经 济 控 制 论

第一章 经济系统	1
§1 经济耦合	1
§2 经济反馈	5
§3 稳定性	7
§4 状态空间描述	11
§5 离散经济系统的状态空间表达式	17
§6 状态方程的规范形式及自由运动	20
§7 相图	52
第二章 经济信息	32
§1 信息和熵	32
§2 自由度, 剩余度和信息效率	33
§3 编码	34
§4 经济信息传输	35
§5 大系统经济信息传输效率和可选择性并联耦合	37
§6 控制的限度	39
第三章 变分法和最大值原理	41
§1 问题的提起	41
§2 基本的数学概念	43
§3 泛函取极值的必要条件: 欧拉方程	46
§4 端点条件	53
§5 约束条件	60
§6 最大值原理	63
§7 最大值原理之端点条件	68
§8 控制问题的约束条件	71
§9 离散型最大值原理	74

第四章	经济管理中的最优控制问题	76
§1	最优经济增长和最优积累率	76
§2	最优库存控制	81
§3	最优设备更新和维修计划	90
§4	最优销售计划	95
§5	最优销售、研制计划和最优产品寿命	103
第五章	随机经济控制	110
§1	经济系统的能控性与能观性	110
§2	随机经济系统	113
§3	滤波理论	118
§4	动态规划与随机最优经济控制	126
§5	二次线性问题与必然等价原理	132
附录 A	函数取极大(小)值的充分条件, 函数的凹性、凸性与集合的凸性、凹性	137
附录 B	极大化之 Kuhn-Tucker 条件	142
参考文献目录		144

第一章 经济系统

§1 经济耦合

经济耦合是指各经济因素或子系统之间的因果关系链。最常见的经济耦合就是分工协作造成的投入产出关系。对于一个没有分工协作的自给自足社会来说,各个子系统之间互相独立,没有联系。这种经济系统叫做相对独立系统,如图1—1所示。图中有三个独立的子系统A, B, C, 分别表示三个自给自足的家庭, x_1, x_2, x_3 分别表示三个家庭对自然界的劳动输出, 而 y_1, y_2, y_3 分别表示三个家庭从自然界得到的收获(输入)。一般而言, **输出是系统对外界的作用, 而输入是外界对系统的作用**。严格地说, 图1—1中的三个家庭与自然界之间有一定的耦合关系, 它表现为输入输出之间的某种因果关系。其中 S_i 将输入 y_i 变换成输出 x_i , 而 R_i 将输出变换成反馈输入。但是三个家庭互相之间是独立的, 没有耦合。

图1—2中是三个分工协作的人, A 为专门生产粮食的农民, B 为专门生产纺织品的手工业者, C 为专门生产肉食的牧民。 x_1, x_2, x_3 分别为粮食、布匹和肉食的产量。图1—2叫

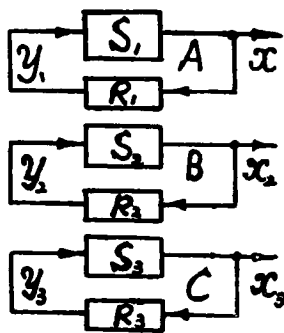


图 1—1

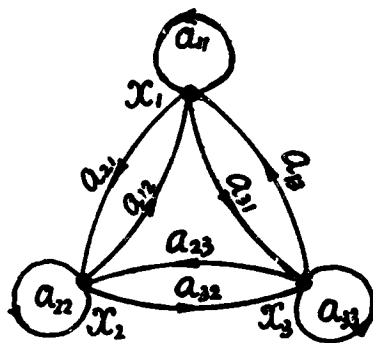


图 1—2

信号流图, 其中 A, B, C 三点叫结点, 结点标有变量 x_1, x_2, x_3 , a_{ij} 叫**传输系数**。图中的网络是下列联立方程组之拓扑图形:

$$\begin{aligned} x_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned} \quad (1-1)$$

易知, 这就是封闭的投入产出模型, 可见投入产出关系正是分工协作造成的输入、输出耦合关系, 这种耦合网络可以用图1—3中等价的**方框图**表示。图中 $x_{11} = a_{11}x_1$, $x_{12} = a_{12}x_2$, $x_{13} = a_{13}x_3$; $x_{21} = a_{21}x_1$, $x_{22} = a_{22}x_2$, $x_{23} = a_{23}x_3$; $x_{31} = a_{31}x_1$, $x_{32} = a_{32}x_2$, $x_{33} = a_{33}x_3$ 。而 A, B, C

的传输系数都是 1。式(1-1)，图1-2，图1-3这三个系统之间可以用一一对应的变换将一个系统化为另一个系统，这在控制论中叫做**同构**。

如果我们有二次收益函数

$$R = px = p(ap + b) = ap^2 + bp \quad a < 0, b > 0 \quad (1-2)$$

其中 $x = ap + b$ 为线性需求函数，于是边际收益为

$$\frac{dR}{dp} = 2ap + b \quad (1-3)$$

它与物理中的均加速（或减速）运动微分方程

$$\frac{dS}{dt} = 2a't + b' \quad (1-4)$$

就是完全同构的。我们只要用一一对应的变换，令总收益 R 等于距离 S ，令价格 p 等于时间 t ，且 $a = a'$ ， $b = b'$ ，则(1-3)就能变换成(1-4)。对于同构的系统，我们可以在一个系统中**模拟**另一个系统。

耦合分为**串联耦合**，**并联耦合**，以及**开环耦合**、**闭环耦合**（**反馈耦合**）。

从图1-4的方框图和信号流图可知，系统的初始输入为 x_1 ，第一个子系统之传输系数

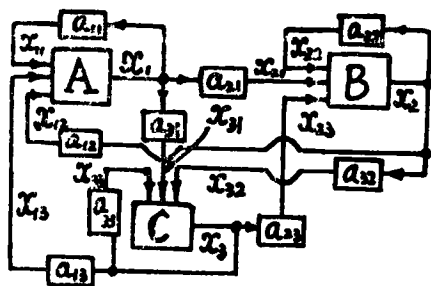


图 1-3

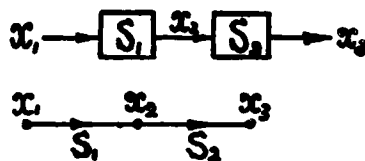


图 1-4

为 S_1 ，它的输出 $x_2 = S_1 x_1$ 就是第二个子系统的输入，一个系统之输出是与之耦合的系统之输入，这就是**串联耦合**。从图中不难看出，最终输出 $x_3 = S_2 x_2$ ，将 $x_2 = S_1 x_1$ 代入，则

$$\begin{aligned} x_3 &= S_2 x_2 \\ &= S_2 S_1 x_1 \end{aligned} \quad (1-5)$$

即串联系统的总传输系数等于子系统传输系数之积。

图1-5中的方框图和等价的信号流图，两个子系统将输入 x 变换成两个并行的输出 y_1 ，

y_2 , 这就叫**并联耦合**。从图中可以看出:

$$\begin{aligned} y_1 &= S_1 x, \\ y_2 &= S_2 x, \\ y &= y_1 + y_2 \\ &= S_1 x + S_2 x \\ &= (S_1 + S_2) x \quad (1-6) \end{aligned}$$

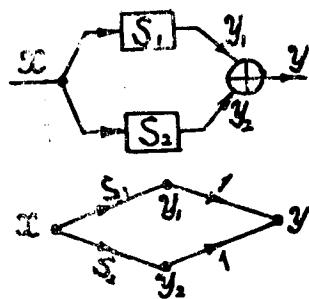


图 1-5

即, **并联耦合系统之总传输系数等于子系统传输系数之和**。

回顾(1-1), 图1-2, 图1-3, 不难看出分工协作的投入产出系统正是既有串联耦合又有并联耦合的系统。(1-1)中的多项求和正显含并联耦合。对于此式我们还可以进行无穷次迭代, 例如将三式代入一式中的 x_1, x_2, x_3 , 则有

$$\begin{aligned} x_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ &= a_{11}(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) + a_{12}(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) + a_{13}(a_{31}x_1 + a_{32}x_2 + a_{33}x_3) \\ &= a_{11}a_{11}x_1 + a_{11}a_{12}x_2 + a_{11}a_{13}x_3 \\ &\quad + a_{12}a_{21}x_1 + a_{12}a_{22}x_2 + a_{12}a_{23}x_3 \\ &\quad + a_{13}a_{31}x_1 + a_{13}a_{32}x_2 + a_{13}a_{33}x_3 \\ &= (a_{11}a_{11} + a_{12}a_{21} + a_{13}a_{31})x_1 + (a_{11}a_{12} + a_{12}a_{22} + a_{13}a_{32})x_2 + (a_{11}a_{13} + a_{12}a_{23} + a_{13}a_{33})x_3 \end{aligned}$$

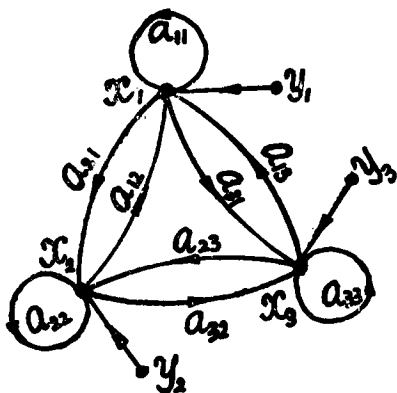


图 1-6

从以上推导的结果不难看出系数乘积反映了串联耦合, 系数求和反映了并联耦合。在图1-2和图1-3中我们不断重复子系统相互之间的投入产出关系, 很容易对照上式理解这种既有串联耦合又有并联耦合的特点。

图1-4和图1-5中的耦合叫做**开环耦合**。而图1-2, 图1-3中 A 之输出 x_1 , 通过**反馈传输系数** a_{11} 变换成 A 之输入 $x_{11} = a_{11}x_1$, 这种**把输出变换成输入**反馈系统的耦合叫**反馈耦合**, 即**闭环耦合**。

设我们有下列开放型投入产出系统, 其拓扑图形如图1-6所示。

$$\begin{aligned} x_1 &= y_1 + a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2 &= y_2 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3 &= y_3 + a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned} \quad (1-7)$$

上式可写成矩阵形式

$$X = Y + AX \quad (1-8)$$

其中

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

从(1—8)可解得

$$X = \frac{1}{I - A} Y = (I - A)^{-1} Y \quad (1-9)$$

相应于(1—8)和(1—9)之信号流图如图1—7所示。这就是反馈耦合，图中的系统将输出 X 通过反馈传输系数 A 变换成输入。而(1—9)就是反馈系统的基本公式。 $(I - A)^{-1}$ 为**反馈因子**，它就是著名的列昂节夫逆阵。

(1—9)可展开成无穷级数

$$X = (I - A)^{-1} Y = (I + A + A^2 + \dots) Y$$

其中 $A^2 = AA$, $A^3 = AAA \dots$ ，这些传输系数之乘积反映了串联耦合的特点，而无穷级数反映了反馈耦合的特点，它说明了闭环回路中输入输出的无穷次交互影响。将矩阵与矢量的乘积展开后，又可从求和式中看出并联耦合的特点。我国过去采用的计划平衡重复法就是用无穷级数的头一二项近似(1—9)。可见投入产出分析的重大意义在于它反映了分工协作的经济系统中串联耦合、并联耦合和反馈耦合的错综复杂网络关系。

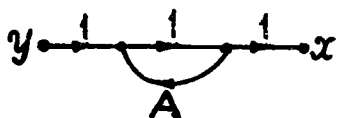


图 1—7

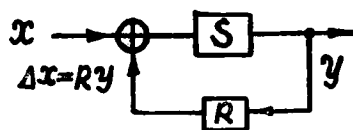


图 1—8

一般而言，设输入为 x ，输出为 y ，被调节系统的传输系数为 S ，反馈传输系数为 R ，则反馈输入

$$\Delta x = Ry \quad (1-11)$$

总输出 y 为初始输入 x 和反馈输入 Δx 之和乘上传输系数 S ：

$$y = S(x + \Delta x) \quad (1-12)$$

将(1—11)代入，则

$$y = Sx + SRy$$

$$y = \frac{S}{1 - SR} x \quad (1-13)$$

这就是一般的反馈调节公式，反馈因子为 $\frac{1}{1 - SR}$ ，而 $\frac{y}{x} = \frac{S}{1 - SR}$ 为**调节系统之传输系数**，(1—13)就是经典控制论中的**传递函数**。

反馈耦合的另一个例子是凯恩斯乘数原理。设国民收入为 Y ，消费为 C ，投资为 I ，则

$$Y = C + I \quad (1-14)$$

又设消费函数为

$$C = cY \quad (1-15)$$

其中 c 为消费倾向，它满足 $0 < c < 1$ ，将上式代入(1-14)，则

$$Y = cY + I$$

$$Y = \frac{1}{1-c} I \quad (1-16)$$

$\frac{1}{1-c}$ 就是反馈因子，它就是著名的凯恩斯乘数。相应的方框图如图1-9。

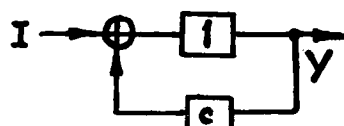
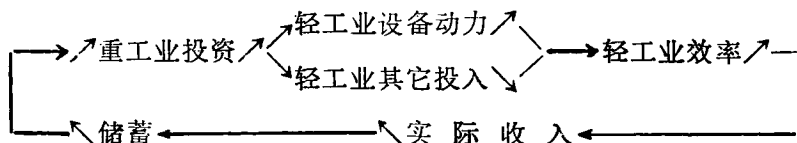


图 1-9

在经济分析中，我们常常假设系统内部结构为未知的“黑箱”，而只研究输出对输入的反应，通过这种研究来了解系统特性，是谓“黑箱原理”。宏观经济学中的总量分析就大量运用了这一原理。

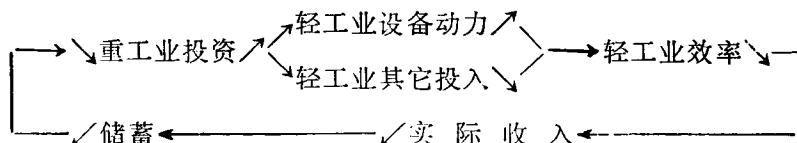
§2 经济反馈

经济反馈分为正反馈和负反馈。若系统输出与反馈输入符号相同，则为正反馈，若系统输出与反馈输入符号相反，则为负反馈。我们用下列动态图描述反馈特性。



这个动态图描述了重工业投资效果很好时，轻工业与重工业之间的耦合。如果重工业投资增多，这一方面使轻工业设备、动力增多，另一方面由于资金有限，使轻工业的其它投入下降，投资效果很好，则前一种影响大于后一种影响，因此两个并联的总输出为正，使轻工业效率上升，人民实际收入增加，又使储蓄上升，反过来使重工业投资来源增多。这就是轻重工业之间的良性循环，它是经济起飞的原因之一。“良性循环”，“起飞”，这就是正反馈，即动态图中“重工业投资”两边的箭头方向相同。初始输入符号为正（动态图中箭头向上）的正反馈叫**正向正反馈**，若初始输入符号为负（动态图中箭头向下）的正反馈叫**负向正反馈**，负向正反馈就是恶性循环。

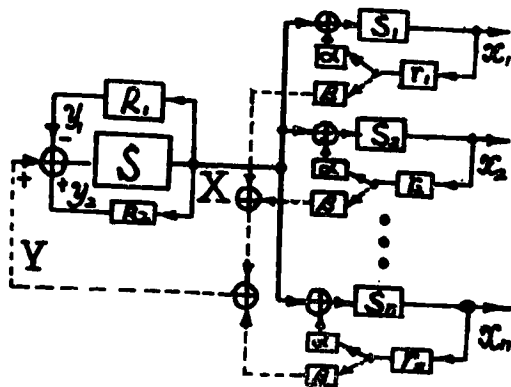
如果重工业投资效果不好，则有下列动态图：



```

graph LR
    HI[重工业投资] --> LI[轻工业效率]
    HI --> LID[轻工业设备动力]
    HI --> LIOT[轻工业其它投入]
    LID --> LI[轻工业效率]
    LIOT --> LI[轻工业效率]
    LI --> AI[实际收入]
    AI --> S[储蓄]
    S --> HI
  
```

在一般工程控制中正反馈是坏事，稳定性负反馈才是好事，而经济控制中，很多正向正反馈是人们所希望的，而某些停滞性负反馈却是人们企图避免的。另外经济系统中有大量非线性元件，可以将负反馈变成正反馈，或把正反馈变成负反馈，例如重工业投资效果就是这类参数，投资效果提高，就可以将负反馈变成正反馈。



下面我们用工框图来描述正反馈与负反馈的差别。图 1—10 中的实线部分是没有专利制度时科技发明活动中，发明人与使用人之间的耦合。加上虚线部分后，就是在专利制度调节下科技发明活动中的耦合。图中的 S 为从事科技发明的人将科技发明的总输入（收入与费用之代数和）变换成科技发明成果（输出） X 之传输系数。在没有专利制度时，科技发明对于发明人的直接收益，即发明人使用这一成果的直接好处（例如发明电

6

$$Y = \sum_{i=1}^n \beta r_i x_i.$$

一般 $Y \gg y_1$, 所以总输入 $Y - y_1 + y_2 > 0$, 系统为闭环的正反馈, 输出 X 与输入 $Y - y_1 + y_2$ 符号相同。这种正反馈使科技能加速发展。发明人希望自己的成果尽量推广, 因为掌握这种新技术的人越多, 则发明人的收入也越多, 收入越多, 发明人更有资金和积极性创造发明, 这种良性循环造成人材辈出, 群星灿烂的局面。

§3 稳定性

若经济系统有平衡状态, 当外界对系统的干扰使系统背离平衡状态时, 系统是否能消除干扰恢复平衡, 这就是所谓稳定性问题。

我们通过市场反馈调节的“蛛网定理”来说明稳定性问题。

设市场对某种商品的需求函数为

$$x_t = a - b p_t, \quad a, b > 0, \quad (1-17)$$

其中 p_t, y_t 分别为 t 时刻之价格和需求量, 又设 t 时刻此种商品的供给量 x'_t 是由 $t-1$ 时刻的价格决定, 因为生产要对价格作出反应有一时间滞后, 于是供给函数为

$$y'_t = \alpha + \beta p_{t-1}, \quad \alpha, \beta > 0 \quad (1-18)$$

在供求均衡的情况下

$$x_t = x'_t, \quad (1-19)$$

这时价格也不再变动, 即

$$p_t = p_{t-1} = \bar{p} \quad (1-20)$$

其中 $\bar{p}$ 表示均衡价格, 联立(1-17), (1-18), (1-19)和(1-20), 则可解出均衡价格和均衡供求量分别为

$$\bar{p} = \frac{a - \alpha}{b + \beta} \quad \bar{x} = \alpha + \beta \bar{p} = \frac{\alpha b + a \beta}{b + \beta} \quad (1-21)$$

其中 $\bar{x}$ 为均衡供求量。

如果市场机制使供求量在非均衡值上达到瞬态相等, 即

$$x_t = a - b p_t = \alpha + \beta p_{t-1} = x'_t \quad (1-22)$$

而均衡供求量为

$$\bar{x} = a - b \bar{p} = \alpha + \beta \bar{p} \quad (1-23)$$

(1-22)与(1-23)之差为

$$\bar{x}_t - \bar{x} = -b(p_t - \bar{p}) = \beta(p_{t-1} - \bar{p}) \quad (1-24)$$

令 $\hat{x}_t = x_t - \bar{x}$ 为 t 时刻需求量与均衡供求量之偏差, 且 $\hat{p}_t = p_t - \bar{p}$, $\hat{p}_{t-1} = p_{t-1} - \bar{p}$ 分别 t 和 $t-1$ 时刻市场价格与均衡价格之偏差, 代入(1-24), 我们有:

$$\hat{x}_t = -b\hat{p}_t = \beta\hat{p}_{t-1}, \quad (1-25)$$

所以

$$\hat{p}_t = A\hat{p}_{t-1}$$

其中 $A = -\frac{\beta}{b}$, 这是一个一阶齐次线性差分方程, 设 $\hat{p}_0 = p_0 - \bar{p}$ 为初始时刻价格与均衡价格之差, 则

$$\hat{p}_1 = A\hat{p}_0,$$

$$\hat{p}_2 = A\hat{p}_1 = A(A\hat{p}_0) = A^2\hat{p}_0$$

.....

$$\hat{p}_t = A\hat{p}_{t-1} = A^2\hat{p}_{t-2} = \dots = A^t\hat{p}_0 \quad (1-27)$$

上式给出了线性蛛网模型之通解, 求解的迭代过程可表为图 1-11 中的蛛网图和时间序列图, 由于 $b > 0, \beta > 0$, 所以 $A = -\frac{\beta}{b} < 0$, 当 t 不断从偶数变为奇数, 又从奇数变为偶数

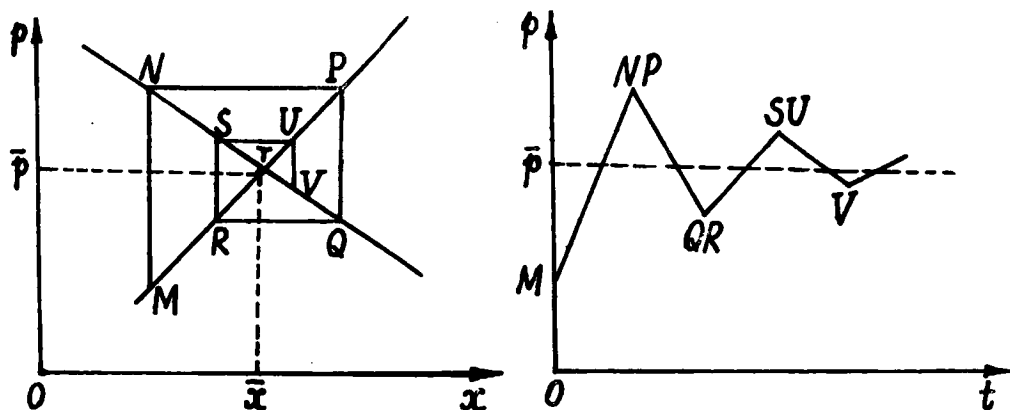


图 1-11

时, A^t 就不断从正数变为负数, 又从负数变为正数, 因而 $\hat{p}_t$ 一会儿为正, 一会儿为负, 即 p_t 一会儿高于 $\bar{p}$ 一会儿低于 $\bar{p}$ 。所以实际价格在均衡价格上下振荡。当价格处于图中 M 点时, 实际价格低于均衡价格 $\bar{p}$, 实际供给量也低于均衡供求量 $\bar{x}$, 而实际需求却大于 $\bar{x}$, 所以供不应求, 迫使价格从 M 点上升至 N 点, 在此点, 价格高于均衡价格 $\bar{p}$, 需求量低于 $\bar{x}$, 供给量由 P 点确定, 它高于 $\bar{x}$, 所以供过于求, 这又使价格从 P 点下跌到 Q 点,

这时价格低于均衡价格 $\bar{p}$ ，需求量大于 $\bar{x}$ ，供给量由 R 确定，它小于 $\bar{x}$ ，所以出现供不应求，又使价格上升到 S 点，如此循环下去。图中右部为这一振荡的时间轨迹，如果 $\beta < b$ ，则 $|A| = \left| -\frac{\beta}{b} \right| < 1$ ，当 $t \rightarrow \infty$ 时， $|A^t| = \left| \left(-\frac{\beta}{b} \right)^t \right| \rightarrow 0$ ，因而 $\hat{p}_t = A^t \hat{p}_0 \rightarrow 0$ ，即 $p_t \rightarrow \bar{p}$ 。这种市场调节的负反馈时程是种减幅振荡，它最终使价格稳定在均衡价格上，因而供求量也趋于均衡。不管干扰使初始价格处于何种状态，例如干扰使初始价格不是在 M 点，而是在 K 点或其它点，系统还是能通过减幅振荡的负反馈调节恢复均衡状态，因此这种市场反馈调节系统是稳定的。我们可以用猪肉的市场价格说明这种系统的特点，猪肉的当前价格直接决定了需求量，但猪的繁殖育肥有一个时间滞后（例如一年），猪肉价格变化会引起下一阶段猪肉供给量变化，如果初始时刻猪肉价格太低，造成供不应求，以后价格又提得过高，出现供过于求，这时出现正式或变相降价，于是人们杀母猪，减少小猪繁育，但由于供给有时滞，已生产的猪肉仍很多，使市场供过于求的现象不能马上消失，于是猪肉继续变相降价，例如收购时压级压价，拒不收购，使养猪实际收入下降，这都使猪肉实际价格低于均衡价格，因而使母猪和小猪数量低于均衡量，下一年又造成猪肉供不应求，不过这种振荡一般是收敛的， $|A|$ 越小，收敛过程越快。

如果 $|A| = \left| -\frac{\beta}{b} \right| > 1$ ，则当 $t \rightarrow \infty$ 时， $|A^t| = \left| \left(-\frac{\beta}{b} \right)^t \right| \rightarrow \infty$ 相应的蛛网图和时间序列图如图 1—12 所示。从图中可看出这是种增幅振荡的负反馈，一旦有干扰使系统脱离均衡状态，则这种反馈不能使系统回复到均衡状态，所以这种系统是不稳定的。

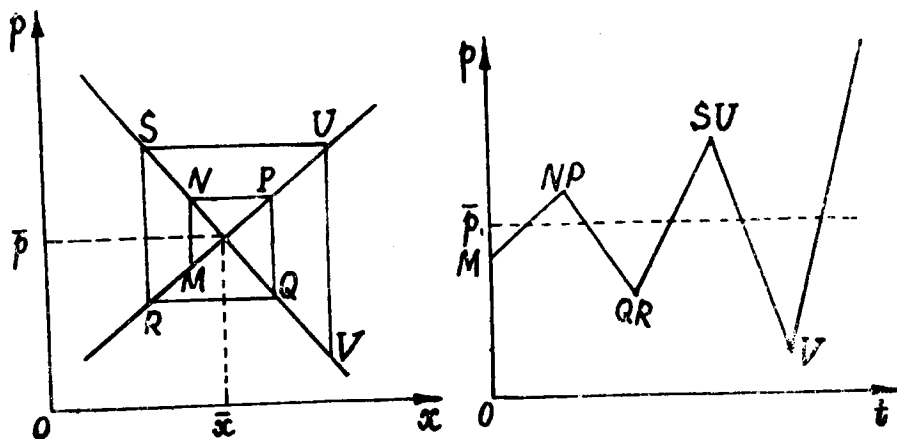


图 1—12

如果 $|A| = \left| -\frac{\beta}{b} \right| = 1$ ，则当 $t \rightarrow \infty$ 时，恒有 $|A^t| = \left| \left(-\frac{\beta}{b} \right)^t \right| = 1$ ，所以当 t 为奇数时， $\hat{p}_t = -\hat{p}_0$ ，当 t 为偶数时 $\hat{p}_t = \hat{p}_0$ ，相应的蛛网图和时间序列图如图 1—13 所示，这种系统在外界干扰下没有能力恢复供求均衡，但是它在不产生新的干扰时总是保持等幅振荡。

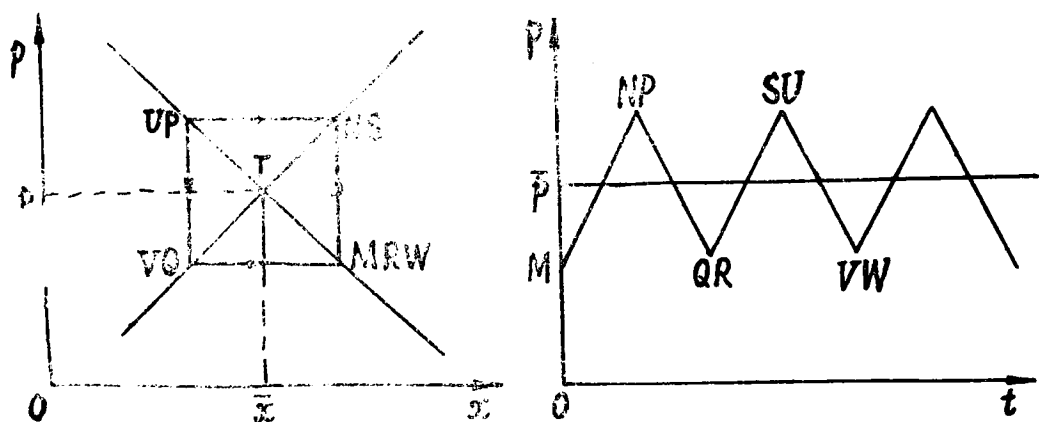


图 1-13

以上三种情况中, 由于 $A < 0$, 所以对于系统 $\hat{p}_t = A\hat{p}_{t-1}$ 来说, 输入 $\hat{p}_{t-1}$ 与输出 $\hat{p}_t$ 符号总是相反, 因此是种负反馈调节。如果 $A > 0$, 则 $\hat{p}_t$ 与 $\hat{p}_{t-1}$ 符号相同, 则系统是正反馈调节。

$A > 0$ 意味着 $-\frac{\beta}{b} > 0$, 则 β 与 b 的符号相反, 因此以前关于 $b > 0$, $\beta > 0$ 的假定不再都成立。若 $b < 0$, $\beta < 0$, 则意味着商品价格 p 越高, 需求量 $x = a - bp$ 越大; 若 $\beta > 0$, $b > 0$, 则意味着价格 p 越高, 供给量 $x' = a + \beta p$ 反而越少, 这两种情况对于正常商品来说不会出现。不过对于某些特殊商品, 也可能出现两种情形中的一种。

$A > 0$ 的正反馈与负反馈类似, 也分为三种类型。

当 $A = -\frac{\beta}{b} < 1$, 则正反馈如图 1-14 所示, 这是种**收敛的正反馈**, 这种正反馈总是使系统趋于均衡状态, 因而它是稳定的。

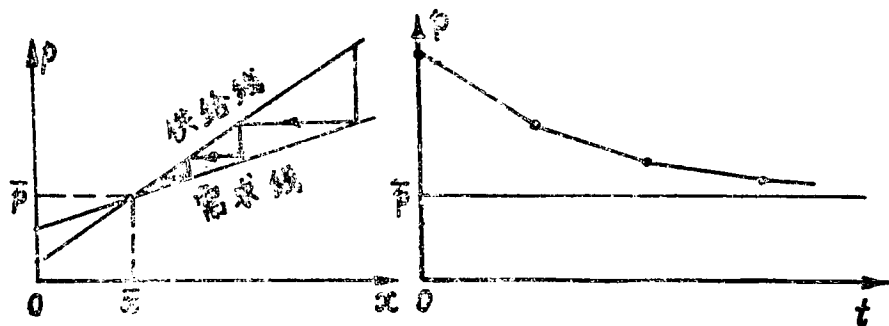


图 1-14

若 $A = -\frac{\beta}{b} > 1$, 则正反馈如图 1-15 所示, 这是种**发散正反馈**, 这种正反馈系统在

干扰下只会使实际价格和供求量与均衡价格和均衡供求量之差距越来越大，因此系统是不稳定的。

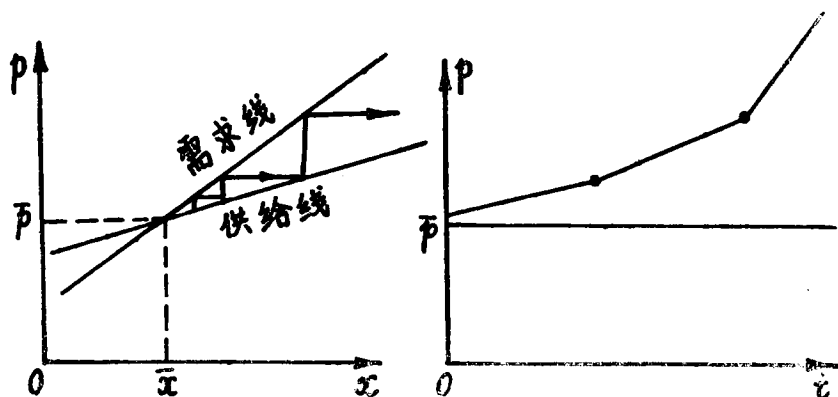


图 1—15

若 $A = -\frac{\beta}{b} = 1$ ，则正反馈如图 1—16 所示，这时可能供给函数与需求函数互相平行（如图左），这类系统根本没有均衡点；也可能供给函数和需求函数完全重合（如图右），这是种随处平衡系统，这两种情况一般都不可能出现。

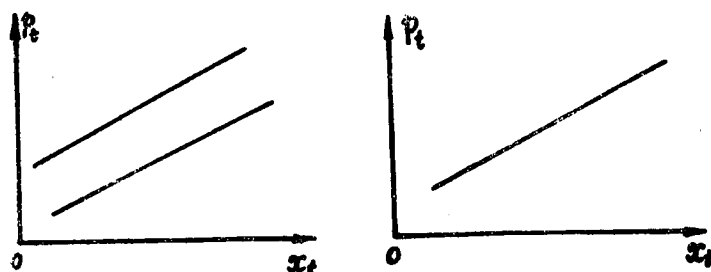


图 1—16

对于上述线性需求函数和线性供给函数的市场反馈调节我们可作如下总结：

A 之符号决定了反馈是正反馈还是负反馈，当 $A < 0$ 时，系统为负反馈，当 $A > 0$ 时，系统为正反馈；

A 之绝对值小于 1 还是大于 1 决定了反馈是收敛的还是发散的。若 $|A| < 1$ ，则反馈是收敛的，系统具有稳定性，若 $|A| > 1$ ，则反馈是发散的，系统不具备稳定性。若 $|A| = 1$ ，则系统或为等幅振荡（当 $A < 0$ 时）或为随处平衡或无均衡点（当 $A > 0$ 时）。

§ 4 状态空间描述

以兰格为代表的经典经济控制论主要是来用传递函数分析法，而现代经济控制论主要采用的是状态空间方法。

一个经济控制过程通常由**控制器**和**被控过程**组成，而被控过程包括**执行机构**、**被控系统**和**量测机构**，如图 1—17 所示。其中系统输入相当于经济计量学中的政策变量、工具变量或可控外生变量，系统干扰相当于不可控外生变量，而状态变量和输出变量相当于内生变量。

经济系统的**状态**是完整地描述系统之时域行为的最小一组变量，如果给定 $t = t_0$ 时刻此组变量的值和 $t \geq t_0$ 时输入的时间函数，那么系统在 $t \geq t_0$ 的任何瞬时的行为就完全确定了。

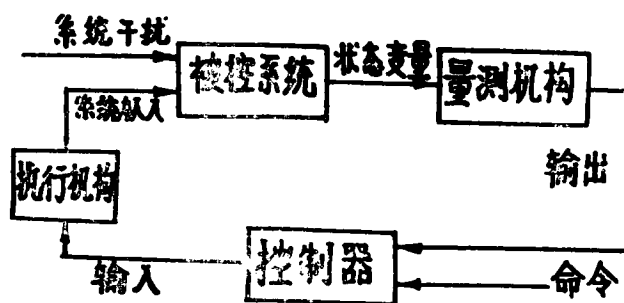


图 1—17

这一组变量称为**状态变量**，多个状态变量组成的矢量，叫**状态矢量**，设这一状态矢量为

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad (1-30)$$

当 $t_0 \leq t < \infty$ ， $x(t)$ 由 $x(t_0)$ 和 $t \geq t_0$ 时的输入函数 $u(t)$ 唯一地确定，而与 t_0 前时刻的状态和输入无关。

以 $x_1, x_2, \dots, x_n$ 为坐标构成的欧氏空间称为**状态空间**，状态空间中的每一点都代表了状态变量的唯一和特定的一组值，即系统的一个状态。若给定 t_0 瞬时系统的状态及 $t \geq t_0$ 时的输入，则 $t \geq t_0$ 各瞬时系统的状态，就构成状态空间中的一条轨线，

状态空间描述不但便于计算，而且在处理非线性系统时比传递函数或频率特性更具优越性。一个 n 阶系统，有且仅有 n 个状态变量可以选择，但状态变量组的选择不是唯一的。

经济系统之动态特性一般可表为 n 个一阶常微分方程构成的方程组：

$$\dot{x}_i = f_i(x_1, \dots, x_n; u, \dots, u_r; t), \quad i = 1, \dots, n \quad (1-31)$$

而经济系统之输出特性，可表为

$$y_j = g_j(x_1, \dots, x_n; u_1, \dots, u_r; t) \quad j = 1, \dots, m \quad (1-32)$$

上两式可简记为矢量形式

$$\dot{x} = f(x, u, t) \quad (1-31')$$

$$y = g(x, u, t) \quad (1-32')$$

其中

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad (1-33)$$

$$f(x, u, t) = \begin{pmatrix} f_1(x_1, \dots, x_n; u_1, \dots, u_r; t) \\ \dots \\ f_n(x_1, \dots, x_n; u_1, \dots, u_r; t) \end{pmatrix} \quad (1-34)$$

$$g(x, u, t) = \begin{pmatrix} g_1(x_1, \dots, x_n; u_1, \dots, u_r; t) \\ \dots \\ g_m(x_1, \dots, x_n; u_1, \dots, u_r; t) \end{pmatrix}$$

(1-31')叫系统之**状态方程**，而(1-32')叫系统之**输出方程**，这两个方程构成了系统在状态空间的描述，若 f, g 为 x, u 的非线性函数，则这类经济系统叫**非线性系统**，若被控过程是线性的，则叫**线性系统**，其状态方程和输出方程可表为

$$\begin{cases} \dot{x} = A(t)x + B(t)u \\ y = C(t)x + D(t)u \end{cases} \quad (1-35)$$

其中

$$A(t) = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix}, \quad B(t) = \begin{bmatrix} b_{11}(t) & \dots & b_{1r}(t) \\ \vdots & \ddots & \vdots \\ b_{n1}(t) & \dots & b_{nr}(t) \end{bmatrix}$$

$$C(t) = \begin{bmatrix} c_{11}(t) & \dots & c_{1n}(t) \\ \vdots & \ddots & \vdots \\ c_{m1}(t) & \dots & c_{mn}(t) \end{bmatrix}, \quad D(t) = \begin{bmatrix} d_{11}(t) & \dots & d_{1r}(t) \\ \vdots & \ddots & \vdots \\ d_{m1}(t) & \dots & d_{mr}(t) \end{bmatrix}$$

上述非线性和线性系统中 f, g 或 A, B, C, D 的元素显含时间因素 t ，这种系统叫做**时变系统**或**非定常系统**，而 f, g ，或 A, B, C, D 中不显含时间因素 t 的系统叫**时不变或定常系统**，其状态空间表达式为：

$$\begin{cases} \dot{x} = f(x, u) \\ y = g(x, u) \end{cases} \quad (1-36)$$

或

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases} \quad (1-37)$$

连续型经济系统之时域模型，可用单变量高阶微分方程表为

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \dots + b_{n-1} \dot{u} + b_n u \quad (1-39)$$

则我们可以通过适当的变换, 将其写成线性时不变状态空间表达式

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}u \end{cases} \quad (1-39)$$

若(1-38)中等式右边没有 u 之导数项, 则时域模型变为

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = bu \quad (1-40)$$

我们按照下列步骤进行变换:

(1) 选择状态变量, 由于 n 阶系统有 n 个状态变量, 若给定 $y(0), \dot{y}(0), \dots, y^{(n-1)}(0)$ 和 $t \geq 0$ 的输入 $u(t)$ 时, 系统在 $t \geq 0$ 时的运动就完全确定了, 因此可取 $y, \dot{y}, \dots, y^{(n-1)}$ 为系统的一组状态变量, 且令

$$\begin{cases} x_1 = y \\ x_2 = \dot{y} \\ \dots \\ x_n = y^{(n-1)} \end{cases} \quad (1-41)$$

(2) 利用以上定义, 我们可将(1-40)中的所有高阶导数化为新引入变量的一阶导数, 因而可以将高阶系统变换成一阶系统。从(1-40)和(1-41)可导出

$$\begin{cases} \dot{x}_1 = \dot{y} = x_2 \\ \dot{x}_2 = \ddot{y} = x_3 \\ \dots \\ \dot{x}_{n-1} = y^{(n-1)} = x_n \\ \dot{x}_n = y^{(n)} = -a_1 y^{(n-1)} - \dots - a_{n-1} \dot{y} - a_n y + bu \\ \quad = -a_1 x_n - \dots - a_{n-1} x_2 - a_n x_1 + bu \end{cases} \quad (1-42)$$

而系统之输出方程式为

$$y = x_1 \quad (1-43)$$

(3) 令

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{u} = [u], \quad \mathbf{y} = [y], \quad \mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix} \\ B &= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad C = [1 \ 0 \dots 0] \end{aligned}$$

则(1—40)可写为状态方程

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \quad (1-44)$$

和输出方程

$$\mathbf{y} = \mathbf{C}\mathbf{x} \quad (1-45)$$

下面我们再讨论(1—38)的一般时域模型怎样化为状态方程，我们将(1—38)重写如下：

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \dots + b_{n-1} \dot{u} + b_n u \quad (1-38)$$

变换步骤是：

(1) 令

$$\begin{cases} x_1 = y - \beta_0 u \\ x_2 = \dot{y} - \beta_0 \dot{u} - \beta_1 u \\ x_3 = \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u \\ \dots\dots\dots \\ x_n = y^{(n-1)} - \beta_0 u^{(n-1)} - \beta_1 u^{(n-2)} - \dots - \beta_{n-2} \dot{u} - \beta_{n-1} u \end{cases} \quad (1-48)$$

(2) 从上式可推出

$$\begin{aligned} a_n y &= a_n x_1 + a_n \beta_0 u \\ a_{n-1} \dot{y} &= a_{n-1} x_2 + a_{n-1} \beta_0 \dot{u} + a_{n-1} \beta_1 u \\ a_{n-2} \ddot{y} &= a_{n-2} x_3 + a_{n-2} \beta_0 \ddot{u} + a_{n-2} \beta_1 \dot{u} + a_{n-2} \beta_2 u \\ a_1 y^{(n-1)} &= a_1 x_n + a_1 \beta_0 u^{(n-1)} + \dots + a_1 \beta_{n-2} \dot{u} + a_1 \beta_{n-1} u \\ y^{(n)} &= x_{n+1} + \beta_0 u^{(n)} + \beta_1 u^{(n-1)} + \dots + \beta_{n-1} \dot{u} + \beta_n u \end{aligned}$$

(3) 不难看出(1—47)各方程等式左边相加即等于(1—38)的等式左边，因此(1—47)等式右边相加应等于(1—38)等式右边，于是我们有

$$\begin{aligned} &\{x_{n+1} + a_1 x_n + \dots + a_{n-1} x_2 + a_n x_1\} + \{\beta_0 u^{(n)} + (\beta_1 + a_1 \beta_0) u^{(n-1)} \\ &+ (\beta_2 + a_1 \beta_1 + a_2 \beta_0) u^{(n-2)} + \dots + (\beta_{n-1} + a_1 \beta_{n-2} + \dots + a_{n-2} \beta_1 + a_{n-1} \beta_0) \dot{u} \\ &+ (\beta_n + a_1 \beta_{n-1} + \dots + a_{n-1} \beta_1 + a_n \beta_0) u\} = b_0 u^{(n)} + b_1 u^{(n-1)} + \dots + b_{n-1} \dot{u} + b_n u, \end{aligned}$$

上式中 $u^{(k)} (k=0, 1, \dots, n)$ 之系数应相等，即

$$\begin{cases} \beta_0 = b_0 \\ \beta_1 = b_1 - a_1 \beta_0 \\ \beta_2 = b_2 - a_1 \beta_1 - a_2 \beta_0 \\ \dots\dots\dots \\ \beta_n = b_n - a_1 \beta_{n-1} - \dots - a_{n-1} \beta_1 - a_n \beta_0 \end{cases} \quad (1-48)$$

于是我们推出了计算(1—46)中变换系数 $\beta_k (k=0, 1, n)$ 的公式。

(4) 从(1—46)和(1—38)可推出

$$\begin{cases} \dot{x}_1 = \dot{y} - \beta_0 \dot{u} = x_2 + \beta_1 u \\ \dot{x}_2 = \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} = x_3 + \beta_2 u \\ \dots\dots\dots \\ \dot{x}_{n-1} = \dots = x_n + \beta_{n-1} u \\ \dot{x}_n = -a_n x_1 - a_{n-1} x_2 - \dots - a_1 x_n + \beta_n u \end{cases} \quad (1-49)$$

和 $y = x_1 + \beta_0 u \quad (1-50)$

这就是状态方程和输出方程，它们可写为矩阵形式

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}u \end{cases} \quad (1-51)$$

其中 $\dot{\mathbf{x}}, \mathbf{x}, u, \mathbf{y}, \mathbf{A}, \mathbf{C}$ 之定义与前相同，但

$$\mathbf{B} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \mathbf{D} = [\beta_0]$$

例 一个经济系统之动态模型为

$$\ddot{y} + 18\dot{y} + 192y = 160\dot{u} + 640u,$$

即相应于(1—38)中的已知系数为

$$a_1 = 18, \quad a_2 = 192, \quad a_3 = 640, \quad b_0 = 0, \quad b_1 = 0, \quad b_2 = 160, \quad b_3 = 640,$$

利用(1—48)可算出

$$\beta_0 = b_0 = 0$$

$$\beta_1 = b_1 - a_1 \beta_0 = 0$$

$$\beta_2 = b_2 - a_1 \beta_1 - a_2 \beta_0 = 160$$

$$\beta_3 = b_3 - a_1 \beta_2 - a_2 \beta_1 - a_3 \beta_0 = 640 - 18 \times 160 = -2240$$

于是(1—46)变换式为

$$x_1 = y - \beta_0 u = y$$

$$x_2 = \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{y}$$

$$x_3 = \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} + \beta_2 u = \ddot{y} - 160u$$

据此,可推出状态方程和输出方程为

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -64 & -192 & -18 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 160 \\ -2240 \end{pmatrix} [u]$$

$$[y] = [1 \ 0 \ 0] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

§ 5 离散经济系统的状态空间表达式

由于统计数据往往是按月按季或按年给出,所以经济系统的模型常常是离散型的,因此离散型状态空间表达式在经济分析中非常重要。

设一个经济系统可用下列高阶差分方程描述:

$$y(k+n) + a_1 y(k+n-1) + \cdots + a_{n-1} y(k+1) + a_n y(k) = bu(k) \quad (1-52)$$

将其变换成状态空间描述的步骤如下:

(1) 选择状态变量, 令

$$\begin{cases} x_1(k) = y(k) \\ x_2(k) = y(k+1) \\ \cdots \cdots \\ x_n(k) = y(k+n-1) \end{cases} \quad (1-53)$$

(2) 从(1-52), (1-53)导出

$$\begin{cases} x_1(k+1) = y(k+1) = x_2(k) \\ x_2(k+1) = y(k+2) = x_3(k) \\ \cdots \cdots \\ x_{n-1}(k+1) = y(k+n-1) = x_n(k) \\ x_n(k+1) = y(k+n) = -a_n x_1(k) - a_{n-1} x_2(k) - \cdots - a_1 x_n(k) + bu(k) \end{cases} \quad (1-54)$$

和

$$y(k) = x_1(k) \quad (1-55)$$

(3) 令

$$\mathbf{x}(k+1) = \begin{pmatrix} x_1(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix}, \quad \mathbf{x}(k) = \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdots & \cdots & -a_0 \end{bmatrix},$$

$$B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ b \end{pmatrix}, \quad u = [u], \quad C = [1 \ 0 \ \cdots \ 0], \quad y(k) = [y(k)]$$

于是可得到状态空间方程

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + Bu$$

和

$$y(k) = C\mathbf{x}(k) \quad (1-56)$$

如果描述经济系统的高阶差分方程中包含控制变量的高阶差分, 则我们有经济系统之离散模型

$$\begin{aligned} & y(k+n) + a_1y(k+n-1) + \cdots + a_{n-1}y(k+1) + a_ny(k) \\ & = b_0u(k+n) + \cdots + b_nu(k) \end{aligned} \quad (1-57)$$

与连续系统的情形类似, 我们有将其化为状态空间描述的如下步骤:

(1) 令

$$\begin{cases} x_1(k) = y(k) - h_0u(k) \\ x_2(k) = x_1(k+1) - h_1u(k) \\ x_3(k) = x_2(k+1) - h_2u(k) \\ \dots\dots\dots \\ x_n(k) = x_{n-1}(k+1) - h_{n-1}u(k) \end{cases} \quad (1-58)$$

(2) 待定系数 $h_0, \dots, h_n$ 之计算公式为

$$\begin{cases} h_0 = b_0 \\ h_1 = b_1 - a_1h_0 \\ h_2 = b_2 - a_1h_1 - a_2h_0 \\ \dots\dots\dots \\ h_n = b_n - a_1h_{n-1} - \cdots - a_{n-1}h_1 - a_nh_0 \end{cases} \quad (1-59)$$

(3) 从(1-57)和(1-58)导出

$$\begin{cases} x_1(k+1) = x_2(k) + h_1u(k) \\ x_2(k+1) = x_3(k) + h_2u(k) \\ \dots\dots\dots \\ x_{n-1}(k+1) = x_n(k) + h_{n-1}u(k) \\ x_n(k+1) = -a_nx_1(k) - a_{n-1}x_2(k) - \cdots - a_1x_n(k) + h_nu(k) \end{cases} \quad (1-60)$$

$$(4) \text{ 令 } \mathbf{x}(k+1) = \begin{pmatrix} x_1(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdots & \cdots & -a_1 \end{pmatrix}, \quad \mathbf{x}(k) = \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix}$$

$$H = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \quad u(k) = [u(k)], \quad y(k) = [y(k)], \quad C = [1 \ 0 \ \cdots \ 0], \quad h_0 = [h_0]$$

则状态方程和输出方程可写为

$$\begin{cases} \mathbf{x}(k+1) = A\mathbf{x}(k) + H\mathbf{u}(k) \\ \mathbf{y}(k) = C\mathbf{x}(k) + h_0\mathbf{u}(k) \end{cases} \quad (1-61)$$

$$(1-62)$$

例 设有经济系统

$$\begin{cases} Y_t = C_t + I_t + G_t \\ C_t = c_0 + c_1 Y_{t-1} \\ I_t = i_0 + i_1 (Y_{t-1} - Y_{t-2}) \end{cases}$$

其中 Y_t, C_t, I_t, G_t 分别为 t 时刻之国民收入、消费、投资和政府支出, c_0, c_1, i_0, i_1 为大于 0 之常数, 从上式可推出二阶差分方程

$$Y_{t+2} + (c_1 + i_1)Y_{t+1} + i_1 Y_t = c_0 + i_0 + G_t$$

$$x_1(t) = Y_t$$

我们令

$$x_2(t) = Y_{t+1}$$

则可导出

$$x_1(t+1) = Y_{t+1} = x_2(t)$$

$$x_2(t+1) = Y_{t+2} = (c_1 + i_1)x_2(t) - i_1 x_1(t) + c_0 + i_0 + G_t$$

于是我们有状态方程

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -i_1 & c_1 + i_1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [c_0 + i_0 + G_t]$$

及输出方程

$$[Y(t)] = [1 \ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

状态方程用信号流图描述非常方便。(1-37)的状态空间表达式可表为图1-18中的信号

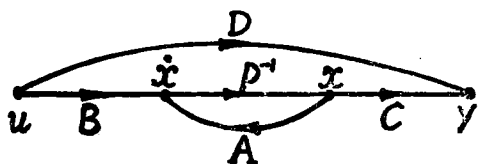


图 1-18

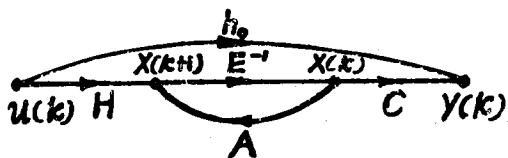


图 1-19

流图, 其中的 P^{-1} 为积分算子, 即为微分算子 P 之逆, $P = \frac{d}{dt}$.

令后向移位算子为 E^{-1} , $E^{-1}x(k+1) = x(k)$, 则离散状态空间表达式 (1—61) 和 (1—62) 可表为图 1—19 中的信号流图。

§ 6 状态方程的规范形式及自由运动。

设有线性时不变系统

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \quad (1-63)$$

其中 $\mathbf{A}$ 为 $n \times n$ 常阵, $\mathbf{B}$ 为 $n \times r$ 常阵。这一系统的特征值定义为 $\mathbf{A}$ 之**特征值**, 就是特征方程

$$|\lambda \mathbf{I} - \mathbf{A}| = 0 \quad (1-64)$$

之根, 系统之特征值有如下性质:

- (1) 对系统 (1—63) 作线性非奇异变换, 其特征值不变;
- (2) 若 λ_i 为 $\mathbf{A}$ 的一个特征值, 且存在某个 n 维非 0 矢量 $\mathbf{v}_i$, 使下式成立:

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i,$$

则称 $\mathbf{v}_i$ 为 $\mathbf{A}$ 之属于 λ_i 之特征矢量。从上式可推出

$$[\lambda \mathbf{I} - \mathbf{A}]\mathbf{v} = 0,$$

由于 $\mathbf{v}$ 为非 0 矢量, 所以只有 (1—64) 成立, 即 $[\lambda \mathbf{I} - \mathbf{A}]$ 为奇异矩阵, 上式才成立。

- (3) 若特征值 $\lambda_1, \dots, \lambda_n$ 两两相异, 则 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 必线性无关。

$$(4) \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdots & \cdots & -a_1 \end{bmatrix} \quad \text{之特征多项式为}$$

$$|\lambda \mathbf{I} - \mathbf{A}| = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n \quad (1-65)$$

- (5) 若 (1—63) 之特征值两两相异, 则定义如下状态方程

$$\dot{\hat{\mathbf{x}}} = \hat{\mathbf{A}} \hat{\mathbf{x}} + \hat{\mathbf{B}}u \quad (1-66)$$

为系统的**对角线规范形式**, 其中 $\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}$

为对角线矩阵, 而其对角线元素即系统之特征值, 由于对角线规范形式中, 每个方程只含有一个状态变量, 状态变量各个方程之间的联系被分离了, 形成 n 个独立的状态变量方程, 从而可使求解过程大大简化。对于时不变线性系统 (1—63), 只要其特征值是两两相异的, 则必可通过一个非奇异变换矩阵 $\mathbf{P}$ 将其化为 (1—66) 的对角线规范形式, 使得 $\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, $\hat{\mathbf{B}} = \mathbf{P}^{-1}\mathbf{B}$ 。

若系统矩阵

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

则使系统状态方程对角线规范化的非奇异阵

$$P = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}, \quad (1-67)$$

对最后个性质我们可证明如下:

根据特征矢量定义 $Av_i = \lambda_i v_i$, 即

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & \cdots & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ \vdots \\ v_{ni} \end{bmatrix} = \lambda_i \begin{bmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ \vdots \\ v_{ni} \end{bmatrix}, \quad (i = 1, \cdots, n)$$

从上式可导出

$$\begin{cases} v_{2i} = \lambda_i v_{1i} \\ v_{3i} = \lambda_i v_{2i} \\ \cdots \cdots \\ v_{ni} = \lambda_i v_{(n-1)i} \\ -(a_n v_{1i} + a_{n-1} v_{2i} + \cdots + a_1 v_{ni}) = \lambda_i v_{ni} \end{cases} \quad (i = 1, \cdots, n)$$

上式可简化为

$$\begin{cases} (\lambda_i^n + a_{n-1} \lambda_i^{n-1} + \cdots + a_1 \lambda_i + a_n) v_{1i} = 0 \\ v_{2i} = \lambda_i v_{1i} \\ v_{3i} = \lambda_i^2 v_{1i} \\ \cdots \cdots \\ v_{ni} = \lambda_i^{n-1} v_{1i} \end{cases} \quad (i = 1, \cdots, n)$$

令 $v_{1i} = 1$, 则得到

$$\begin{cases} v_{1i} = 1 \\ v_{2i} = \lambda_i \\ \cdots \cdots \\ v_{ni} = \lambda_i^{n-1} \end{cases} \quad (i = 1, \cdots, n)$$

即

$$v_i = \begin{pmatrix} 1 \\ \lambda_i \\ \lambda_i^2 \\ \vdots \\ \lambda_i^{n-1} \end{pmatrix}$$

$\because P = (v_1, v_2, \dots, v_n)$, $\therefore$ 性质(5)得证

例 设有时不变线性系统

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} x + \begin{pmatrix} 9 \\ 7 \\ 15 \end{pmatrix} u$$

其特征值为 $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = -1$, (两两相异), 故可取

$$P = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ 4 & 1 & 1 \end{pmatrix},$$

$$P^{-1} = \begin{pmatrix} -\frac{1}{3} & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix}.$$

使得

$$\begin{aligned} \hat{A} = P^{-1}AP &= \begin{pmatrix} -\frac{1}{3} & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ 4 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{3} & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 4 & 1 & 1 \\ 8 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

$$\hat{B} = P^{-1}B = \begin{pmatrix} -\frac{1}{3} & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 9 \\ 7 \\ 15 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$$

如果系统中没有控制变量 u 和扰动则叫**自由系统**, 若其中包括依赖于控制或扰动的项, 则

强制系统。

设有线性时不变自由系统，其对角规范形式为

$$\dot{\mathbf{x}} = \hat{A} \mathbf{x} \quad (1-68)$$

其中 $\dot{\mathbf{x}} = \begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\hat{A} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{pmatrix}$, 由于 $\hat{A}$ 为对角形，因此上式为 n 个互

相独立的一阶线性齐次微分方程，若初始状态为 $\mathbf{x}_0 = \begin{pmatrix} x_{10} \\ \vdots \\ x_{n0} \end{pmatrix}$ ，则(1-68)之解为

$$\mathbf{x} = e^{At} \mathbf{x}_0 \quad (1-69)$$

其中

$$e^{At} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & e^{\lambda_2 t} & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$$

(1-68)稳定性的条件为， A 之本征值为负，所以稳定性的两个必要条件为

- (1) $\sum_{i=1}^n \lambda_i < 0$;
 - (2) 当 n 为偶数时， $\prod_{i=1}^n \lambda_i = 0$;
- 当 n 为奇数时， $\prod_{i=1}^n \lambda_i < 0$

由于 $\hat{A} = P^{-1}AP$ ，所以对 $\hat{A}$ 元素 λ_i 的要求可变成对 A 之元素的要求，即必要条件变成

$$t_r(A) < 0$$

且

$$\begin{aligned} &\text{当 } n \text{ 为偶数时, } |A| > 0 \\ &\text{当 } n \text{ 为奇数时, } |A| < 0 \end{aligned} \quad (1-71)$$

若系统状态方程表为一阶齐次线性差分方程：

$$\mathbf{x}(t+1) = A\mathbf{x}(t) \quad (1-72)$$

则其解为

$$\mathbf{x}(t) = A^t \mathbf{x}(0) \quad (1-73)$$

其中 $\mathbf{x}(0)$ 为初始时刻系统之状态，由于

$$\hat{A} = P^{-1}AP$$

$\therefore$

$$A = P\hat{A}P^{-1}$$

又 $\because$

$$A^2 = (P\hat{A}P^{-1})(P\hat{A}P^{-1}) = P\hat{A}^2P^{-1},$$

类似地, 我们有 $A^t = P \hat{A}^t P^{-1}$. (1-74)

其中 $\hat{A}$ 为对角阵, 其 i 个对角线元素为 λ_i .

$\therefore \hat{A}^t$ 亦为对角阵, 其对角线元素为 λ_i^t . 由于 $P, P^{-1}, \mathbf{x}(0)$ 都是具体的数量矩阵或矢量, 所以解 $\mathbf{x}(t)$ 为 λ_i^t 之线性组合.

$$\begin{aligned} \text{例如 设 } \mathbf{x}(t) &= \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix}, \quad P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} p^{11} & p^{12} & p^{13} \\ p^{21} & p^{22} & p^{23} \\ p^{31} & p^{32} & p^{33} \end{pmatrix}, \\ \hat{A} &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \mathbf{x}_0 = \begin{pmatrix} x_{10} \\ x_{20} \\ x_{30} \end{pmatrix}, \quad \text{则 } \mathbf{x}(t+1) = A\mathbf{x}(t) \text{ 之解为} \\ \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} &= P \hat{A}^t P^{-1} \mathbf{x}_0 = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \begin{pmatrix} \lambda_1^t & 0 & 0 \\ 0 & \lambda_2^t & 0 \\ 0 & 0 & \lambda_3^t \end{pmatrix} \begin{pmatrix} p^{11} & p^{12} & p^{13} \\ p^{21} & p^{22} & p^{23} \\ p^{31} & p^{32} & p^{33} \end{pmatrix} \begin{pmatrix} x_{10} \\ x_{20} \\ x_{30} \end{pmatrix} \\ &= \begin{pmatrix} p_{11}\lambda_1^t & p_{12}\lambda_2^t & p_{13}\lambda_3^t \\ p_{21}\lambda_1^t & p_{22}\lambda_2^t & p_{23}\lambda_3^t \\ p_{31}\lambda_1^t & p_{32}\lambda_2^t & p_{33}\lambda_3^t \end{pmatrix} \begin{pmatrix} p^{11}x_{10} + p^{12}x_{20} + p^{13}x_{30} \\ p^{21}x_{10} + p^{22}x_{20} + p^{23}x_{30} \\ p^{31}x_{10} + p^{32}x_{20} + p^{33}x_{30} \end{pmatrix} \\ &= \begin{pmatrix} p_{11}z_{10}\lambda_1^t + p_{12}z_{20}\lambda_2^t + p_{13}z_{30}\lambda_3^t \\ p_{21}z_{10}\lambda_1^t + p_{22}z_{20}\lambda_2^t + p_{23}z_{30}\lambda_3^t \\ p_{31}z_{10}\lambda_1^t + p_{32}z_{20}\lambda_2^t + p_{33}z_{30}\lambda_3^t \end{pmatrix} \end{aligned} \quad (1-75)$$

其中 $z_{i0} = p^{i1}x_{10} + p^{i2}x_{20} + p^{i3}x_{30}$ (1-76)

将上例推广, 不难导出解式

$$x_{it} = \sum_{k=1}^n p_{ik} z_{k0} \lambda_k^t = \sum_{k=1}^n c_{ik} \lambda_k^t, \quad \text{其中 } c_{ik} = p_{ik} z_{k0}, \quad (1-77)$$

其中 λ_k 为 $n \times n$ 矩阵 A 之第 k 个特征根, p_{ik} 为对应于特征根 λ_k 的特征矢量 $\mathbf{p}_k$ 第 i 个元素, z_{k0} 由(1-76)定义.

从(1-77)不难推出, 若所有特征根 $\lambda_k (k=1, \dots, n)$ 之绝对值都小于 1, 则解 $x_{it} (i=1, \dots, n)$ 是收敛的, 因此系统是稳定的. 若任一特征根 (具有非 0 系数) 之绝对值大于 1, 则此解是发散的, 所以系统不稳定. 如果有些根是负的, 则当 t 在奇数和偶数之间变换时 λ_k^t 一会儿为负, 一会儿为正, 因此解之行为是振荡的. 如果有些根是复数, 亦会出现振荡. 复根总是成对出现, 例如若有复根

$$\lambda_1 = a + bi,$$

其中 a, b 为常数, $i = \sqrt{-1}$, 则必有共轭复根

$$\lambda_2 = a - bi.$$

λ_1 在复平面中的模为

$$|\lambda| = \sqrt{a^2 + b^2}$$

从图1—20易知

$$a = |\lambda| \cos \theta, \quad b = |\lambda| \sin \theta,$$

利用上式, 可将 λ 改写为

$$\lambda = |\lambda| (\cos \theta + i \sin \theta). \quad (1-78)$$

将 $\cos \theta + i \sin \theta$ 展成泰勒级数, 并将 $e^{i\theta}$ 展成泰勒级数, 不难证明

$$e^{i\theta} = \cos \theta + i \sin \theta$$

所以 λ 又可写为

$$\lambda_1 = |\lambda| e^{i\theta}. \quad (1-78')$$

相应的共轭复根为

$$\lambda_2 = a - bi = |\lambda| (\cos \theta - i \sin \theta) = |\lambda| e^{-i\theta}. \quad (1-79)$$

根据(1-77), 这两个特征根在解 x_{it} 的表达式中可写为

$$c_{i1}\lambda_1^t + c_{i2}\lambda_2^t \quad (1-80)$$

如果 λ_1 和 λ_2 为共轭复数且(1-80)之值为实数, 则系数 c_{i1} 和 c_{i2} 必为共轭复数。设 $\lambda_1 = re^{i\theta}$, $\lambda_2 = re^{-i\theta}$, $r = |\lambda_1| = |\lambda_2|$ 。设 $c_{i1} = s_1 e^{i\psi_1}$, $c_{i2} = s_2 e^{-i\psi_1}$, 则(1-80)可改写为

$$\begin{aligned} c_{i1}\lambda_1^t + c_{i2}\lambda_2^t &= s_1 r^t e^{i(\psi_1 + \theta t)} + s_1 r^t e^{-i(\psi_1 + \theta t)} \\ &= 2s_1 r^t \cos(\psi_1 + \theta t) \end{aligned} \quad (1-81)$$

从上式可看出, 解是否稳定由 r^t 决定, 若 $r = |\lambda_1| = |\lambda_2| < 1$, 则解是收敛的, 所以系统是稳定的; 若 $r > 1$, 解发散, 因此系统不稳定; 若 $r = 1$, 则解既不收敛也不发散。余弦函数的周期性说明解作周期振荡, 其振荡周期为 $2\pi/\theta$ 。因为对于 $\cos k$, 当 $k = 0, 2\pi, 4\pi, \dots$, $\cos k$ 之值相等, 若 $k = \theta t = 0, 2\pi, 4\pi, \dots$, $\cos \theta t$ 亦取等值, 也就是说, 当 $t = 0$,

$\frac{2\pi}{\theta}, \frac{4\pi}{\theta}, \dots$, $\cos \theta t$ 之值相等, 因此 $\cos \theta t$ 之振荡周期为 $2\pi/\theta$, 即频率为 $\theta/2\pi$ 。又

因为 $t = 0$ 时, $\cos \theta t = 1$, 而 $t = -\frac{\psi}{\theta}$ 时, $\cos(\theta t + \psi) = 1$, 所以 $\cos(\theta t + \psi)$ 超前于 $\cos \theta t$ $\frac{\psi}{\theta}$ 时间单位, 而 $\cos(\theta t + \psi)$ 与 $\cos \theta t$ 周期相同, 都为 $2\pi/\theta$ 。

§7 相图

如果在一个空间, 系统的运动能表示为互不相交的轨迹, 使得在不变的外部作用下, 系统的每一初始状态都唯一地确定系统的过程, 那么这种空间叫**相空间**, 这一空间的坐标叫**相坐标**, 表现系统运动的相轨迹族叫**相图**。相图实际上就是微分方程或差分方程之图解, 利用它可以不求解微分方程, 而对系统之动态特性(例如稳定性)进行定性分析。§3中的图

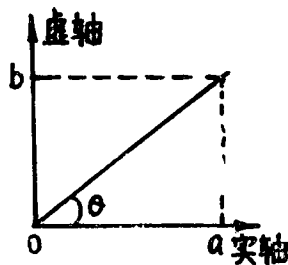


图 1—20

1—11, 图1—12, 图1—13, 图1—14, 图1—15, 图1—16都是相图之一种类型。

我们先考虑一对齐次微分方程

$$\begin{cases} \dot{y}_1 = a_{11}y_1 + a_{12}y_2 \\ \dot{y}_2 = a_{21}y_1 + a_{22}y_2 \end{cases} \quad (1-82)$$

其解为

$$\begin{cases} y_1 = k_{11}e^{\lambda_1 t} + k_{12}e^{\lambda_2 t} \\ y_2 = k_{21}e^{\lambda_1 t} + k_{22}e^{\lambda_2 t} \end{cases} \quad (1-83)$$

将上式中的 t 消去, 在 y_1 - y_2 坐标上我们可作出(1—82)之相图。现在我们要不求解(1—82), 而作出其相图。设(1—82)有如下具体形式:

$$\begin{cases} \dot{y}_1 = 5y_1 + 3y_2 \\ \dot{y}_2 = 2y_1 + 10y_2, \end{cases} \quad (1-84)$$

作图步骤如下:

(1) 在 y_1 - y_2 相空间作等倾线。若经过一条直线的所有相轨线都有相等的斜率(即不同的相轨线互相平行), 则此直线叫**等倾线**。我们先从(1—84)导出

$$\frac{\dot{y}_2}{\dot{y}_1} = \frac{dy_2}{dy_1} = \frac{2y_1 + 10y_2}{5y_1 + 3y_2} = k \quad (1-85)$$

令 $y_2 = ay_1$, 代入上式, 则

$$\frac{2y_1 + 10ay_1}{5y_1 + 3ay_1} = k, \quad a = \frac{5k - 2}{10 - 3k},$$

当斜率 k 为不同的实常数时, 等倾线为

$$y_2 = \frac{5k - 2}{10 - 3k}y_1 \quad (1-86)$$

例如, $k = \frac{2}{5}$ 时, 等倾线为 $y_2 = 0$, 当 $k = \frac{10}{3}$ 时, 等倾线为 $y_1 = 0$, 当 $k = \frac{5}{7}$ 时, 等倾线为 $y_2 = \frac{1}{5}y_1$, 这三条等倾线如图 1—21 所示。

(2) 从(1—84)可推出, 若 $\dot{y}_1 = 0$, 则有 $5y_1 + 3y_2 = 0$, 即 $y_2 = -\frac{5}{3}y_1$ 。当 $y_2 > -\frac{5}{3}y_1$ 时, $\dot{y}_1 > 0$, 相曲线之横坐标 y_1 随时间渐增, 当 $y_2 < -\frac{5}{3}y_1$ 时, $\dot{y}_1 < 0$, y_1 随时间渐减; 类似地, 若 $\dot{y}_2 = 0$, 则有 $y_2 = -\frac{1}{5}y_1$, 当 $y_2 > -\frac{1}{5}y_1$ 时, $\dot{y}_2 > 0$, 相曲线之纵坐标 y_2 随时间渐升,

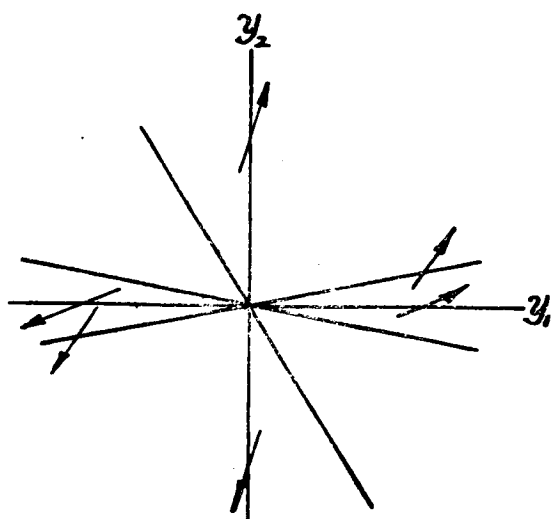


图 1—21

若 $y_2 < -\frac{1}{5}y_1$, 则 $\dot{y}_2 < 0$, y_2 随时间渐减。

上述两条直线将相空间划分为 4 个区域, 我们可根据以上分析判断相曲线在各个区域中箭头之大致指向, 如图 1—22 所示。图中的箭头说明, 这个系统是不稳定的, 在均衡点 (此例即原点) 外的任一点, 随着时间的推移, 系统越来越远离均衡状态。

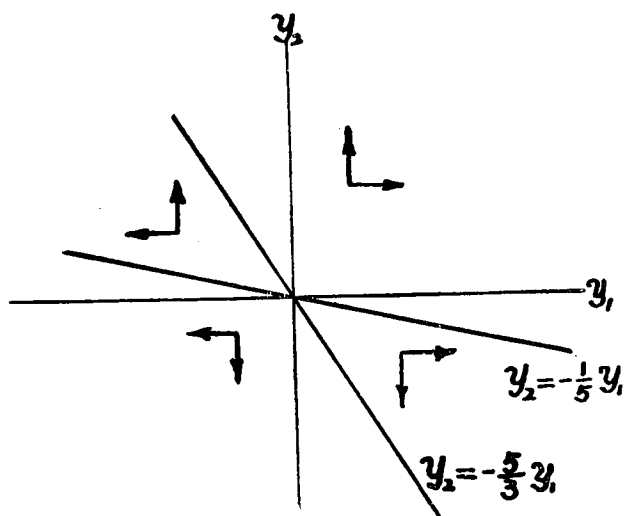


图 1—22

(3) 作相空间中相曲线之分界线。 y_1 - y_2 相空间中斜率与相曲线斜率相等的直线叫分界

线。设这条直线为

$$y_2 = ky_1,$$

将上式代入(1—85), 则有

$$k = \frac{2+10k}{5+3k},$$

解出

$$k = -\frac{1}{3}, 2,$$

所以分界线为

$$y_2 = -\frac{1}{3}y_1, y_2 = 2y_1.$$

假如相曲线有点在分界线上, 则由于相曲线斜率与直线斜率相等, 此点随着时间流逝只能在这条直线上运动, 而不会离开这条直线, 对于真正的曲线而言, 这当然不可能。所以相曲线与分界线没有交点。分界线上的等倾线就是分界线本身。

(4) 将上列三步给出的信息相结合, 我们可作出相图之定性分析略图, 如图 1—23 所示。

例 设有微分方程组

$$\begin{cases} \dot{y}_1 = -5y_1 + 3y_2 \\ \dot{y}_2 = -2y_1 + 10y_2 \end{cases}$$

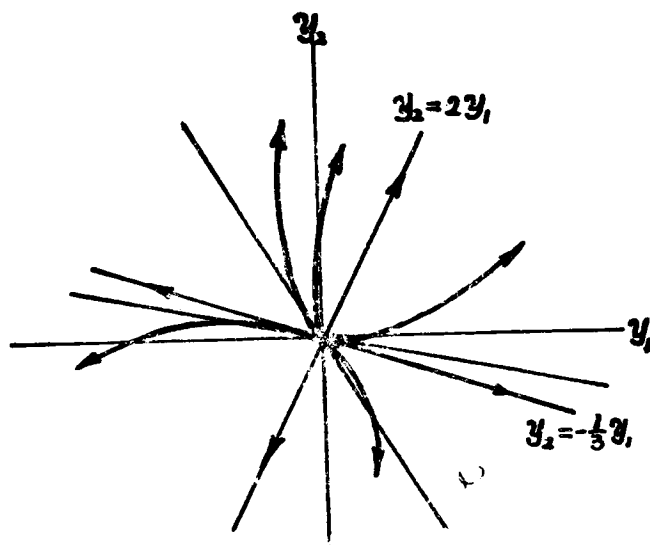


图 1—23

当 $\dot{y} = 0$ 时, 我们有 $y_2 = \frac{5}{3}y_1$

当 $\dot{y}_2 = 0$ 时，我们有 $y_2 = \frac{1}{5}y_1$ ，这两条直线将相空间分为 4 个区域，在各个区域，相曲线上箭头指向如图 1—24 所示。

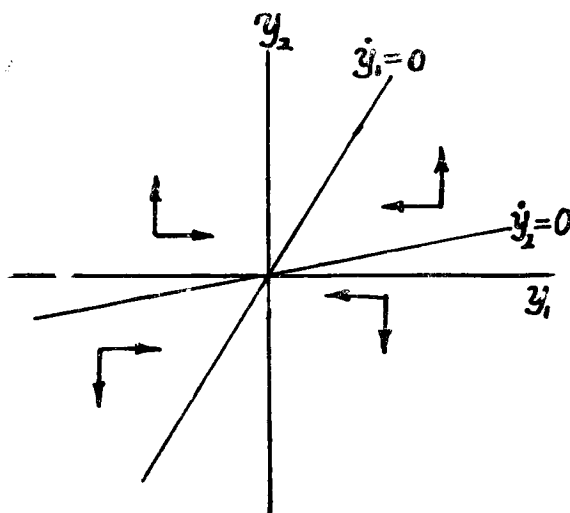


图 1—24

利用 $\frac{-2y_1 + 10ky_1}{-5y_1 + 3ky_1} = k$ ，解出 $k = 4.86, 0.14$ ，据此作分界线

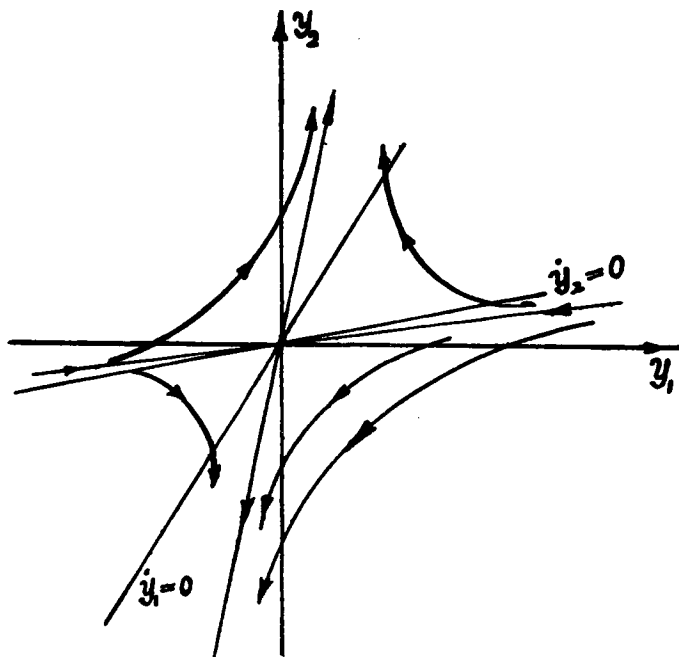


图 1—25

$$y_2 = 4.86y_1 \quad \text{和} \quad y_2 = 0.14y_1$$

根据以上信息作出图1-25中相图定性略图，从图中不难看出系统是不稳定的。

离散经济系统之相图作法从§3已可略知一二。另一种作法是以 y_t 为横坐标，以 y_{t+1} 为纵坐标。设有离散状态方程：

$$y_{t+1} = \frac{1}{2}y_t, \quad (1-87)$$

在 y_{t+1}, y_t 平面作直线 $y_{t+1} = \frac{1}{2}y_t$ 和 $y_{t+1} = y_t$ ，见图1-26。这两条直线的交点为

$$\Delta y_t = y_{t+1} - y_t = 0,$$

这就是系统之均衡状态。从初始状态 $y_t = y_0$ ，再取纵坐标 $y_2 = \frac{1}{2}y_1 = \frac{1}{4}y_0$ ，又取横坐标 $y_2 = \frac{1}{4}y_0$ ，再又取纵坐标 $y_3 = \frac{1}{2}y_2 = \frac{1}{8}y_0, \dots$ ，如此循环下去，不难看出这是一个没有振荡的收敛的正反馈过程。这一系统是稳定的。图1-27是 $y_{t+1} = 2y_t - 2$ 之相图，这个系统是不

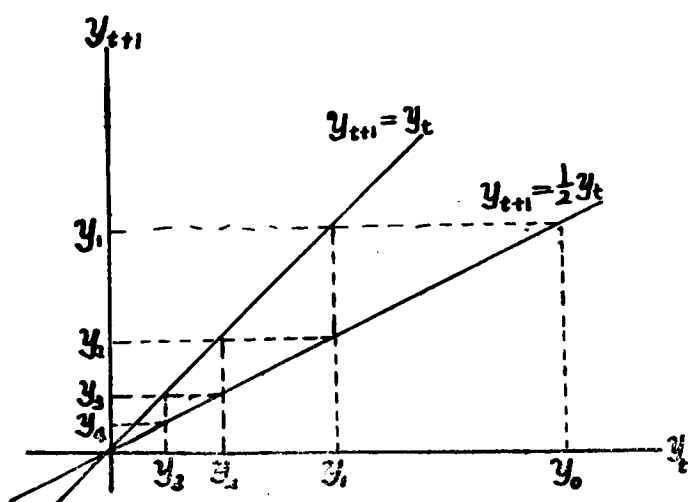


图 1-26

稳定的正反馈系统。而图1-28是 $y_{t+1} = -2y_t$ 之相图，这个系统是不稳定的负反馈，它表现为增幅振荡。图 1-29 是 $y_{t+1} = -\frac{1}{3}y_t + 12$ 相图，这个系统是稳定的负反馈，它是一个减幅振荡过程。

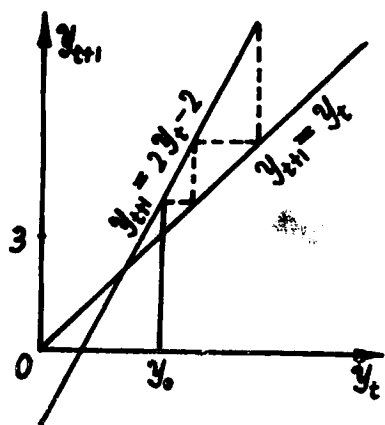


图 1—27

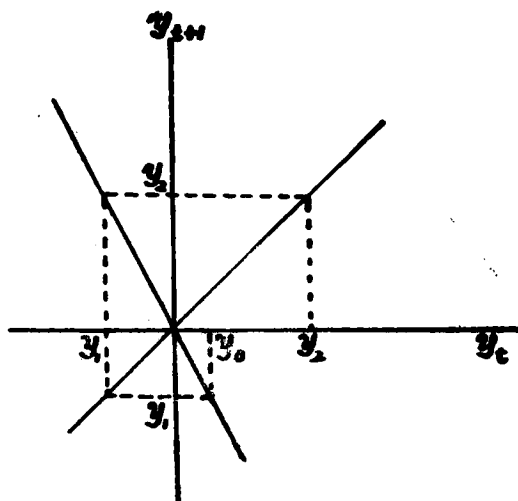


图 1—28

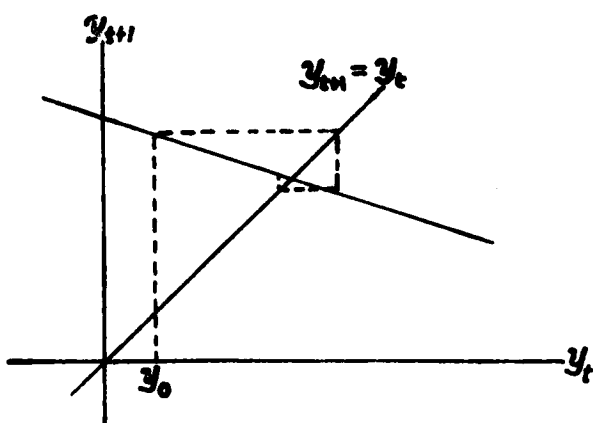


图 1—29

第二章 经济信息

§1 信息和熵

信息是种**差异度**的量度。经济系统中可分辨的状态越多,则其差异度越大,信息量也越大,可分辨的状态又叫**信息源**。离散信息源有一个含有若干**字符**的**字符表**。例如跳着变的价格之不同水平就是字符,而不同价格水平的集合就构成一个**字符表**。若干字符组成的序列叫**词**。如可能的价格水平有50种,则字符表中有50个字符,如果每天可能出现50种价格水平中的任意15种,则一天的价格构成15字符长的词。又例如某厂工资水平有10个级别,则字符表中有10个字符,而54人的车间之工资级别构成54字符长的词。字符表中字符的个数叫**码基**。字符所有可能出现的可重复排列就组成了所有可分辨的状态,因而决定了系统的差异度。例如某种产品计划价格只有5种,每个月允许变动两次价格,则一个月内,可能取的价格水平就是5因素的可重复2次排列,即这个价格系统有 $5^2=25$ 种可分辨的状态,其差异度为25。此系统有5个字符,即码基为5,每月的价格水平是2字符长的一个词。若一个信源由若干字符或词组成一个序列,与一消息对应,则可一般地称之为**信源系列**。

但是若两个经济系统都有两种可能的状态,而在第一个系统中每种状态出现的概率分别为 $\frac{1}{2}, \frac{1}{2}$,在第二个系统中两种状态出现的概率分别为 $\frac{1}{1000}, \frac{999}{1000}$,这时两个系统的信息量是大不相同的。不难理解,对于两种状态出现的概率相差很悬殊的系统来说,我们更容易预测其实际出现的状态,也就是说这种系统的确定性更大,即不定性很小。对于不定性很小的系统,则排除这种不定性后,得到的信息也不多,而不定性越大,则排除其不定性后得到的信息量也越大。所以我们用反映不定性的熵来描述**信息容量**,而用反映确定性的**负熵**来描述**信息量**。熵是利用*i*种状态($i=1, \dots, N$)出现的概率 p_i 来定义的。熵由下列定义式给出。

$$H = - \sum_{i=1}^N p_i \log p_i, \quad \sum_{i=1}^N p_i = 1, \quad (2-1)$$

其中 $\log$ 可以是以 e 为底的自然对数,也可以是以2为底的对数。若以 e 为底,则算出的熵之单位为nit,若以2为底,算出的熵之单位为bit,此式就是信息容量测度公式,信息量由负熵定义,即信息量为

$$-H = \sum_{i=1}^N p_i \log p_i \quad (2-2)$$

此书除作特殊说明外,一般以bit为信息量单位。不难验证,若所有状态出现的概率都相等,即 $p_i = \frac{1}{N} (i=1, \dots, N)$,则

$$H = -N \left(\frac{1}{N} \log \frac{1}{N} \right) = \log N \quad (2-3)$$

其中 N 为可能出现的状态数, 即差异度。当 $N = 2$ 时, 不难将(2-1)式对 p_i 求导数, 且令其等于 0, 解出 $p_1 = p_2$, 因而证明了当各种状态出现的概率都相等时, 不定性(熵) H 达到极大值。这一证明不难推广到 N 为任意正整数的情形。

§2 自由度、剩余度和信息效率

当经济系统中各种可能状态的验前概率都相等时, 系统达到其最大熵 $H_{\max}$ 。实际上由于系统受到各种约束, 验前概率不可能完全相等, 因此经济系统的实际熵总是小于其最大熵 $H_{\max}$ 。这时我们可以用状态较少, 但其验前概率完全相等的一个系统获得与状态较多而受束缚的系统之信息容量同样大的熵。不受束缚之经济系统的自由度等于其码基, 而受约束的经济系统之自由度小于其码基。实际熵与最大熵之比叫**信息效率**, 设其为 E , 则

$$E = \frac{H}{H_{\max}} \quad (2-4)$$

E 满足 $0 \leq E \leq 1$ 。我们又设 R 为**剩余度**,

$$R = 1 - E = \frac{H_{\max} - H}{H_{\max}} \quad (2-5)$$

信息效率反映了经济竞争的水平。竞争越平等, 则各种经济要素组合形式出现的验前概率就越接近。这种“机会均等”表现为信息效率 E 很高。例如三个使用机床的企业(用 A, B, C 表示)都要购买三个机床厂(用字母排序的三个位置表示)生产的同类设备, 可能的状态是:

AAA	AAB	CBB
BBB	AAC	BCC
CCC	BBC	ACC
ABC	BBA	ABA
ACB	CCA	ACA
BAC	CCB	BAB
BCA	CAA	BCA
CAB	BAA	CAC
CBA	ABB	CBC

共有 $3^3 = 27$ 种状态。若完全自由选购和自由推销, 即各种状态验前概率完全相等, 则最大熵

$$\begin{aligned} H_{\max} &= - \sum_{i=1}^{27} \frac{1}{27} p_i \log p_i \\ &= \log 27 = 4.755 \text{ bit.} \end{aligned}$$

若对于买方选择卖方或对卖方选择买方有某种限制, 则有些状态出现的概率等于 0, 各企业购买某种商品的机会不均等, 实际熵就会小于其最大熵。例如 27 种状态中由于统购统销的约束, 有 10 种状态根本不能出现, 其概率为 0, 有 7 种状态各以 0.037 的概率出现。还有 4 种

状态各以0.074的概率出现,则实际熵

$$\begin{aligned} H &= -10 \times 0 \log 0 - 10 \times 0.074 \log 0.074 - 7 \times 0.037 \log 0.037 \\ &= 2.62 \text{ bit}, \end{aligned}$$

则这个经济系统的信息效率

$$E = \frac{H}{H_{\max}} = \frac{2.62}{4.755} = 0.551$$

剩余度

$$R = 1 - E = 0.449.$$

§3 编码

用一种事物可分辨的状态集对应于另一事物的可分辨的状态集,我们从前一事物的差异度了解后一事物的差异度叫**消息**。体现消息的物质载体叫**信号**。例如货币体现的价格信号、利润信号,我们从这些信号得到有关经济效果的消息。而发生信号传输的系统叫**信道**。

把事物的状态集 $X = (x_1, \dots, x_m)$ 变成用信号表示的消息集 $X' = (x'_1, \dots, x'_n)$ 要经过一个一一对应的变换,这个变换由变换算子 T 产生, T 使得

$$X' = TX \quad (2-6)$$

把任何状态变换成消息的算子 T 叫**码**,而这种变换运算叫**编码**。逆变换 T^{-1} 叫**解码** 或**译码**。编码和译码都要是一一对应的变换,才能保持通讯中信息量不会减少。

发报机发出的无线电波与英文字母这两个系统之间在物质形态上毫无共同之处,但在差异这一点上,各系统却有其共通之处。如果人为地将各系统的差异进行一一对应的约定,就可以进行编码、解码和信息传输,经济系统中的经济效果、价格、收入也是不同的系统,但对这些系统的差异进行一一对应的约定,也可以进行编码、解码和信息传输。

例如把一个企业的经济活动效果与厂长的收入之间建立一一对应关系这就是编码。厂长根据收入增减判断经济效果的高低就是解码。如果经济效果有120种不同的水平(其差异度为120),而容许的收入水平只有3种(差异度为3),这时编码解码就不可能是一一对应的,这就使人们无法从自己的收入得到信息源的真实消息,因而产生了信息损失。

对于定长信源序列而言,编码和解码中不产生信息损失的必要条件由信源序列编码定理给出。

定理 设信源每一状态的平均熵为 $H(u)$, 信源序列为定长 n , 而构成信号的码字个数为 M , 信号序列为定长 m , 则不产生信息损失,即保证编码、解码一一对应的必要条件为

$$\frac{m}{n} \log M \geq H(u) \quad (2-7)$$

其中 $H(u) = -\sum_k p_k \log p_k$, p_k 为状态 k 出现之概率, n 种状态的熵为 $-\sum_k n p_k \log p_k$, 而通道最大熵为 $\log M^m$, 因此信源状态能与通道状态能一一对应的必要条件为 $\log M^m \geq$

$-\sum_k n p_k \log p_k$, 据此可导出(2-7)。因而定理得证。

信源序列编码定理说明若经济效果之信源序列熵 $H(u)$ 很大时, 则价格和收入的可选择水平 M 要足够多, 才能使经济信息传输中不致产生错失。如果将序列长度与时间挂上勾, 则定理意味着, 若经济效果信源在短时间产生了很长的序列 (n 很大), 则价格和收入系统也必须在同样时间内产生足够多的不同水平 (m 足够大), 才能使经济信息的编码解码保持一一对应。也就是说当经济效果迅速变化时, 价格和收入水平也要跟着迅速变化才能避免信息损失。换言之, 传输通道的自由度和熵要足够大才能使入口处熵很大的消息不失真地得到传输。如果 $H(u)$ 很大, M 很小 (即可选择的价格、收入水平很少), 要使传输不失真, 则要求 m 很大, 即价格、收入变化快。若收入, 价格变化速度不可能很快, 那么 m 很大就表现为信息不失真地传输需要很长时间, 即产生很大时滞, 如上一章所讨论的, 时滞正是反馈调节中产生振荡 (左右摇摆) 的原因。

例 设价格差异度 $M=4$, 经济效果熵 $H(u)=15.7$ bit, 每单位时间经济效果产生的信源序列长度为 $n=5$, 而单位时间价格序列长度 $m=5.5$, 问在这种价格体制下, 经济信息是否能依赖纯经济方式不失真地得到适时传输?

根据信源序列编码定理, 要求 $\frac{m}{n} \log M \geq H(u)$, 由于 $\frac{m}{n} \log M = \frac{5.5}{5} \log 4 = \frac{11}{5}$, 而 $H(u)=15.7$, 所以必要条件不具备, 可见在这种价格体制下不能完全靠收入、价格等物质刺激机制来传输经济效果信息, 要使经济效果信息不失真地适时传输, 必须改革现行价格体制, 否则有必要借助行政方式。

§4 经济信息传输

经济信息传输中的信息损失是利用条件熵概念来描述的。设有经济效果信号 $x_j (j=1, \dots, n)$ 和价格或收入信号 $y_i (i=1, \dots, m)$, 则收入信号 y_i 出现时, 相应经济效果信号 x_j 出现的条件概率为 $p(x_j|y_i)$, 借此我们可以计算条件熵, 条件熵越大, 说明经济效果信号与收入信号之间的一一对应之不确定性越大, 因为熵是种不确定性的量度。条件熵大也就是说“干好干坏”一个样, 即经济效果与收入之间的一一对应很差。条件熵小说明经济效果与收入之间的一一对应很严密。传输系统的条件熵 $H(X|Y)$ 是将 y_i 出现时 x_j 出现之熵用 y_i 出现的概率加权平均, 即

$$H(X|Y) = -\sum_i \frac{m}{n} \sum_j \frac{n}{m} p(y_i) p(x_j|y_i) \log p(x_j|y_i) \quad (2-8)$$

这就是传输中的信息损失。信息源的信息容量为

$$H(X) = -\sum p(x_j) \log p(x_j) \quad (2-9)$$

因此信道中实际传输的信息量为

$$\begin{aligned} I(X, Y) &= H(X) - H(X|Y) \\ &= -\sum_j p(x_j) \log p(x_j) + \sum_i \frac{m}{n} \sum_j \frac{n}{m} p(y_i) p(x_j|y_i) \log p(x_j|y_i) \end{aligned} \quad (2-10)$$

$I(X, Y)$ 叫通道的信息传输速率。 $I(X, Y)$ 与信源信息容量 $H(X)$ 之比叫传输效

率, 令其为 P , 则

$$P = \frac{I(X, Y)}{H(X)} = \frac{H(X) - H(X|Y)}{H(X)}$$

$$= 1 - \frac{H(X|Y)}{H(X)} \quad (2-11)$$

P 又是通道传输信息可靠性的量度, 所以又可将 P 叫做 **可靠性**, 它通常满足 $0 < P < 1$. 而通道传输信息的错失率定义为

$$Q = 1 - P = \frac{H(X|Y)}{H(X)} \quad (2-12)$$

Q 又是通道传输信息不可靠性的量度, 它通常亦满足 $0 < Q < 1$.

例 下表列出当收入出现 y_1, y_2, y_3 时, 实际曾发生的经济效果为 x_1, x_2, x_3 的条件概率

收入水平	y_1			y_2			y_3		
经济效果水平	x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
$p(x_j y_i)$	$\frac{1}{3}$	0	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{9}$	0	$\frac{1}{18}$	0	$\frac{1}{3}$

又设 $p(x_1) = p(x_2) = p(x_3) = \frac{1}{3}$

$$p(y_1) = \frac{1}{4}, \quad p(y_2) = \frac{1}{2}, \quad p(y_3) = \frac{1}{4}$$

则信息源的熵

$$H(X) = -3\left(\frac{1}{3} \log \frac{1}{3}\right) = \log 3 = 1.585 \text{ bit}$$

条件熵

$$\begin{aligned} H(X|Y) &= p(y_1) \sum p(x_j|y_1) \log(x_j|y_1) \\ &\quad + p(y_2) \sum p(x_j|y_2) \log(x_j|y_2) \\ &\quad + p(y_3) \sum p(x_j|y_3) \log(x_j|y_3) \\ &= \frac{1}{4} \left(-\frac{1}{3} \log \frac{1}{3} - 0 \log 0 - \frac{1}{9} \log \frac{1}{9} \right) \\ &\quad + \frac{1}{2} \left(-\frac{1}{18} \log \frac{1}{18} - \frac{1}{9} \log \frac{1}{9} - 0 \log \frac{1}{3} \right) \\ &\quad + \frac{1}{4} \left(-\frac{1}{18} \log \frac{1}{18} - 0 \log 0 - \frac{1}{3} \log 0 \right) \\ &= 0.6105 \text{ bit}, \end{aligned}$$

这就是传输中的信息损失。

信息传输速率

$$I(X, Y) = H(X) - H(X|Y)$$

$$= 1.585 - 0.6105$$

$$= 0.9745 \text{ bit,}$$

$$P = \frac{I(X, Y)}{H(X)} = \frac{0.9745}{1.585} = 0.6148,$$

通道错失率为

$$Q = 1 - P = 1 - 0.6148 = 0.3852.$$

§5 大系统经济信息传输效率和可选择并联耦合

在第一章中我们讨论了分工协作的经济系统 实质上是既有串联耦合又有并联耦合的系统, 分工越发达, 社会化水平越高, 则串联耦合的工序、部门越多。我们可以将串联的工序和部门视为大系统中的部件, 每两个串联的部件之间有一个经济信息传输通道, 物资从前一部件流向后一部件, 而经济信息从后一部件反馈传输到前一部件, 即物流与货币信号流同时反向流通。设每个通道的信息传输效率平均为 p , 而大系统由 n 个部件串联组成。根据第一章(1—5), 即串联系统的总传输系数等于子系统传输系数之积, 所以分工协作经济系统总的信息传输效率为

$$P = p^n \quad (2-13)$$

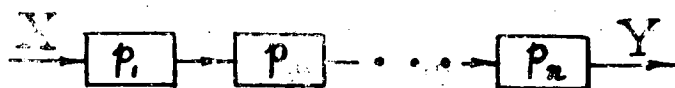


图 2-1

由于 $0 < p < 1$, 所以 n 越大, 则 p^n 越小。即分工社会化水平越高, 经济大系统总的信息传输效率下降, 而经济信息在传输中的错失率上升, 要去掉经济发达的这一负作用必须把**可选择并联耦合**引入经济系统。可选择并联耦合在经济系统中就是在有一定生产过剩的情况下同行的多极平等竞争。可选择并联耦合与一般并联耦合的差别在于耦合中的并联元件区分为**工作元件**和**备用元件**。在经济系统中备用元件就是与过剩的生产能力联结的通道, 工作元件就是与竞争中获胜的生产能力联结的通道, 当充当工作元件的通道失效时, 备用元件就会来代替它。所以对于初始经济信息 X_i 只有当所有并联的通道(包括工作元件和备用元件)都将其传输错了时, 这个信

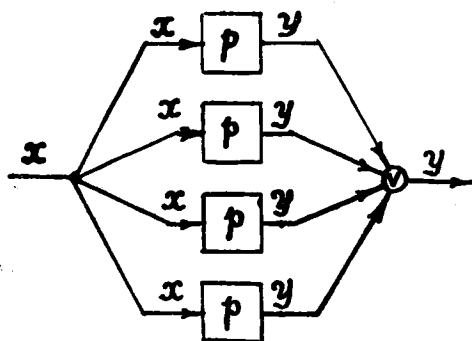


图 2-2

息才会在传输中产生错失。设每个信道的平均错失率为 q ，而每个工序或部门中可选择并联的通道共有 m 个，则每一工序错失率为 q^m ，而此工序（或部门）的传输效率为

$$p = 1 - q^m \quad (2-14)$$

如果共有 n 个工序或部门串联耦合在一块，总的传输效率

$$\begin{aligned} P &= p^n \\ &= (1 - q^m)^n \end{aligned} \quad (2-15)$$

如果每个通道(元件)的传输效率且每一工序可选择并联的通道数不等，则我们有图2—3

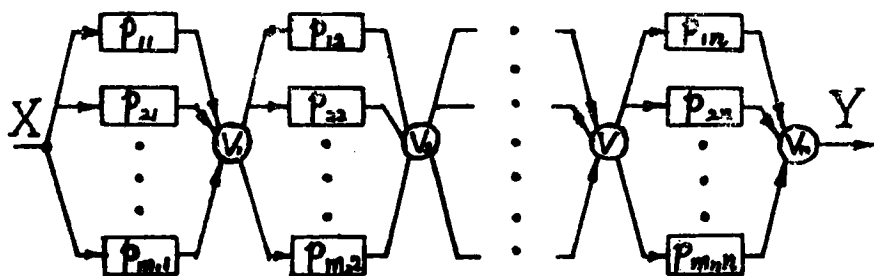


图 2—3

的一般系统，其中第 j 个工序(或部门)中的第 i 个通道之错失率为 q_{ij} ，传输效率为 $p_{ij} = 1 - q_{ij}$ ， j 部门中并联的通道有 m_j 个。图中 V 为选言算子，它在所有备用元件中挑选工作元件，当工作元件失效时，用备件元件代替之。这个工序（或部门）总的错失率为

$$\prod_{i=1}^{m_j} q_{ij} = \prod_{i=1}^{m_j} (1 - p_{ij}) \quad (2-16)$$

此工序的信息传输效率为

$$p_j = 1 - \prod_{i=1}^{m_j} q_{ij} \quad (2-17)$$

而 j 个 ($j = 1, \dots, n$) 工序（或部门）串联在一块的总传输效率为

$$\begin{aligned} P &= \prod_{j=1}^n p_j \\ &= \prod_{j=1}^n (1 - \prod_{i=1}^{m_j} q_{ij}) \end{aligned} \quad (2-18)$$

若可选择并联耦合信道的平均错失率为 q ，且各工序元并联耦合元件平均数为 m ，则 n 个串联工序的总信息传输效率由(2—15)给出，将其等式两边求对数，则

$$\begin{aligned} \log P &= n \log (1 - q^m) \\ 1 - q^m &= \log^{-1} \left(\frac{\log P}{n} \right) = P \\ q^m &= 1 - P^{\frac{1}{n}}, \end{aligned} \quad (2-19)$$

$$q = (1 - P^{\frac{1}{n}})^{\frac{1}{m}}. \quad (2-20)$$

将(2-19)再求一次对数, 则有

$$\begin{aligned} m \log q &= \log (1 - P^{\frac{1}{n}}) \\ m &= \frac{\log (1 - P^{\frac{1}{n}})}{\log q} \end{aligned} \quad (2-21)$$

我们可以利用(2-20)在已知串联的工序数 n , 和并联耦合的通道数 m 的情况下, 计算通道平均错失率的上限。而(2-21)可用来在已知通道平均错失率 q 及串联工序数 n 的情况下, 计算使大系统经济信息传输效率达到 P 时各工序要并联多少个可选择通道。

例 设某种商品从原材料到半成品, 再变成成品, 再卖给消费者, 共要经过半成品生产厂, 成品生产厂和商业部门三道工序。每道工序的信息反馈传输效率平均为0.7。消费者使用此种商品的效果通过购买时销售量对价格的反应, 将信息传给初始生产者, 总的传输效率为

$$0.7 \times 0.7 \times 0.7 = 0.343.$$

这个系统信息传输总效率极低, 生产者得到的经济效率消息之可信度只有0.343。如果每道工序有3个平等竞争的企业, 则反馈信号在每道工序有3个可选择并联耦合通道, 每个通道的平均错失率为 $q = 1 - p = 1 - 0.7 = 0.3$, 将 $m = 3$, $n = 3$, $q = 0.3$ 代入(2-15), 则得到

$$P = (1 - 0.3^3)^3 = (1 - 0.027)^3 = 0.973^3 = 0.921$$

若要求 $p = 0.976$, 则将 $n = 3$, $m = 3$, $P = 0.976$ 代入(2-20), 算得通道错失率之上限

$$q = (1 - P^{\frac{1}{n}})^{\frac{1}{m}} = (1 - 0.976^{\frac{1}{3}})^{\frac{1}{3}} = 0.2.$$

即每个通道传输效率至少应为0.8。

若通道错失率仍为0.3, 要使总传输效率达到0.976, 则要求可选择并联耦合的企业数

$$\begin{aligned} m &= \frac{\log (1 - P^{\frac{1}{n}})}{\log q} = \frac{\log (1 - 0.976^{\frac{1}{3}})}{\log 0.3} \\ &= 4(\text{个}). \end{aligned}$$

§6 控制的限度

经济控制能力并不受物质形式 差别的局限, 我们可以用电信号 控制机械、原子弹和飞船, 也可以用声音控制经济组织活动; 经济控制能力也不受能量限制, 我们可以用微弱的电子信号控制氢弹爆炸或强大的电网, 也可以用微弱的声音控制壮观的经济建设活动。但是控制能力却受到信息传输效率的限制。在经济系统中信息传输效率决定了消费者控制生产者的能力的限度。正如上节所述, 如果分工很发达, 而平等竞争的企业不多, 则总的信息传输效率下降, 消费者选择生产者的熵不大, 这就使消费者控制生产者的能力受到限制。

设 $H(X)$ 为消费者进行选择的熵, $H(Y)$ 为只有生产者对结局施加影响时结局的不定程度, $H(Y|X)$ 为消费者对结局施加影响后, 生产者使结局产生的不定度, 则消费者的选

择、控制使结局不定性减少的量为

$$H(Y) - H(Y|X) = I(X, Y) \quad (2-22)$$

不定性减少的量也就是实际传输的信息量，所以上列等式成立。而根据(2-10)

$$I(X, Y) = H(X) - H(X|Y),$$

将(2-22)代入上式，则

$$H(Y) - H(Y|X) = H(X) - H(X|Y) \quad (2-23)$$

所以

$$H(Y|X) = H(Y) - H(X) + H(X|Y) \quad (2-24)$$

此式表示在消费者和生产者的双重作用下，结局的实际不定性，提高消费者控制生产者的能力（即改变“皇帝女儿不愁嫁”的局面），就要使 $H(Y|X)$ 减小。 $H(Y|X)$ 是在消费者对结局施加影响的条件下，生产者对结局施加影响后，使结局产生的不定性（条件熵）。要使 $H(Y|X)$ 缩小，则要求增大消费者进行选择熵 $H(X)$ ，其信息容量决定了 对控制能力的限制。另一方面也可缩小生产者使结局产生的熵 $H(Y)$ 以及生产者对 结局施加影响熵 $H(X|Y)$ 。

第三章 变分法和最大值原理

§1 问题的提起

在初等微积分中我们学过中值定理，并且十分熟悉求函数极值的方法。在初等微积分中，求出的函数极值点往往是不变常数，例如若有需求函数：

$$x = b - ap \quad (3-1)$$

总销售额，即收益是销售量 x 与价格 p 之乘积，即总收益 R 为

$$\begin{aligned} R &= xp \\ &= (b - ap)p \\ &= bp - ap^2 \end{aligned} \quad (3-2)$$

又设费用函数为

$$C = cx + A = c(b - ap) + A \quad (3-3)$$

其中 c 为单位产品之可变费用， A 为固定费用。

于是利润 π 为

$$\begin{aligned} \pi &= R - C \\ &= bp - ap^2 - c(b - ap) - A \end{aligned} \quad (3-4)$$

利润最大化的一阶条件为

$$\frac{d\pi}{dp} = b - 2ap + ac = 0 \quad (3-5)$$

解出最优价格为

$$p^* = \frac{b + ac}{2a} \quad (3-6)$$

由于 a, b, c 都是用经济计量学算出的常数，所以利润极大化的最优价格 p^* 也是常数。

如果需求量不但与价格有关，而且与价格变化的速度有关，那么问题就复杂得多了。价格上涨得快时，很多消费者估计以后购买此种商品会需更多货款，所以即使不需要此种商品，也要抢着购买，贮存起来以备将来使用。反过来，当价格下跌很快时，很多消费者估计以后购买此种商品会更占便宜，所以尽量将购买计划推迟，因而使当前需求量减少，可见需求量正比于价格变化速率。这与当前价格对需求量的影响正好相反。当前价格越高，需求量越少，即需求量反比于当前价格。所以我们调价时，就不能光考虑将价格调为最优价格，而且要考虑调节价格的行为本身造成消费者之反应对决策过程的影响。因此我们的目的不是

求解静态最优价格，而是求解动态最优价格，即最优价格变化的时间轨线，以使所考虑的整个时期内，总销售利润最大。

例如，若有如下需求函数

$$x(t) = b - ap(t) + \alpha \frac{dp(t)}{dt} \quad (b, a, \alpha > 0) \quad (3-7)$$

于是收益函数为

$$\begin{aligned} R(t) &= x(t)p(t) \\ &= \left[b - ap(t) + 2\alpha \frac{dp(t)}{dt} \right] p(t) \end{aligned} \quad (3-8)$$

若费用函数仍为

$$\begin{aligned} C(t) &= cx(t) + A \\ &= c \left[b - ap(t) + \alpha \frac{dp(t)}{dt} \right] + A \end{aligned} \quad (3-9)$$

则利润函数为

$$\begin{aligned} \pi(t) &= R(t) - C(t) \\ &= \left[b - ap(t) + \alpha \frac{dp(t)}{dt} \right] p(t) - c \left[b - ap(t) + \alpha \frac{dp(t)}{dt} \right] - A \end{aligned} \quad (3-10)$$

上式可记为更一般的

$$\pi(t) = \pi[p(t), \dot{p}(t), t] \quad (3-11)$$

于是 $0 \sim T$ 时域内的总利润为

$$I = \int_0^T \pi[p(t), \dot{p}(t), t] dt \quad (3-12)$$

我们的目的是在可能的无穷多个 $p(t)$ 轨线中，找出一个 $p^*(t)$ ，以使 I 极大化， I 是 $p(t)$ ， $\dot{p}(t)$ ， t 的一个泛函，即它是函数的函数。

可见变分法是将静态极值问题推广到动态极值。

当然变分法不一定包括时间因素。例如图3—1中的曲线 L 绕 x 轴旋转形成的曲面面积，随曲线形状不同而大相庭径。如果要求出一个曲线函数 $y = f(x)$ ，以使曲面面积最大，则这就是一个变分问题。图中的微小弧段长度为

$$ds = \sqrt{1 + y'^2} dx \quad (3-13)$$

而垂直于 x 轴横截这一曲面可得一个圆，其周长为 $2\pi y$ ，于是以 ds 为宽的微小环带面积为

$$2\pi y ds = 2\pi y \sqrt{1 + y'^2} dx \quad (3-14)$$

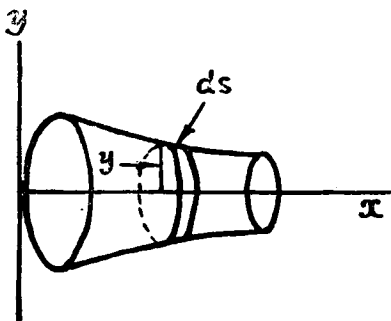


图 3—1

从 x_0 到 x_1 之曲面面积为

$$I = \int_{x_0}^{x_1} 2\pi y \sqrt{1+y'^2} dx \quad (3-15)$$

我们的目的就是要在无穷多种可能的 $y = f(x)$ 中选择一个曲线函数

$$y^* = f^*(x)$$

以使泛函 I 极大化。这个问题中没有时间因素，我们可以将最优价格问题中的 t 看作是 x ，于是我们将变分问题写成更为一般的形式：

从 $y = f(x)$ 中选择 $y^* = f^*(x)$ ，以使目标泛函

$$I[y] = \int_{x_0}^x F(x, y, y') dx \quad (3-16)$$

取极值（极大或极小）。

§2 基本的数学概念

定义 若所有 $y(x)$ 附近的函数 $z(x)$ 使泛函值之差 $I[z] - I[y] \leq 0$ ，则函数 $y(x)$ 是极大化函数；若所有 $y(x)$ 附近的函数 $z(x)$ 使泛函值之差 $I[z] - I[y] \geq 0$ ，则函数 $y(x)$ 是极小化函数。

泛函 $I[y]$ 的连续性 若对于任一 $\varepsilon > 0$ ，有 $\delta > 0$ ，使得对于 $|y_0 - y| < \delta$ ， $|I[y] - I[y_0]| < \varepsilon$ ，则 $I[y]$ 在 y_0 连续。

回顾初等微积分中关于函数连续性的定义，不难看出泛函连续性的定义不过是函数连续性的推广，除了将函数换上泛函，将点变成函数外，以上定义并没有新增什么东西。

点间距离 先考虑二维空间中的点间距离。设二维空间两点由二维矢量表示，例如 $A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ ， $B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ ，矢量的模由其长度确定，例如 $\|A\| = \sqrt{a_1^2 + a_2^2}$ ， $\|B\| = \sqrt{b_1^2 + b_2^2}$ ，矢量 A 与矢量 B 的距离为

$$D = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

根据矩阵代数

$$A - B = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \end{bmatrix},$$

所以

$$D = \|A - B\|.$$

即两点间的距离为两矢量差之模。这一定义可推广到多维空间。

点集间距离 考虑作为矢量集合的矩阵，给定一个 $2 \times n$ 矩阵 $A = [A_1, \dots, A_n]$ ，其中 $A_i =$

$$\begin{bmatrix} a_{1i} \\ a_{2i} \end{bmatrix}. \text{ 矩阵 } A \text{ 的模定义为}$$

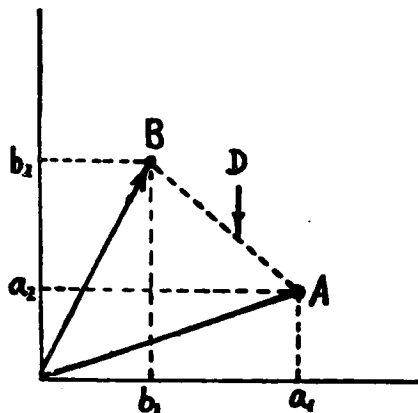


图 3-2

$$\|A\| = \max[\|A_1\|, \dots, \|A_n\|],$$

即矩阵的模等于其中模最大的矢量之模。例如对于 $2 \times n$ 矩阵 A , A 中的 n 列是二维空间中的 n 个矢量, $\|A\|$ 由 n 个矢量中最长者之模表示。若 B 亦为 $2 \times n$ 矩阵, 其 n 列也可表为二维空间之 n 个矢量, 因此可定义 A 矩阵到 B 矩阵离为

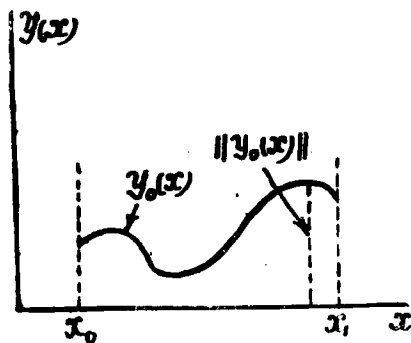
$$\|A - B\| = \max[\|A_1 - B_1\|, \dots, \|A_n - B_n\|],$$

即两矩阵的距离等于两矩阵相对应列之间的最大距离。

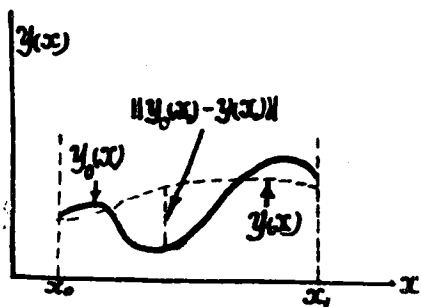
函数也是点集, 所以测度函数的模及函数间的距离可用与上类似的方法。

函数之间的距离 一连续函数 $y_0(x)$ 在某一区域: $x_0 \leq x \leq x_1$ 的模定义为函数在这一区域的最大值:

$$\|y_0(x)\| = \max_{x_0 \leq x \leq x_1} |y_0(x)|.$$



a



b

图 3-3

这实际上是函数上的点到 x 轴的最大距离。两函数在这一区域的距离为两函数上点间的最大距离, 即

$$\|y_0(x) - y(x)\| = \max_{x_0 \leq x \leq x_1} |y_0(x) - y(x)|.$$

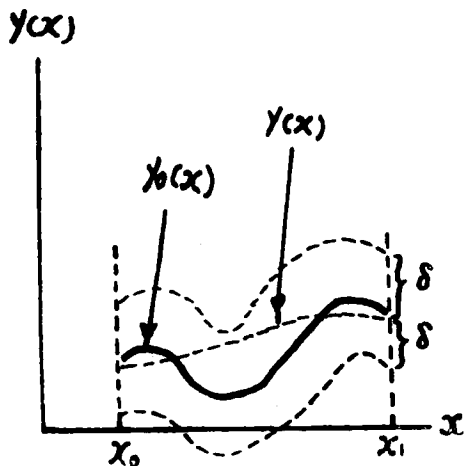


图 3-4

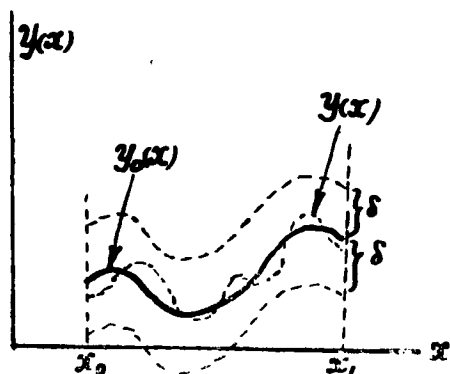


图 3-5

若 y 之图形在一个以 $y_0(x)$ 为中线的 2δ 宽的带状区域内, 则 $y_0(x)$ 与 $y(x)$ 之间的距离不会超过 δ 。我们说 $y(x)$ 处于 $y_0(x)$ 的 δ 邻域内, 但是用以上函数间距离的定义来描述两个函数之间的接近有不尽妥当之处, 图3--5中的两个函数虽然都在 2δ 的带状区域内, 但一条曲线十分平滑, 一条曲线凸凹不平, 这两条曲线不管 δ 如何小, 但只要 δ 不等于 0, 总是不能将二者视为互相接近的, 因为二者的一阶导数大相径庭。 $y'_0(x)$ 在 1 和 -1 之间, 而 $y(x)$ 的切线在某些地方几乎是垂直的, 这时 $y'(x)$ 几乎为 $+\infty$ 或 $-\infty$ 。若有目标泛函

$$I[y] = \int_{x_0}^{x_1} F(x, y, y') dx,$$

由于其中包含 y' , 在这种情况下 $I[y]$, 就会出现不连续的情况, 若目标泛函是使曲线弧长最小, 则 $y(x)$ 虽然处于 $y_0(x)$ 的 δ 邻域内, 但 $y(x)$ 比 $y_0(x)$ 长得多, 于是对于泛函 $I[y]$, 即使 $\|y_0(x) - y(x)\| < \delta$, 并不必定有 $|I[y] - I[y_0]| < \varepsilon$, 即 $I[y]$ 并不必定连续, 所以函数的模有必要考虑引入函数之导数, 例如函数的模为

$$\max_{x_0 \leq x \leq x_1} |y(x)| + \max_{x_0 \leq x \leq x_1} |y'(x)|$$

还可以将高阶导数包括在模的定义中, 我们用脚码表示模中包含的导数之最高阶数, 例如:

$$\|y\|_0 = \max |y(x)|$$

$$\|y\|_1 = \max |y(x)| + \max |y'(x)|$$

$$\|y\|_2 = \max |y(x)| + \max |y'(x)| + \max |y''(x)|$$

.....

$$\|y\|_n = \max |y(x)| + \max |y'(x)| + \dots + \max |y^{(n)}(x)|$$

同样, 函数之间的距离也可包含高阶导数, 也可用脚码表示其中包含的导数之最高阶数, 不过一般包含一阶导数就足够了。即函数间距离定义为:

$$\|y_0(x) - y(x)\| = \max |y_0(x) - y(x)| + \max |y'_0(x) - y'(x)|.$$

根据如此定义的函数间距离, 则处于 $y_0(x)$ 的 δ 邻域中的 $y(x)$ 就不可能是图3—5那种形状迥异的曲线啦, 于是也避免了目标泛函 $I[y]$ 的不连续问题。据此我们定义 $y_0(x)$ 的 δ 邻域为

$$\|y_0(x) - y(x)\|_1 < \delta.$$

有了这个定义我们就能根据连续性定义准确地判断泛函 $I[y]$ 的连续性了。当 δ 充分小时, $y(x)$ 趋近于 $y_0(x)$ 。如上定义的一阶趋近在几何上可用三维 x, y, y' 广义相空间描绘。若 $\|y_0(x) - y(x)\|_1 < \delta$, 则 $y(x)$ 在以 $y_0(x)$ 为轴心, 半径为 δ 的“管子”中。在 $x-y$ 平面(垂直于 y' 轴) $|y_0(x) - y(x)| < \delta$, 在 $x-y'$ 平面(垂直于 y 轴), $|y'_0(x) - y'(x)| < \delta$ 。

弱最大 若存在一个 $\varepsilon > 0$, 使得对于所有 $\|y_0(x) - y(x)\|_1 < \varepsilon$, 有 $I[y_0] - I[y] \geq 0$, 则 y_0 使泛函 I 达到弱极大化。

强最大 若存在一个 $\varepsilon > 0$, 使得对于所有 $\|y_0(x) - y(x)\|_0 < \varepsilon$, 有 $I[y_0] - I[y] \geq 0$, 有 $I[y_0] - I[y] \geq 0$, 则 y_0 使泛函 I 达到强极大化。

§3 泛函取极值的必要条件: 欧拉方程

为了避免繁复的脚码, 我们用 $y(x)$ 而不是用 $y_0(x)$ 表示最优函数, 而用 $z(x)$ 而不是用 $y(x)$ 表示在扰动下偏离最优函数的函数, 令 $\delta y(x)$ 为 $y(x)$ 的变分, 这与一般微分中用 Δx 表示 x 的微分意义类同, 于是

$$z(x) = y(x) + \delta y(x) \quad (3-17)$$

或

$$\delta y(x) = z(x) - y(x)$$

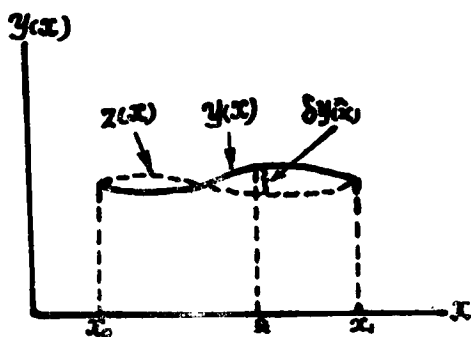


图 3-6

显然变分 $\delta y(x)$ 正是 $z(x)$ 和 $y(x)$ 在任一点 $x(x_0 \leq x \leq x_1)$ 的垂直距离, 见图3-6, 对于与 $y(x)$ 一阶趋近的 $z(x)$, $\delta y(x)$ 表示对于特定 x 值的 $y(x)$ 及其导数 $y'(x)$ 的一个无穷小变化。我们研究这个无穷小变化的后果类似于微分中研究 Δx 变化引起 Δy 变化。

从图3-6我们可以看出变分是 x 的函数, 我们可以视其为函数 $\phi(x)$ 乘上一个微小参数 ε , 即

$$\delta y(x) = \varepsilon \phi(x) \quad (3-18)$$

其中 $\phi(x_0) = 0$, $\phi(x_1) = 0$, 将上式代入 (3-17) 则

$$z(x) = y(x) + \varepsilon \phi(x) \quad (3-19)$$

当 $\varepsilon \rightarrow 0$ 时, $z(x) \rightarrow y(x)$, 即

$$\lim_{\varepsilon \rightarrow 0} z(x) = y(x).$$

根据(3-19), 则目标泛函

$$I = \int_{x_0}^{x_1} F[x, z(x), z'(x)] dx$$

可写为

$$I = \int_{x_0}^{x_1} F[x, y(x) + \varepsilon \phi(x), y'(x) + \varepsilon \phi'(x)] dx \quad (3-20)$$

显然, 对于任一固定的 $\phi(x)$, I 完全依赖于 ε 的大小, 因为 $y(x)$ 是最大化函数, 因此在 $\varepsilon = 0$ 时, I 取极大值,

$$\therefore \frac{dI}{d\varepsilon} = 0,$$

即

$$\left. \frac{dI}{d\varepsilon} \right|_{\varepsilon=0} = 0 \quad (3-21)$$

此处我们假定 $\frac{dI}{d\varepsilon}$ 存在。根据对积分求导数的法则，若有

$$I(\varepsilon) = \int_{x_0}^{x_1} F(\varepsilon) dx,$$

由于积分是求和的极限，根据和的导数等于导数之和法则，于是我们有

$$\frac{dI}{d\varepsilon} = \int_{x_0}^{x_1} \left(\frac{dF}{d\varepsilon} \right) dx,$$

令 (3-20) 中的

$$y(x) + \varepsilon \phi(x) = \Phi(\varepsilon),$$

且令

$$y'(x) + \varepsilon \phi'(x) = \Phi'(\varepsilon),$$

然后将

$$I = \int_{x_0}^{x_1} F[x, \Phi(\varepsilon), \Phi'(\varepsilon)] dx$$

对 ε 求导数，则我们有

$$\frac{dI}{d\varepsilon} = \int_{x_0}^{x_1} \left(\frac{\partial F}{\partial \Phi} \frac{d\Phi}{d\varepsilon} + \frac{\partial F}{\partial \Phi'} \frac{d\Phi'}{d\varepsilon} \right) dx \quad (3-22)$$

由于 $\frac{d\Phi}{d\varepsilon} = \phi(x)$ ，且 $\frac{d\Phi'}{d\varepsilon} = \phi'(x)$ ，于是上式变成

$$\frac{dI}{d\varepsilon} = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial \Phi} \phi(x) + \frac{\partial F}{\partial \Phi'} \phi'(x) \right] dx \quad (3-23)$$

利用分部积分公式，可得

$$\int_{x_0}^{x_1} \left[\frac{\partial F}{\partial \Phi'} \phi'(x) \right] dx = \left. \frac{\partial F}{\partial \Phi'} \phi(x) \right|_{x_0}^{x_1} - \int_{x_0}^{x_1} \left[\frac{d}{dx} \frac{\partial F}{\partial \Phi'} \phi(x) \right] dx, \quad (3-24)$$

曲线 $z(x)$ 不论取何种形式，都要经过两个端点 (x_0, y_0) , (x_1, y_1) ，而 $\phi(x_0) = 0$ ，且 $\phi(x_1) = 0$ ，所以上式等号右边第一项等于 0，于是 (3-23) 化简为

$$\frac{dI}{d\varepsilon} = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial \Phi} - \frac{d}{dx} \frac{\partial F}{\partial \Phi'} \right] \phi(x) dx \quad (3-25)$$

当 $\varepsilon = 0$ 时， $\Phi(\varepsilon) = y(x)$ ， $\Phi'(x) = y'(x)$ ，于是上式变成

$$\left. \frac{dI}{d\varepsilon} \right|_{\varepsilon=0} = \int_{x_0}^{x_1} \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) \phi(x) dx \quad (3-26)$$

利用分部积分，亦可将 (3-23) 化为上式，当 $\varepsilon = 0$ 时 $\frac{dI}{d\varepsilon} = 0$ ，于是

$$\int_{x_0}^{x_1} \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) \phi(x) dx = 0 \quad (3-27)$$

其中的函数 $\phi(x)$ 是任意连续函数, 且受 $y(x)$ 邻域的形状和范围影响 (在这个邻域中 $z(x)$ 存在)。由于 $\phi(x)$ 为任意的, 则 $\phi(x)$ 可以为正, 也可以为负, 还可以部分为正, 部分为负。在这种情况下, 要 (3-27) 成立, 则括号中的式子必须等于 0, 于是我们有:

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = 0, \quad x_0 \leq x \leq x_1 \quad (3-28)$$

或写成

$$F_y - \frac{d}{dx} F_{y'} = 0, \quad (3-28')$$

这就是著名的**欧拉方程**, 它是最大化函数 $y(x)$ 要满足的必要条件, 它亦是最小化函数的必要条件。因为从 (3-21) ~ (3-28') 的所有推理对于极大化问题和极小化问题都成立。

满足欧拉方程的函数 $y(x)$ 叫**极值曲线**。

(3-28) 给出的关于 $y(x)$ 的必要条件意味着 F_y 和 $F_{y'}$ 必须在整个 $x_0 \leq x \leq x_1$ 区域存在。一般这两个函数依赖于 x, y, y' , 实际上 $\frac{d}{dx} F_{y'}$ 要求 $F_{y'}$ 对 x 的导数在整个 $x_0 \leq x \leq x_1$ 区域存在, 这又以 $y'(x)$ 在此区域存在为条件, 因为 $F_{y'}$ 中的所有项对 x 的导数将包含 y' , 所以我们假定 F 的二阶偏导数连续。利用求导数的链法则, 可以将 (3-28') 中的第 2 项展开为

$$\frac{d}{dx} F_{y'} = \frac{\partial F_{y'}}{\partial y'} \frac{dy'}{dx} + \frac{\partial F_{y'}}{\partial y} \frac{dy}{dx} + \frac{\partial F_{y'}}{\partial x} \frac{dx}{dx}$$

即

$$\frac{d}{dx} F_{y'} = (F_{y'y'}) (y'') + (F_{y'y}) y' + F_{y'x} \quad (3-29)$$

于是欧拉方程也可写为

$$(F_{y'y'}) y'' + (F_{y'y}) y' + (F_{y'x} - F_y) = 0 \quad (3-30)$$

这一般是二阶非线性微分方程, 其边界条件为 $y(x_0) = y_0$ 和 $y(x_1) = y_1$ 。这个方程一般不太好求解, 困难之处在于它是非线性的, 且 $y(x)$ 之边界值在 x_0 和 x_1 给出, 而不是象一般二阶微分方程那样 $y(x)$ 和 $y'(x)$ 在 $x = x_0$ 处给出, 因此有必要求助于数值解法。

为了导出目标泛函取极值的二阶条件, 我们用另一种方法推导欧拉方程。

复习一下泰勒级数, 函数 $f(x)$ 可在 $x = a$ 展为

$$f(x) = f(a) + f'(a) \frac{x-a}{1!} + f''(a) \frac{(x-a)^2}{2!} + \cdots + f^{(n-1)}(a) \frac{(x-a)^{n-1}}{(n-1)!} + R. \quad (3-31)$$

对于多变量函数 $f(X)$, $X = (x_1, \dots, x_n)$, 可在 X^* 展成

$$f(X) = f(X^*) + df(X^*) + \frac{1}{2!}d^2f(X^*) + \dots + \frac{1}{(n-1)!}d^{(n-1)}f(X^*) + R_n, \quad (3-32)$$

引入梯度概念 $\nabla f = [\partial f / \partial x_1, \dots, \partial f / \partial x_n]'$, 然后记 $\partial f / \partial x_i = f_i$, 则梯度可简记为 $\nabla f = [f_1, \dots, f_n]'$. 又令

$$dX = [dx_1, \dots, dx_n]',$$

则

$$df(X^*) = [\nabla f(X^*)] \cdot [dX]. \quad (3-33)$$

令 $\partial^2 f / \partial x_i \partial x_j = f_{ij}$, 且用 H 表示二阶偏导数之 hessian 矩阵:

$$H = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{pmatrix} \quad (3-34)$$

则 $f(X)$ 在 X^* 的二阶全微分可写为

$$d^2f(X^*) = [dX]'H(X^*)[dX] \quad (3-35)$$

其中 $H(X^*)$ 表示在 X^* 取值的 hessian 矩阵。

若 X^* 为极大值点, 则

$$f(X) - f(X^*) = df(X^*) + \frac{1}{2!}d^2f(X^*) + \dots + \frac{1}{(n-1)!}d^{(n-1)}f(X^*) + R_n \leq 0 \quad (3-36)$$

极大化的一阶条件为

$$df(X^*) = [\nabla f(X^*)] \cdot [dX] = 0 \quad (3-37)$$

二阶条件为

$$d^2f(X^*) = [dX]'H(X^*)[dX] < 0 \quad (3-38)$$

由于(3-35)为二次式, 其符号由 $H(X^*)$ 之主子式决定。为了说明二次式符号与 $H(X^*)$ 主子式符号的关系, 我们首先考虑二变量二次式。

$$Q(x_1, x_2) = X'HX \quad (3-39)$$

其中 $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $H = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. $\because H$ 为对称矩阵 ($f_{12} = f_{21}$), 所以上式可展为

$$\begin{aligned} & a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2, \\ & = a_{11}x_1^2 + 2a_{12}x_1x_2 + \frac{a_{12}^2}{a_{11}}x_2^2 + a_{22}x_2^2 - \frac{a_{12}^2}{a_{11}}x_2^2 \\ & = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \left(\frac{a_{11}a_{22} - a_{12}^2}{a_{11}}\right)x_2^2 \end{aligned} \quad (3-40)$$

因此二次式的符号由一阶主行列式 $|H_1| = |a_{11}|$ 和二阶行列式 $|H_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 的符号决定。若 $|H_1| > 0$, $|H_2| > 0$, 则二次式之值肯定大于 0, 若 $|H_1| < 0$, $|H_2| > 0$, 则 $|H_2|/|H_1| < 0$, 所以二次式之值肯定为负。换言之, H 正定时, 二次式之值大于 0, H 负定时, 二次式之值小于 0。以上推导可推广到多元函数的情形。因此 (3—38) 成立的条件为 $H(X^*)$ 负定, 这也就是极大化之二阶条件, 极小化之二阶条件为 (3—35) 中的二次式 > 0 , 即要求 $H(X^*)$ 正定。

我们现在回到对目标泛函极大化条件的讨论, 这个条件为

$$I[z] - I[y] \leq 0,$$

令 $\Delta I = I[z] - I[y]$ 为 I 之增量即 I 之全变分, $y(x)$ 使 I 极大化则意味着

$$\Delta I \leq 0 \quad (3-41)$$

不难看出, 由于

$$I[z] = \int_{x_0}^{x_1} F(x, z, z') dx,$$

所以 ΔI 可写成:

$$\Delta I = \int_{x_0}^{x_1} [F(x, z, z') - F(x, y, y')] dx \quad (3-42)$$

回顾 (3—32), (3—33), (3—35), 则 $F(x, z, z')$ 可展为二阶近似:

$$\begin{aligned} F(x, z, z') &= F(x, y, y') + [\nabla F(y)]' [Z - Y] + \frac{1}{2} [Z - Y]' H(y) [Z - Y] \\ &= F(x, y, y') + F_y(z - y) + F_{y'}(z' - y') \\ &\quad + \frac{1}{2} [F_{yy}(z - y)^2 + 2F_{yy'}(z - y)(z' - y') + F_{y'y'}(z' - y')^2] \end{aligned} \quad (3-43)$$

其中

$$\nabla F(y) = \begin{pmatrix} \partial F / \partial y \\ \partial F / \partial y' \end{pmatrix} = \begin{pmatrix} F_y \\ F_{y'} \end{pmatrix}$$

$$[Z - Y] = \begin{pmatrix} z(x) - y(x) \\ z'(x) - y'(x) \end{pmatrix},$$

$$H(y) = \begin{pmatrix} F_{yy} & F_{yy'} \\ F_{y'y} & F_{y'y'} \end{pmatrix}$$

将 $y(x)$ 的变分 $z(x) - y(x) = \varepsilon \phi(x)$ 和 $z'(x) - y'(x) = \varepsilon \phi'(x)$ 代入 (3—43), 并将 $F(x, y, y')$ 移至左边, 则

$$F(x, z, z') - F(x, y, y') = (F_y) \varepsilon \phi(x) + (F_{y'}) \varepsilon \phi'(x) \\ + \frac{1}{2} \{ (F_{yy}) \varepsilon^2 [\phi(x)]^2 + 2(F_{yy'}) \varepsilon^2 \phi(x) \phi'(x) + (F_{y'y'}) \varepsilon^2 [\phi'(x)]^2 \} \quad (3-44)$$

将上式代入(3-42), 则有

$$\Delta I = \varepsilon \int_{x_0}^{x_1} [F_y \phi(x) + F_{y'} \phi'(x)] dx \\ + \varepsilon^2 / 2 \int_{x_0}^{x_1} \{ F_{yy} [\phi(x)]^2 + 2F_{yy'} \phi(x) \phi'(x) + F_{y'y'} [\phi'(x)]^2 \} dx \quad (3-45)$$

用类似微分的术语, 可将上式第一项的全变分式称为一阶变分, 用 δI 表示, 第二项称为二阶变分, 用 $\delta^2 I$ 表示, 在二阶近似下, 全变分 $\Delta I = \delta I + \delta^2 I$, 对于一阶近似 $\Delta I = \delta I$, 这时若要(3-41)成立, 则要求 $\delta I \leq 0$, 即

$$\delta I = \varepsilon \int_{x_0}^{x_1} [F_y \phi(x) + F_{y'} \phi'(x)] dx \leq 0 \quad (3-46)$$

由于 ε 在符号和大小上是任意的, 所以当且仅当下式成立时, (3-46)成立:

$$\int_{x_0}^{x_1} [F_y \phi(x) + F_{y'} \phi'(x)] dx = 0 \quad (3-47)$$

以下推导完全与(3-23)~(3-28)相同。利用分部积分公式

$$\int_{x_0}^{x_1} F_{y'} \phi'(x) dx = F_{y'} \phi(x) \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} F_{y'} \phi(x) dx,$$

由于 $\phi(x_0) = \phi(x_1) = 0$, 上式右边第一项消去, 于是(3-47)变成

$$\int_{x_0}^{x_1} \left[F_y - \frac{d}{dx} F_{y'} \right] \phi(x) dx = 0$$

由于 $\phi(x)$ 是任意的 (除端点外), 从上式可导出欧拉方程

$$F_y - \frac{d}{dx} F_{y'} = 0.$$

利用 $\Delta I = \delta I + \delta^2 I \leq 0$, 由于 $\delta I = 0$, 所以极大化的二阶条件为

$$\delta^2 I = \frac{1}{2} [X - X^*]' H(X^*) [X - X^*] < 0 \quad (3-48)$$

因此极大化之充分条件为 hessian 矩阵 $H(X^*)$ 负定, 而极小化的必要条件与极大化相同, 只是二阶条件变成 $H(X^*)$ 正定。

设变分问题中不只一个待选择的函数 $y(x)$, 而是有多个函数 $y_i(x)$, $i = 1, \dots, n$, 则目标泛函为

$$I = \int_{x_0}^{x_1} F(x, y_1, \dots, y_n, y'_1, \dots, y'_n) dx \quad (3-49)$$

其中 $y_i(x_0) = y_{i0}$, $y_i(x_1) = y_{i1}$, 令 $Y = [y_1, \dots, y_n]$, 令 $\dot{Y} = [y'_1, \dots, y'_n]$, 于是目标泛函

变成

$$\text{Max}_x: I[Y] = \int_{x_0}^{x_1} F(x, Y, \dot{Y}) dx \quad (3-50)$$

其中

$$Y(x_0) = Y_0, \quad Y(x_1) = Y_1,$$

最大化的必要条件是, 对于每个 $y_i(x)$, 都要满足欧拉方程:

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx} \frac{\partial F}{\partial \dot{y}_i} = 0, \quad i = 1, \dots, n, \quad (3-51)$$

这类似于一般多元微积分中的极大化为必要条件: 一阶偏导数等于 0。展开后共有 n 个二阶 (一般为非线性) 微分方程, 共有 $2n$ 个边界条件。

例 动态最优价格问题

某种商品由一厂商独家经营, 每周市场需求量与其价格及买者对未来价格的预期有关, 需求量 $z(t)$ 正比于价格变化速率 $\frac{dp}{dt}$, 负比于价格 $p(t)$, 用经济计量学搞出的需求函数之具体形式为

$$z(t) = -p + 100 \frac{dp}{dt} + 100, \quad (1)$$

生产费用函数为

$$c(z) = \frac{1}{2} z^2 + 2z + 40 \quad (2)$$

t 时刻的利润为

$$\pi(p, p', t) = z(t)p(t) - c(z) \quad (3)$$

厂商的目的是在 t_0 到 t_1 时域选择价格变化的轨线, 使利润最大化。设 $t_0 = 0$, $p(t_0) = p_0 = 70$, $t_1 = 52$, $p(t_1) = p_1 = 100$, 将 (1) 代入 (2), 再代入 (3), 则用 $p(t)$ 和 $p'(t)$ 表示的利润函数为

$$\pi(t) = -\frac{3}{2} p^2 + 202p - 5000(p')^2 - 10,200p' + 200pp' - 5240 \quad (4)$$

目标泛函为

$$\text{Max}_x: I = \int_{t_0}^{t_1} \pi(p, p') dt \quad (5)$$

其中 $t_0 = 0$, $p_0 = 70$, 且 $t_1 = 52$, $p_1 = 100$, 上式中的 π 相当于欧拉方程 (3-28') 中的 F 。

$$\pi_p = -3p + 202 + 200p'$$

$$\pi_{p'} = -10,000p' - 10,200 + 200p,$$

且

$$-\frac{d}{dt}\pi p' = -10,000p'' + 200p'$$

所以欧拉方程(3-28')在此例中为

$$10,000p'' - 3p = -202 \quad (6)$$

其齐次解为

$$p(t) = k_1 e^{0.0173t} + k_2 e^{-0.0173t} \quad (7)$$

(6) 之非齐次特解为一常数

$$p^*(t) = k_3 = 67.333 \quad (8)$$

于是(6)之通解为

$$p(t) = k_1 e^{0.0173t} + k_2 e^{-0.0173t} + 67.333 \quad (9)$$

利用边界条件 $t_0 = 0$, $p_0 = 70$, $t_1 = 52$, $p_1 = 100$, 则可得到方程组:

$$\begin{cases} k_1 + k_2 = 2.667 \\ 2.459k_1 + 0.407k_2 = 32.667 \end{cases}$$

解出 $k_1 = 15.391$ 和 $k_2 = -12.724$, 代入(9), 则通解为

$$p(t) = 15.391e^{0.0173t} - 12.724e^{-0.0173t} + 67.333 \quad (10)$$

从上式可求出 $p'(t)$, 然后将 $p(t)$, $p'(t)$ 代入(4) 然后将其从 $t_0 = 0$ 到 $t_1 = 52$ 积分, 求得最大利润为 170,000.

§4 端点条件

上节例中的动态最优价格问题可能端点并不是固定的(固定端点情形见图 3-7), 这又分为三种情况:

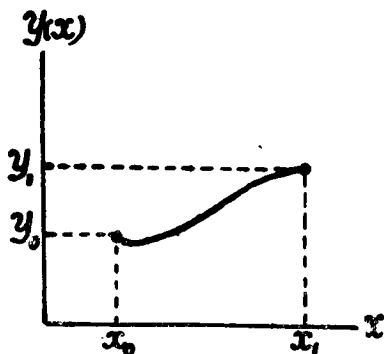


图 3-7

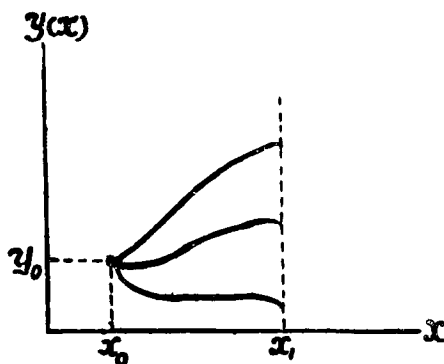


图 3-8

(1)可能终点条件不确定,但终端时间是固定的,例如 $t_1 = 52$ 时,价格是多少不知道,见图3—8,这就是所谓自由右边界问题;

(2)也可能最终价格 $p(t_1) = 100$,但什么时候达到这一价格却不知道,即 t_1 不定,见图3—9,这就是所谓自由终端时间问题;

(3)也许最终要达到的价格 $p(t_1)$ 和终端时间 t_1 都不知道,但是要求 p 与 t 满足一定的关系,例如要求价格上涨率在终端时间为 0.01,即 $p = (1 + 0.01)^t$ 成立,这就是固定终端曲面问题,见图3—10,在二维问题中,终端曲面实际上是曲线。

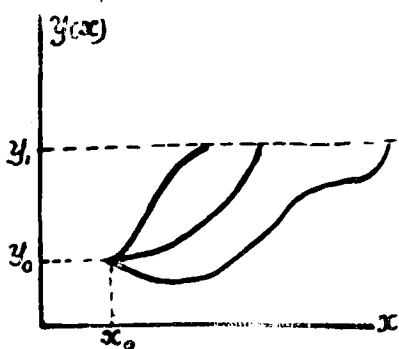


图 3—9

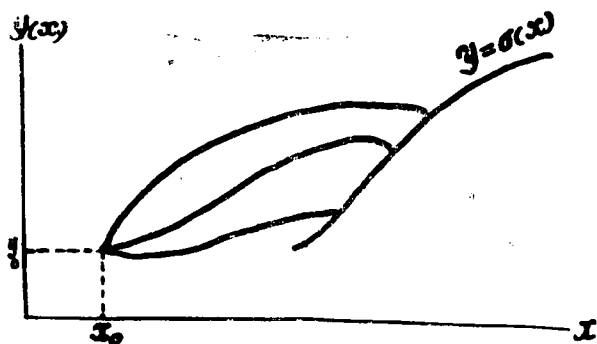


图 3—10

图3—8, 3—9不过是图3—10的特例。设图3—10中的终端曲面取一般形式 $y = \sigma(x)$, 则图3—8中的终端曲面为一垂直线 $x = x_1$, 而图3—9中的终端曲面则是函数 $y = y_1$ 。

一般地,多维空间中的目标泛函 I 依赖于若干个 $y_i(x)$, 终端曲面可写为 $S(x, y) = 0$, y 为矢量。于是图3—8中的终端曲面可写为 $S(x, y) = x - x_1$, 图3—9中的终端曲面可写为 $S(x, y) = y - y_1$, 图3—10中的终端曲面可写为 $S(x, y) = y - \sigma(x)$ 。

左端点也可按与上类同的方法处理,设初始曲面为 $y = \rho(x)$, 则可变边界问题可以用最一般的形式表为 $y(x)$ 的端点由某一初始曲面 $y - \rho(x) = 0$ 和某一终端曲面 $y - \sigma(x) = 0$ 确定。

在固定端点问题中,进行一阶变分式的推演时(3—47)为

$$\int_{x_0}^{x_1} [F_y \phi(x) + F_{y'} \phi'(x)] dx = 0 \quad (3-52)$$

通过分部积分可将上式等号左边第二项化简,即

$$\int_{x_0}^{x_1} F_{y'} \phi'(x) dx = F_{y'} \phi(x) \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} F_{y'} \phi(x) dx \quad (3-53)$$

由于 $z(x) = y(x) + \epsilon \phi(x)$, 且任一 $z(x)$ 必定通过特定的端点 $\phi(x_0) = \phi(x_1) = 0$, 所以(3—53)等号右边第一项消去, 于是(3—52)可写为

$$\int_{x_0}^{x_1} \left(F_y - \frac{d}{dx} F_{y'} \right) \phi(x) dx = 0 \quad (3-54)$$

于是推出欧拉方程

$$F_y - \frac{d}{dx} F_{y'} = 0$$

现在端点不是固定的, 于是(3-52)就不能变成(3-54)啦。所以我们要重新定义初始条件为 $y = \rho(x)$, 终端条件为 $y = \sigma(x)$ 的一阶全变分。图3-11说明了问题的几何意义。我们

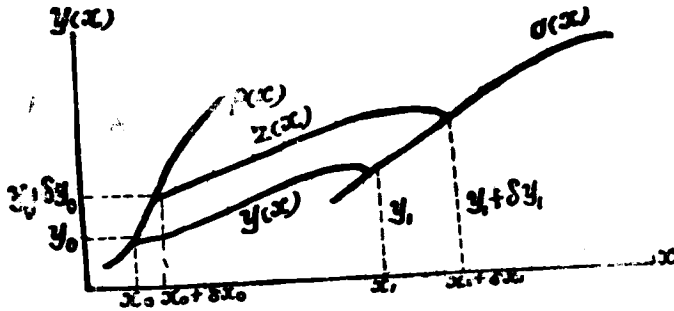


图 3-11

用 δx_0 表示 x 初始值的变化, 用 δy_0 表示邻近函数 $z(x)$ 初始值的变化, 这时 $y(x)$ 在 (x_0, y_0) 起始, 而 $z(x)$ 在函数 $y = \rho(x)$ 上的 $(x_0 + \delta x_0, y_0 + \delta y_0)$ 起始。同样, 当 δx_1 很小时, $y(x)$ 的右端点延伸至 $x_1 + \delta x_1$, 于是 $y(x)$ 和 $z(x)$ 都以 $x_1 + \delta x_1$ 为积分上限, 于是 (3-56) 可改写为

$$\begin{aligned} \Delta I &= \int_{x_0}^{x_1} [F(x, z, z') - F(x, y, y')] dx \\ &+ \int_{x_1}^{x_1 + \delta x_1} F(x, z, z') dx - \int_{x_0}^{x_0 + \delta x_0} F(x, z, z') dx \quad (3-57) \end{aligned}$$

上式等号右边第一项正是固定端点问题中的 ΔI , 根据上节(3-46), 此项可写为

$$\varepsilon \int_{x_0}^{x_1} [F_{y'} \phi(x) + F_{y''} \phi'(x)] dx,$$

所以, (3-57)之一阶近似可写为

$$\begin{aligned} \Delta I &= \varepsilon \int_{x_0}^{x_1} [F_{y'} \phi(x) + F_{y''} \phi'(x)] dx + \int_{x_1}^{x_1 + \delta x_1} F(x, z, z') dx \\ &- \int_{x_0}^{x_0 + \delta x_0} F(x, z, z') dx. \quad (3-58) \end{aligned}$$

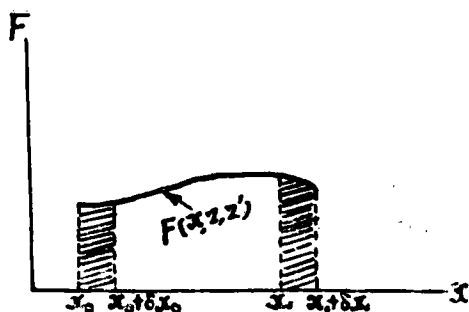


图 3-12

之差, 于是(3-58)可写成

$$\begin{aligned}\Delta I &\cong \varepsilon \int_{x_0}^{x_1} [F_y \phi(x) + F_{y'} \phi'(x)] dx + \int_{x_1}^{x_1 + \delta x_1} F(x, y, y') dx \\ &= \int_{x_0}^{x_0 + \delta x_0} F(x, y, y') dx\end{aligned}\quad (3-59)$$

此外, 当 δx_1 很小时, x_1 和 $x_1 + \delta x_1$ 很接近。 $\int_{x_1}^{x_1 + \delta x_1} F(x, y, y') dx$ 表示的面积可以用

$\delta(x_1) F(x, y, y')|_{x=x_0}$ 来近似, 这就是所谓积分学的一阶中值定理。图3-13说明了这个定理的几何意义。

阴影示出的矩形面积为曲线下面积的近似, δx_1 越小, 则这种近似越好, 同样, 左端的曲边梯形面积

$$\int_{x_0}^{x_0 + \delta x_0} F(x, y, y') dx \approx (\delta x_0) F(x, y, y')|_{x=x_0}$$

将这两个近似式代入(3-59), 然后利用分部积分将(3-59)中的第一项积分中的 $[F_y \phi'(x)]$ 变换一下, 于是最后得到

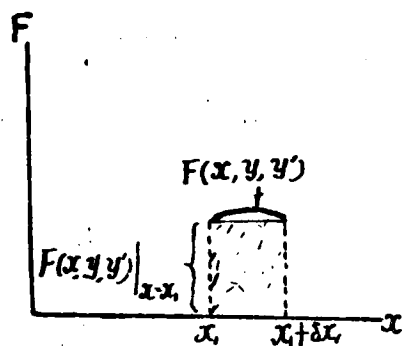


图 3-13

$$\begin{aligned}\Delta I &\cong \varepsilon \int_{x_1}^{x_1} \left(F_y - \frac{d}{dx} F_{y'} \right) \phi(x) dx + \varepsilon \phi(x_1) F_{y'}|_{x=x_1} \\ &\quad - \varepsilon \phi(x_0) F_{y'}|_{x=x_0} + (\delta x_1) F(x, y, y')|_{x=x_1} \\ &\quad - (\delta x_0) F(x, y, y')|_{x=x_0}\end{aligned}\quad (3-60)$$

我们将图3-11左端点放大成图3-14。对于 $z(x) = y(x) + \varepsilon \phi(x)$, 当 δx_0 很小时, δy_0 近似为两个垂直线段 $\varepsilon \phi(x_0)$ 和 $y'(x) \delta x_0$ 之和, 即

$$\varepsilon \phi(x_0) \cong \delta y_0 + y'(x_0) \delta x_0 \quad (3-61)$$

对于右端点也有类似的推理, 即

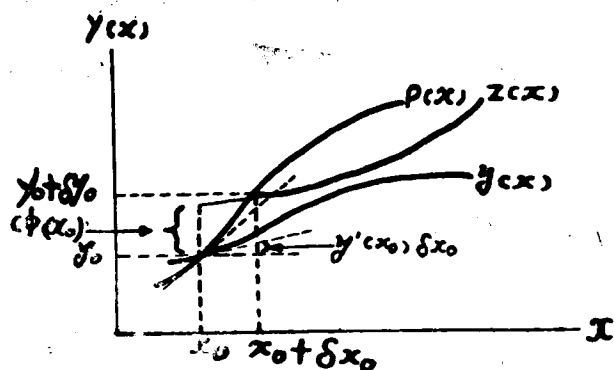


图 3-14

$$\varepsilon \phi(x_1) = \delta y_1 - y'(x_1) \delta x_1 \quad (3-62)$$

将以上近似式代入(3-60), 合并 $\delta x_0, \delta x_1, \delta y_0, \delta y_1$ 的同类项, 定义下式为 I 之一阶变分:

$$\begin{aligned} \delta I = \varepsilon \int_{x_0}^{x_1} \left(F_y - \frac{d}{dx} F_{y'} \right) \phi(x) dx + F_{y'}|_{x=x_0} (\delta y_0) \\ - (F - F_{y'} y')|_{x=x_0} (\delta x_0) = 0 \end{aligned} \quad (3-63)$$

在固定端点问题中, 由于没有 x_0, y_0, x_1, y_1 的变分, 即 $\delta x_0 = \delta y_0 = \delta x_1 = \delta y_1 = 0$, 因此 δI 只有一个积分项, (见上节(3-27)). 对于(3-63)的积分项来说, 由于 $y(x)$ 优于与其共端点 $(x_0, y_0), (x_1, y_1)$ 的邻域函数 $z(x)$, 因此欧拉方程仍应成立, 即

$$F_y - \frac{d}{dx} F_{y'} = 0 \quad (3-64)$$

所以积分项为 0, 于是当 δI 为 0 时, 最优化的必要条件变成

$$\begin{aligned} F_{y'}|_{x=x_1} (\delta y_1) + (F - F_{y'} y')|_{x=x_1} (\delta x_1) \\ - F_{y'}|_{x=x_0} (\delta y_0) - (F - F_{y'} y')|_{x=x_0} (\delta x_0) = 0 \end{aligned} \quad (3-65)$$

现在我们引入端点函数 $\rho(x)$ 和 $\sigma(x)$, 据此我们可将 y_0 的变分与 x_0 的变分联系起来, 也可以将 y_1 的变分与 x_1 的变分联系起来。这里再次采用近似方法, 当 $y(x)$ 之变分很小时, 这种近似更有效。参见图3-14, 考虑 $\rho(x)$ 在点 x_0 的切线, 不难看出

$$\delta y_0 \approx \rho'(x_0) \delta x_0 \quad (3-66)$$

这时假定 $0 < \rho'(x_0) < \infty$, 且 $0 < \sigma'(x_1) < \infty$, 即端点曲线之切线不是水平或垂直的, 于是 δy_0 与 δx_0 通过 ρ 之斜率与左端曲线挂上勾了, 而最大化函数必定经过这条曲线上某点, 对于右端点也有类似的分析, 即

$$\delta y_1 \approx \sigma'(x_1) \delta x_1 \quad (3-67)$$

将(3-66), (3-67)代入(3-65), 并合并 δx_1 和 δx_0 的同类项, 则得到

$$[F + (\sigma' - y')F_{y'}]|_{x=x_1}(\delta x_1) - [F + (\rho' - y')F_{y'}]|_{x=x_0}(\delta x_0) = 0 \quad (3-68)$$

由于 δx_1 和 δx_0 是任意的, 与其相乘的项都必须等于 0, (3-68)才能成立, 即

$$[F + (\sigma - y')F_{y'}]|_{x=x_1} = 0 \text{ 和 } [F + (\rho - y')F_{y'}]|_{x=x_0} = 0 \quad (3-69)$$

这两个条件一个是右端点条件, 一个是左端点条件, 相当于固定端点的简单端点条件 $y(x_0) = y_0$, $y(x_1) = y_1$, 它们就是周知的可变端点问题的**横截性条件**。极大化曲线 $y(x)$ 与函数 $\rho(x)$ 和 $\sigma(x)$ 相交, 即 $y(x)$ 是端点曲线的一个横截函数, 由于(3-69)与初始点的 y' 和 ρ' 有关, 且与终点的 y' 和 σ' 有关, 因此此式说明了这一横截函数是怎样满足与之相交的边界曲线的。

对于图3-8中的自由右端边界问题, 其中 x_1 是固定的, 而 y_1 不是固定的, $\delta x_1 = 0$ 且 $\delta x_0 = 0$, $\delta y_0 = 0$, 由于 $\delta y_1 \neq 0$, 则(3-65)化为

$$F_{y'}|_{x=x_1} = 0 \quad (3-70)$$

这时 $\sigma(x_1) = \infty$, 由于(3-67)中的近似不再适合, 所以(3-69)的结果不再成立。同样, 对于自由终端时间问题(见图3-9), $\delta y_1 = 0$, 且 $\delta x_0 = 0$, $\delta y_0 = 0$, 但 $\delta x_1 \neq 0$, 所以条件(3-65)变成

$$(F - F_{y'}y')|_{x=x_1} = 0 \quad (3-71)$$

这时 $\sigma'(x_1) = 0$, 因此(3-69)中的结果仍有效, 只是(3-67)中 $\delta y_1 = 0$, 加上(3-69), 令 $\sigma'(x_1) = 0$, 就得到上面的(3-71)。

例 动态最优价格问题。

问题与上节的例相同, 只是边界条件变了, 首先我们考虑终端时间为52周, 而此刻的价格不定之情形, 这就是自由右端边界问题, 然后考虑最终价格为 $p_1 = 20$, 但不知何时达到的自由终端时间问题。

(1) 自由右端边界

右端边界为 $t_1 = 52$, 横截条件为(3-70), 即

$$F_{p'}|_{t=t_1} = 0,$$

根据上节动态最优价格的条件

$$F_{p'} = -10,000p' - 10,200 + 200p \quad (1)$$

令其等于 0, 则

$$-10,000p' - 10,200 + 200p = 0 \quad (2)$$

欧拉方程仍成立, 即

$$p(t) = k_1 e^{0.0173t} + k_2 e^{-0.0173t} + 67.333 \quad (3)$$

将上式对 t 求导数, 则

$$p'(t) = 0.0173k_1 e^{0.0173t} - 0.0173k_2 e^{-0.0173t} \quad (4)$$

将 $t_1 = 52$ 代入上两式, 则

$$p(t_1) = 2.4586k_1 + 0.4067k_2 + 67.333 \quad (5)$$

$$p'(t_1) = 0.0425k_1 - 0.007k_2 \quad (6)$$

将上两式代入横截条件整理后得到

$$66.72k_1 + 151.34k_2 = -3266.667 \quad (7)$$

根据左端初始条件 $t_0 = 0$, $p(t_0) = 70$, 则

$$k_1 + k_2 = 2.667. \quad (8)$$

联立上两式解出

$$k_1 = 43.373, \quad k_2 = -407.06,$$

于是

$$p(t) = 43.373t^{0.0173t} - 40.706e^{-0.0173t} + 67.333 \quad (10)$$

当 $t_1 = 52$ 时, $p(t_1) = 157.4$.

将解出的 $p(t)$ 和 $t_0 = 0$, $t_1 = 52$ 代入目标泛函, 积分后得出 t_0 至 t_1 之最大利润约为 243,000, 比上节固定端点中的最大利润 170,000 大。

(2) 自由终端时间

若 $t_0 = 0$, $p(t_0) = 0$, $p(t_1) = 20$, 但 t_1 不确定, 于是横截条件 (3—21) 在这里变成 $(F - F_{p'} p')|_{t=t_1} = 0$. 原问题中

$$F = -\frac{3}{2}p^2 + 202p - 5000p'^2 - 10200p' + 200pp' - 5240, \quad (11)$$

$$F_{p'} = -10,000p' - 10200 + 200p, \quad (12)$$

所以横截条件变成

$$-\frac{3}{2}p^2 + 202p + 5000(p')^2 - 5240 = 0 \quad (13)$$

将 $p(t_1) = 20$ 代入上式, 则

$$p'(t_1) = \pm 0.6$$

欧拉方程仍成立, 即

$$p(t) = k_1 e^{0.0173t} + k_2 e^{-0.0173t} + 67.333 \quad (14)$$

对时间求导数, 则

$$p'(t) = 0.0173k_1e^{0.0173t} - 0.0173k_2e^{-0.0173t} \quad (15)$$

将 $p(t_1) = 20$, $p'(t_1) = \pm 0.6$ 代入上两式, 令 $t = t_1$, 则

$$k_1e^{0.0173t_1} + k_2e^{-0.0173t_1} = -47.333 \quad (16)$$

$$k_1e^{0.0173t_1} - k_2e^{-0.0173t_1} = \pm 34.682 \quad (17)$$

初始条件为 $t_0 = 0$, $p(t_0) = 10$, 代入(14), 则

$$k_1 + k_2 = -57.333 \quad (18)$$

(16), (17), (18)构成三个三元方程, 考虑到 ± 34.682 , 则(17)有两个三元方程, 在 $+34.682$ 情况下, 解出三元方程组中的

$$k_1 = -6.326e^{-0.0173t_1} \quad (19)$$

$$k_2 = -41.008e^{0.0173t_1} \quad (20)$$

代入(18), 则

$$-6.326e^{-0.0173t_1} - 41.008e^{0.0173t_1} = -57.333 \quad (21)$$

用牛顿近似法或用计算器编个小程序进行试凑计算, 可解出

$$t_1 \approx 14.15, \quad (22)$$

再代入(19), (20), 则

$$k_1 = -4.95, \quad k_2 = -52.38 \quad (23)$$

且

$$p(t) = -4.95e^{0.0173t} - 52.38e^{-0.0173t} + 67.333 \quad (24)$$

在 -34.682 的情形, 三元方程组之解为

$$t_1 = 122.2, \quad k_1, k_2 \text{ 值与上相同} \quad (25)$$

由于函数 $p(t)$ 在 $t^* = 68$ 有唯一的最大值35.13, 因此必有一个小于68的 t 和一个大于68的 t , 使 $p(t) = 20$. 因此有两组最优解, 不过在 $t_0 = 0$, $t_1 = 14.15$ 的情形, 达到 $p(t_1) = 20$ 的时间较短。

§5. 约束条件

(1) 整体等式约束

含这类约束的经典问题是由欧拉解决的所谓等周问题, 即在给定长度为 K 的所有可能的封闭曲线中选择使其围成的区域面积最大者。又例如, 联结直线上两固定端点的曲线有定长 K , 选择所有可能的曲线中使其与直线围成的面积最大者。这类问题目标泛函为

$$\max; I(y) = \int_{x_0}^{x_1} F(x, y, y') dx \quad (3-72)$$

其中 $y(x_0) = y_0$, $y(x_1) = y_1$.

约束条件为

$$\int_{x_0}^{x_1} G(x, y, y') dx = K \quad (3-72')$$

在等周问题中, 约束是要求总弧长为常数, 此处 $G(x, y, y') = \sqrt{1+y'^2}$. (3-72') 叫整体约束, 因为它确定了 $y(x)$ 的整个特性. 作拉氏函数:

$$L(x, y, y', \lambda) = F(x, y, y') - \lambda G(x, y, y') \quad (3-73)$$

将欧拉方程中的 F 换成 L , 则得到必要条件:

$$L_y - \frac{d}{dx} L_{y'} = 0 \quad (3-74)$$

即

$$\frac{\partial}{\partial y} (F - \lambda G) - \frac{d}{dx} \left[\frac{\partial}{\partial y'} (F - \lambda G) \right] = 0$$

$$F_y - \lambda G_y - \frac{d}{dx} (F_{y'} - \lambda G_{y'}) = 0$$

$$\left(F_y - \frac{d}{dx} F_{y'} \right) - \lambda \left(G_y - \frac{d}{dx} G_{y'} \right) = 0 \quad (3-75)$$

上式中第一个括号中的式子正是无约束欧拉方程形式, 第二个括号中的式子形式与前者并无差别, 只是将 F 换成了 G .

(2) 局部等式约束

设有两个待选择函数 $y_1(x)$, $y_2(x)$, 可写为矢量形式 $Y(x) = [y_1(x), y_2(x)]'$, 要求使目标泛函极大化, 即

$$\max_y I[y] = \int_{x_0}^{x_1} F(x_1, y_1, y_2, y'_1, y'_2) dx \quad (3-76)$$

其中端点固定, 且服从等式约束:

$$g(x, y_1, y_2) = C,$$

或写成

$$h(x, y_1, y_2) = g(x, y_1, y_2) - C = 0 \quad (3-77)$$

作拉氏函数

$$L(x, Y, \dot{Y}, \lambda) = F(x, Y, \dot{Y}) - [\lambda(x)] h(x, Y, \dot{Y}), \quad (3-78)$$

此式由于将 $\dot{Y}$ 纳入约束, 因此可适于更一般的情况, 由于约束是随 x 变化的, 所以 λ 也是 x 的函数, 将欧拉方程中的 F 换成 L , 则最优化的必要条件为

$$L_{y_i} - \frac{d}{dx} L_{y_i'} = 0 \quad (3-79)$$

且满足 $h(x, Y, \dot{Y}) = 0$ 和固定端点条件

$$Y(x_0) = Y_0, \quad Y(x_1) = Y_1$$

若 $Y = [y_1, \dots, y_n]'$, $n > 2$, 且有 k 个约束条件 $h_1(x, Y) = 0, \dots, h_k(x, Y) = 0, k < n$, 则有 k 个拉氏乘子 $\lambda_1(x), \dots, \lambda_k(x)$, 拉氏函数为

$$L = F - \sum_{i=1}^k \lambda_i(x) h_i \quad (3-80)$$

最优化要求偏导数矩阵 $[\partial g_i / \partial y_j]$ 满秩且各列线性独立。局部约束问题要求待选择函数多于两个, 即约束条件数少于变量数。若有一个约束, $g(x, y) = C$, 一个 $y(x)$ 函数, 则对最优函数没有选择余地, 只能选择约束方程作为唯一可能的解。一般而言, 若约束条件数 k 等于函数个数 n , 也没有什么最优化的选择余地。例如若有两个约束曲面, 两个待选择函数, 则两个约束曲面一般相交于一条曲线, 若这条曲线经过两固定端点, 则只有唯一可行解, 若不过任一端点, 则问题无解。

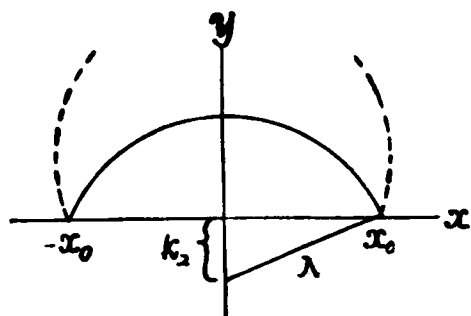


图 3-15

图3—15中联接 $-x_0$ 和 x_0 两点的曲线有定长 K , 求与 x 轴围成的面积最大之曲线。目标泛函为

$$\max, I[y] = \int_{-x_0}^{x_0} y(x) dx \quad (1)$$

其中 $y(-x_0) = 0, y(x_0) = 0$

约束为

$$\int_{-x_0}^{x_0} \sqrt{1+y'^2} dx = K \quad (2)$$

利用(3-75), 且由于 $F = y(x), F_y = 1, F_{y'} = 0, G = \sqrt{1+y'^2}, G_y = 0, G_{y'} = y' / \sqrt{1+y'^2}$, 于是我们有

$$1 + \lambda \left[\frac{d}{dx} \frac{y'}{\sqrt{1+y'^2}} \right] = 0 \quad (3)$$

两边积分得

$$x + \lambda \left[\frac{y'}{\sqrt{1+y'^2}} \right] = k_1 \quad (4)$$

令 $y' = \tan \alpha$, $\because 1 + \tan^2 \alpha = \sec^2 \alpha$, 且 $\sec \alpha = \frac{1}{\cos \alpha}, \frac{\sin \alpha}{\cos \alpha} = \tan \alpha$, $\therefore (4)$ 可写为

$$(x - k_1) = -\lambda \sin \alpha \quad \text{或} \quad x = k_1 - \lambda \sin \alpha$$

$\because y' = \tan \alpha, \therefore dy = \tan \alpha dx$, 又 $\because x = k_1 - \lambda \sin \alpha, \therefore dx = -\lambda \cos \alpha d\alpha, \therefore dy =$

$-\lambda \operatorname{tg} \alpha \cos \alpha d\alpha = -\lambda \sin \alpha d\alpha$, 对此式两边积分, 则

$$y = \lambda \cos \alpha + k_2 \quad \text{或} \quad y - k_2 = \lambda \cos \alpha,$$

即

$$x - k_1 = -\lambda \sin \alpha$$

$$y - k_2 = \lambda \cos \alpha \quad (5)$$

$\therefore k_1 = 0$, 即圆点在 y 轴上,

$$\therefore x_0^2 + k_2^2 = \lambda^2 \quad \text{或} \quad k_2 = \pm \sqrt{\lambda^2 - x_0^2} \quad (6)$$

若 $\lambda = x_0$, 则 $k_2 = 0$, 而 $\pi x_0 = \pi \lambda$ 为曲线周长, 即为圆周长之一半。若 $\lambda > x_0$, 且假定 $K > 2x_0$, 若 $k_2 = \sqrt{\lambda^2 - x_0^2}$, x 轴上部的圆弧长大于整个圆周长之一半, 若 $k_2 = -\sqrt{\lambda^2 - x_0^2}$, 则 x 轴上部的圆弧长小于整个圆周长之一半。

将 $y' = \operatorname{tg} \alpha$, $dx = -\lambda \cos \alpha d\alpha$ 代入 (2), 且当 $x = x_0$ 时, 我们有 $x_0 - k_1 = -\lambda \sin \alpha$, 将 $k_1 = 0$ 代入, 则 $\alpha = \sin^{-1}\left(-\frac{x_0}{\lambda}\right)$, 当 $x = -x_0$ 时, 则 $-x_0 - k_1 = -\lambda \sin \alpha$, 于是 $\alpha = \sin^{-1}\left(\frac{x_0}{\lambda}\right)$, 代入 (2) 中的积分上下限, 约束条件变成

$$\int_{\sin^{-1}\left(\frac{x_0}{\lambda}\right)}^{\sin^{-1}\left(-\frac{x_0}{\lambda}\right)} -\lambda \sqrt{1 + \operatorname{tg}^2 \alpha} \cos \alpha d\alpha = K$$

即

$$\begin{aligned} & \sin^{-1}\left(-\frac{x_0}{\lambda}\right) \cdot \\ & -\lambda \int_{\sin^{-1}\left(\frac{x_0}{\lambda}\right)} d\alpha = K \\ & -\lambda \left[\sin^{-1}\left(-\frac{x_0}{\lambda}\right) - \sin^{-1}\left(\frac{x_0}{\lambda}\right) \right] = K \end{aligned} \quad (7)$$

从上面的三角函数方程利用牛顿近似法或计算机程序试凑计算可解出 λ , 再代入 (6) 则可解出 k_2 , 据此可确定圆弧的圆心和半径, 因而确定圆弧的形状、方位。

§6 最大值原理

上两节讨论的古典变分法不能处理 F_v 和 F_{vv} 为常数的问题及目标函数。约束条件为线性式的问题, 拉氏乘数法则也不能处理目标函数、约束条件为线性式的问题。1950年苏联数学家 Понтрягин 提出了最大值原理, 这种最优控制理论更适于解含约束条件特别是含不等式约束的问题, 并且对函数的连续性和平滑性没有严格约束。

最优控制问题可描述如下:

令 $u(x) = y'(x)$ 为控制变量, 则有控制问题:

$$\text{Max/Min,} \quad I[y, u] = \int_{x_0}^{x_1} F(x, y, u) dx \quad (3-81)$$

其中 $y(x_0) = y_0$, $y(x_1) = y_1$, 且 $y'(x) = u(x)$, x, y 为状态变量。这相当于变量为 y 和 u 的变分法问题, 其约束条件为

$$h(u, y') = y' - u = 0 \quad (3-82)$$

我们在(3-51)和(3-80)中写出了多个函数的泛函极值条件及多个等式约束下的泛函极值条件, 利用这些结果, 拉氏函数为

$$\begin{aligned} L(x, y, u, y') &= F(x, y, u) - \lambda(x)h(u, y') \\ &= F(x, y, u) - \lambda(x)[y' - g(u)] \end{aligned} \quad (3-83)$$

其中 $g(u)$ 是约束条件 $y' = g(u)$ 的更一般形式, 在(3-82)中 $g(u) = u$, 回顾(3-51), 将欧拉方程用于(3-83), 则

$$L_y - \frac{d}{dx}L_{y'} = 0 \quad (3-84a)$$

$$L_u - \frac{d}{dx}L_u = 0 \quad (3-84b)$$

$$L_\lambda - \frac{d}{dx}L_\lambda = 0 \quad (3-84c)$$

对于(3-81)和(3-82)的问题, (3-84)具体可写为

$$L_{y'} = -\lambda(x)$$

$$L_u = 0$$

$$L_\lambda = 0$$

$$L_y = -\lambda'(x), \quad \text{即} \quad -\frac{\partial L}{\partial y} = \lambda'(x) \quad (3-84a)$$

$$L_u = 0, \quad \text{即} \quad \frac{\partial L}{\partial u} = 0 \quad (3-84b)$$

$$L_\lambda = 0, \quad \text{即} \quad \frac{\partial L}{\partial \lambda} = 0 \quad (3-84c)$$

为了导出最大值原理, 我们引入汉密尔顿函数 H , 令

$$H = F + \lambda g \quad (3-85)$$

由于原问题中拉氏函数为

$$L = F - \lambda y' + \lambda g$$

所以

$$L = H - \lambda y'$$

H 仍依赖于 x, y, u 和 λ , 但 H 中不再含 y' , 其次

$$\frac{\partial L}{\partial y} = \frac{\partial H}{\partial y}, \quad \frac{\partial L}{\partial u} = \frac{\partial H}{\partial u}, \quad \frac{\partial L}{\partial \lambda} = \frac{\partial H}{\partial \lambda} - y'.$$

于是必要条件(3—83)变成

$$\left\{ \begin{array}{l} -\frac{\partial H}{\partial y} = \lambda' \\ \frac{\partial H}{\partial u} = 0 \\ \frac{\partial H}{\partial \lambda} = y' \end{array} \right. \quad \begin{array}{l} (3-86a) \\ (3-86b) \\ (3-86c) \end{array}$$

以上必要条件为一对一阶微分方程, (3—81)中的端点条件给定了 $y(x)$ 的初始值和终端值, 但并未给定 x_0 和 x_1 处的 λ 值。在控制理论中, 与拉氏乘数类似的变量 λ 一般叫做协状态变量或伴随变量。除符号外, (3—86a)与(3—86c)是对称的, 条件(3—86b)为无约束极值条件。即要求 H 取极值。

变分法与最大值原理主要有如下差别:

(1) 最大值原理只要求控制变量 $u(x)$ 分段连续, 它可以有跃变, 而变分法却要求所有变量连续。

(2) 最大值原理只要求 $F(x, y, u)$ 的一阶偏导数连续, 而变分法要求 F 的一阶和二阶偏导数连续;

(3) 二者都要求状态变量分段光滑。

(4) 最大值原理对于约束函数 $g(x, y, u)$ 在某些情况下要求有连续的一阶和二阶偏导数。

例 动态最优价格问题

前几节例举的动态最优价格问题用控制理论可作如下描述。

$$\text{Max: } \int_{t_0}^{t_1} \left(-\frac{3}{2} p^2 + 202p - 5000u^2 - 10, 200u + 200pu - 5240 \right) dt$$

$$\text{其中 } t_0 = 0, p(t_0) = 70, t_1 = 52, p(t_1) = 100, p' = u \quad (1)$$

拉氏函数为

$$L = -\frac{3}{2} p^2 + 202p - 5000u^2 - 10, 200u + 200pu - 5240 - \lambda(p' - u),$$

汉密尔顿函数为

$$H = -\frac{3}{2} p^2 + 202p - 5,000u^2 - 10, 200u + 200pu + 5240 + \lambda u \quad (2)$$

最大值原理的必要条件(3—86)为

$$3p - 202 - 200u = u' \quad (3a)$$

$$-10,000u - 10,200 + 200p + \lambda = 0 \quad (3b)$$

$$u = p' \quad (3c)$$

从(3b)可解出 λ, p, u 之表达式, 将 λ 对 t 求导数, 则得到 λ' 的表达式, 令其等于(3a)左边, 然后将 $u = p', u' = p''$ 代入, 则得到二阶微分方程

$$10,000p'' - 3p = -202 \quad (4)$$

这就是§3例题中的(6)式。由于初始和终端条件与(6)式相同, 所以可从(4)得到运用变分法同样的结果

$$p(t) = 15.391e^{0.0173t} - 12.724e^{-0.0173t} + 67.333 \quad (5)$$

对于这类问题, 我们可以有下列更一般的程序, 从(3b)可得

$$u = 0.02p + 0.0001\lambda - 1.02 \quad (6)$$

代入(3a)和(3c)得到下列一对一阶微分方程

$$\begin{cases} p' = 0.02p + 0.0001\lambda - 1.02 \\ \lambda' = -p - 0.02\lambda + 2 \end{cases} \quad (7)$$

根据第一章§5, §6解一阶线性微分方程组的程序, 我们首先要求得一个非奇异矩阵, 以便将(7)式中的系数矩阵 $A = \begin{bmatrix} 0.02 & 0.0001 \\ -1 & -0.02 \end{bmatrix}$ 化为规范型。为了求这个矩阵, 我们首先要求特征矢量。根据前文我们已知特征值为0.0173和-0.0173, 则根据特征矢量定义

$$Av = \lambda_i v_i, \quad (8)$$

我们有

$$\begin{bmatrix} 0.02 & 0.0001 \\ -1 & -0.02 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = 0.0173 \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} \quad (9)$$

和

$$\begin{bmatrix} 0.02 & 0.0001 \\ -1 & -0.02 \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = -0.0173 \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} \quad (10)$$

从中解出一对特征矢量为

$$v_1 = \begin{bmatrix} 1 \\ -27 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -373 \end{bmatrix} \quad (11)$$

奇异变换矩阵

$$p = [v_1, v_2] \quad (12)$$

于是方程组的解为

$$\begin{pmatrix} p \\ \lambda \end{pmatrix} = [v_1, v_2] \begin{pmatrix} e^{0.0173t} & 0 \\ 0 & e^{-0.0173t} \end{pmatrix} [v_1, v_2]^{-1} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} + \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \quad (13)$$

其中由于 $[v_1, v_2]^{-1} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -27 & -373 \end{pmatrix}^{-1} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ 为常数列矢量 $[c_1, c_2]'$, 所以解可写为

$$\begin{pmatrix} p \\ \lambda \end{pmatrix} = c_1 v_1 e^{0.0173t} + c_2 v_2 e^{-0.0173t} + \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \quad (14)$$

而 $\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ 为非齐次式之一对解, 于是有

$$\begin{cases} p = c_1 e^{0.0173t} + c_2 e^{-0.0173t} + h_1 \\ \lambda = -27c_1 e^{0.0173t} - 373c_2 e^{-0.0173t} + h_2 \end{cases} \quad (15)$$

$$\lambda = -27c_1 e^{0.0173t} - 373c_2 e^{-0.0173t} + h_2 \quad (16)$$

令 p 和 λ 为常数, 则非齐次方程组(7)变成

$$\begin{cases} 0.02p + 0.0001\lambda = 1.02 \\ -p - 0.02\lambda = -2 \end{cases} \quad (17)$$

解出 $p = 67.333$, $\lambda = -3266.667$, 这就是 h_1, h_2 , 于是

$$\begin{cases} p = c_1 e^{0.0173t} + c_2 e^{-0.0173t} + 67.333 \\ \lambda = -27c_1 e^{0.0173t} - 373c_2 e^{-0.0173t} - 3266.667 \end{cases} \quad (18)$$

利用端点条件 $p(0) = 70$, $p(52) = 100$, 可得到与变分法解相同的结果:

$$\begin{cases} p(t) = 15.391e^{0.0173t} - 12.724e^{-0.0173t} + 67.333 \\ \lambda(t) = -415.557e^{0.0173t} + 4746.052e^{-0.0173t} - 3266.667 \end{cases} \quad (19)$$

协状态变量 $\lambda(t)$ 反映了价格变化速率对总利润的影响, 它是广义的动态“影子价格”, 即价格的“影子价格”。

若有 n 个状态变量, m 个控制变量和 n 个等式约束, 则最大值原理在引入下列矢量后可写为更一般的形式,

$$\begin{aligned} Y &= [y_1(x), \dots, y_n(x)]', \quad \dot{Y} = [y_1'(x), \dots, y_n'(x)]', \\ U &= [u_1(x), \dots, u_m(x)]', \quad G(x, Y, U) = [g_1(x, Y, U), \dots, g_n(x, Y, U)]', \\ A(x) &= [\lambda_1(x), \dots, \lambda_n(x)]', \quad \dot{A} = [\lambda_1'(x), \dots, \lambda_n'(x)]', \end{aligned}$$

于是:

$$\text{Max}_x: \quad I[Y, U] = \int_{x_0}^{x_1} F(x, Y, U) dx \quad (3-87)$$

其中 $Y(x_0) = Y_0$, $Y(x_1) = Y_1$, $\dot{Y} = G(x, Y, U)$

拉氏函数为

$$L(x, Y, U, \dot{Y}) = F(x, Y, U) - [\Lambda(x)]' \cdot [\dot{Y} - G(x, Y, U)], \quad (3-88)$$

汉氏函数为

$$H(x, Y, U, \Lambda) = F(x, Y, U) + [\Lambda(x)]' \cdot [G(x, Y, U)], \quad (3-89)$$

令

$$H_Y = [\partial H / \partial y_1, \dots, \partial H / \partial y_n]'$$

$$H_\Lambda = [\partial H / \partial \lambda_1, \dots, \partial H / \partial \lambda_n]'$$

$$H_U = [\partial H / \partial u_1, \dots, \partial H / \partial u_m]'$$

于是必要条件(3-86)可写为

$$\begin{cases} -H_Y = \dot{\Lambda} & (3-90a) \\ H_U = 0 & (3-90b) \\ H_\Lambda = \dot{Y} & (3-90c) \end{cases}$$

上列必要条件共有 $2n$ 个微分方程, 其中(3-90b)中的 m 个方程应使 H 极大化。

§7 最大值原理之端点条件

回顾(3-65)表示的端点条件

$$\begin{aligned} & F_{y'}|_{x=x_1}(\delta y_1) + (F - F_{y'}y')|_{x=x_1}(\delta x_1) \\ & - F_{y'}|_{x=x_0}(\delta y_0) - (F - F_{y'}y')|_{x=x_0}(\delta x_0) = 0 \end{aligned} \quad (3-65)$$

若左端点固定, 则 $\delta y_0 = \delta x_0 = 0$, 于是上式变成

$$F_{y'}|_{x=x_1}(\delta y_1) + (F - F_{y'}y')|_{x=x_1}(\delta x_1) = 0 \quad (3-91)$$

若引入 I 的梯度 $\nabla I = \begin{pmatrix} F - F_{y'}y' \\ F_{y'} \end{pmatrix}$, 则上列一阶变分条件式可写为

$$[\nabla I] \cdot \begin{pmatrix} \delta x_1 \\ \delta y_1 \end{pmatrix} = 0 \quad (3-92)$$

即端点条件可更一般地表为矢量正交条件。

最大值原理中有等式约束 $y' = g(x, y, u)$, 而上列端点条件问题中都未考虑约束, 所以我们可以将约束搞进目标泛函, 以便取得与上述问题类似的结构形式。设目标泛函为

$$I[y, u] = \int_{x_0}^{x_1} L(x, y, u, y', \lambda) dx \quad (3-93)$$

其中
$$L(x, y, u, y', \lambda) = F(x, y, u) - \lambda(x)[y' - g(x, y, u)], \quad (3-94)$$

于是梯度

$$\nabla I = \begin{pmatrix} (L - L_{y'}y') \\ L_{y'} \end{pmatrix} \quad (3-95)$$

控制理论公式中，汉氏函数为

$$H = F + \lambda g$$

$$L = H - \lambda y'$$

则
$$L_{y'} = -\lambda(x)$$

$$L - L_{y'}y' = F - \lambda y' + \lambda g + \lambda y' = F + \lambda g = H,$$

所以
$$\nabla I = \begin{pmatrix} H \\ -\lambda \end{pmatrix} \quad (3-96)$$

(1) 固定端点

若 $y(x_0) = y_0$, $y(x_1) = y_1$, 则 $\delta x_1 = 0$, $\delta y_1 = 0$, 端点条件要求在 x_1 满足:

$$[\nabla I] \cdot \begin{pmatrix} \delta x_1 \\ \delta y_1 \end{pmatrix} = 0, \text{ 即 } \begin{pmatrix} H(x_1) \\ -\lambda(x_1) \end{pmatrix} \cdot \begin{pmatrix} \delta x_1 \\ \delta y_1 \end{pmatrix} = 0. \quad (3-97)$$

所以在右端点 x_1 对汉氏函数 H 或协状态变量 λ 的值完全没有约。

(2) 自由右边界

这种情况下 $\delta x_1 = 0$, 但 $y(x)$ 未给定, 只要求其终结于 $x = x_1$, 所以 $\delta y_1 \neq 0$, 因此横截条件(3-97)对汉氏函数 $H(x_1)$ 没有约束, 但要求

$$-\lambda(x_1) = 0, \text{ 即 } \lambda(x_1) = 0 \quad (3-98)$$

于是必要条件中的一对微分方程

$$-\frac{\partial H}{\partial y} = \lambda', \quad \frac{\partial H}{\partial \lambda} = y'$$

之端点条件为 $y(x_0) = y_0$, $\lambda(x_1) = 0$ 。

§4中动态最优价格的自由右边界例子中

$$t_1 = 52, \quad p(t_1) \text{ 不定}.$$

这一问题的控制论横截条件为 $\lambda(t_1) = 0$ 。即端点条件为

$$t_0 = 0, \quad p(t_0) = 70; \quad t_1 = 52, \quad \lambda(t_1) = 0.$$

利用上节例中的(18)式和左端点条件可得到

$$k_1 + k_2 = 2.667$$

利用上节例中的(18)式和右端点条件可得到

$$-66.383k_1 - 151.711k_2 = 3266.667,$$

联立上两式可解得

$$k_1 = 43.026, k_2 = -40.358$$

用变分法解出的

$$k_1 = 43.026, k_2 = -40.706$$

误差是计算中的四舍五入造成的。

(3) 自由终端时间

在这种情况下, $\delta y_1 = 0$, $\delta x_1 \neq 0$, 端点条件(3-97)说明, 在 $x = x_1$, 对 λ 没有约束, 但汉氏函数必须为 0, 即

$$H(x_1) = 0 \quad (3-99)$$

§4中动态最优价格的自由终端时间例子中

$$t_0 = 0, p(0) = 10, p(t_1) = 20, t_1 \text{ 待定}$$

在控制论公式中, 横截性条件要求下式等于 0

$$H = -\frac{3}{2}p^2 + 202p - 5000u^2 - 10,200u + 200pu - 5240 + \lambda u,$$

根据上节例中的(18)

$$\lambda = 10,000u + 10,200 - 200p,$$

$$\therefore \lambda u = 10,000u^2 + 10,200u - 200pu,$$

代入 H , 并令 $H(t_1) = 0$, 则

$$-\frac{3}{2}p_1^2 + 202p_1 + 5000u_1^2 - 5240 = 0,$$

由于 $u = p'$, $p_1 = 20$, 于是

$$p'(t_1) = \pm 0.6$$

将 $p(t_1) = 20$, 代入(18), 则

$$20 = c_1 e^{0.0173t_1} + c_2 e^{-0.0173t_1} + 67.333$$

$$\text{即 } c_1 e^{0.0173t_1} + c_2 e^{-0.0173t_1} = -47.333$$

将(18)中的 p 对 t 求导数, 用 0.0173 除且将 $p'(t_1) = \pm 0.6$ 代入, 则得到

$$p'(t_1) = 0.0173c_1 e^{0.0173t_1} - 0.0173c_2 e^{-0.0173t_1}$$

$$\pm 0.6 = 0.0173c_1 e^{0.0173t_1} - 0.0173c_2 e^{-0.0173t_1}$$

$$c_1 e^{0.0173t_1} - c_2 e^{-0.0173t_1} = \pm 34.682$$

将左端点条件 $t_0 = 0$, $p(t_0) = 10$ 代入上节例中的(18), 则有

$$c_1 + c_2 = -57.333$$

联立含 c_1, c_2, t_1 三个未知量的三个方程可解得

$$c_1 = -4.95, c_2 = -52.38, t_1 \approx 14.15, 122.2.$$

(4) 固定终端曲面

右端点条件仍为(3—92), 即

$$\begin{pmatrix} (F - F_{y'} y') \\ F_{y'} \end{pmatrix} \cdot \begin{pmatrix} \delta x_1 \\ \delta y_1 \end{pmatrix} = 0.$$

不过最优曲线必须终结于 $y = \sigma(x)$, 所以在右端点 y 的变分 δy 与 x 的变分 δx 不是独立的, 对于自变量在 x_1 的微小变分, 我们有

$$\delta y_1 = \sigma'(x_1) \delta x_1, \text{ 即 } \delta y_1 - \sigma'(x_1) \delta x_1 = 0,$$

也就是说 y_1 的变分由 $y = \sigma(x)$ 在 $x = x_1$ 的切线近似, 这意味着 $\begin{pmatrix} \delta x_1 \\ \delta y_1 \end{pmatrix}$ 和 $\begin{pmatrix} 1 \\ \sigma'(x_1) \end{pmatrix}$ 为平行

矢量, 所以 $\begin{pmatrix} 1 \\ \sigma'(x_1) \end{pmatrix}$ 也与 ∇I 正交, 于是横截性条件变成

$$\nabla I \cdot \begin{pmatrix} 1 \\ \sigma'(x_1) \end{pmatrix} = 0. \quad (3-100)$$

对于控制问题, 根据(3—96)

$$\nabla I = \begin{pmatrix} H(x) \\ -\lambda(x) \end{pmatrix}$$

所以控制理论的横截性条件为

$$\begin{pmatrix} H(x_1) \\ -\lambda(x_1) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \sigma'(x_1) \end{pmatrix} = 0 \quad (3-101)$$

即

$$H(x_1) = \lambda(x_1) \sigma'(x_1)$$

§8 控制问题的约束条件

(1) 对控制变量有等式约束

设有两个控制变量 u_1, u_2 , 约束式为

$$C(x, u_1, u_2) = 0 \quad (3-102)$$

显然, 若只有一个控制变量, 则上式确定了唯一的 $u(x)$, 不存在对控制变量进行优选的问

题。

目标泛函为

$$\text{Max: } \int_{x_0}^{x_1} F(x, y_1, y_2, u_1, u_2) dx \quad (3-103)$$

状态方程为

$$y_1'(x) = g_1(x, y_1, y_2, u_1, u_2) \quad (3-104)$$

$$y_2'(x) = g_2(x, y_1, y_2, u_1, u_2)$$

作扩大的汉氏函数

$$H^* = F + \lambda_1 g_1 + \lambda_2 g_2 - vC = H - vC \quad (3-105)$$

其中 H 是一般控制问题中的汉氏函数,

$$\text{由于} \quad \frac{\partial H^*}{\partial \lambda_i} = \frac{\partial H}{\partial \lambda_i}, \quad \frac{\partial H^*}{\partial y_i} = \frac{\partial H}{\partial y_i},$$

所以必要条件(3-86a)和(3-86c)形式不变, 只是将 H 变成 H^* 而已, 但(3-86b)变成

$$\frac{\partial H^*}{\partial u_i} = \frac{\partial F}{\partial u_i} + \lambda_1 \frac{\partial g_1}{\partial u_i} + \lambda_2 \frac{\partial g_2}{\partial u_i} - v \frac{\partial C}{\partial u_i} = 0. \quad (3-105)$$

以上讨论不难推广到含多个控制变量, 多个状态变量和多个对控制变量约束的情形。

若约束条件(3-102)中包含状态变量, 例如

$$P(x, y_1, y_2, u_1, u_2) = 0 \quad (3-106)$$

则汉氏函数为

$$H^* = F + \lambda_1 g_1 + \lambda_2 g_2 - vP \quad (3-107)$$

根据必要条件(3-86), 我们有

$$-\frac{\partial H^*}{\partial y_i} = -\frac{\partial H}{\partial y_i} + v \frac{\partial P}{\partial y_i} = \lambda_i \quad (3-108a)$$

$$\frac{\partial H^*}{\partial u_i} = \frac{\partial H}{\partial u_i} - v \frac{\partial P}{\partial u_i} = 0 \quad (3-108b)$$

$$\frac{\partial H^*}{\partial \lambda_i} = \frac{\partial H}{\partial \lambda_i} = \dot{y}_i \quad (3-108c)$$

若(3-106)那样的约束有多个, 则每一约束都可乘上乘子 v_i 写进汉氏函数, 而 $\frac{\partial H^*}{\partial y_i}$ 和 $\frac{\partial H^*}{\partial u_i}$ 的表达式就扩大了。

(2) 对状态变量的等式约束。

设有约束条件

$$r(x, y_1, y_2) = 0 \quad (3-109)$$

其中只包含状态变量而不含控制变量，从上式可导出

$$\frac{dr}{dx} = \frac{\partial r}{\partial x} + \frac{\partial r}{\partial y_1} \frac{dy_1}{dx} + \frac{\partial r}{\partial y_2} \frac{dy_2}{dx} = 0 \quad (3-110)$$

因为

$$y'_1(x) = g_1(x, y_1, y_2, u_1, u_2)$$

$$y'_2(x) = g_2(x, y_1, y_2, u_1, u_2),$$

所以

$$\frac{\partial r}{\partial x} + \frac{\partial r}{\partial y_1} g_1 + \frac{\partial r}{\partial y_2} g_2 = 0 \quad (3-111)$$

加上条件

$$r[x_0, y_1(x_0), y_2(x_0)] = 0 \text{ 或 } r[x_1, y_1(x_1), y_2(x_1)] = 0 \quad (3-112)$$

此式可用来弥补对 r 求导数时的信息损失，因为从(3-110)回到(3-109)会出现不定积分常数问题。由于(3-111)含有控制变量，所以它实际上是既含控制变量又含状态变量的等式约束 P ，因此可以按(3-108)中的必要条件求解。

例 设动态最优价格问题与前文中的例相同，端点条件为

$$t_0 = 0, p(t_0) = 70; t_1 = 52, p(t_1) = 100, \quad (1)$$

另外加上调价速度不能高于0.6的约束，即

$$u \leq 0.6 \quad (2)$$

根据前例，无约束最优解为

$$p(t) = 15.391e^{0.0173t} - 12.724e^{-0.0173t} + 67.333 \quad (3)$$

由于 $u = p'$ ，而

$$p'(t) = 0.266e^{0.0173t} + 0.220e^{-0.0173t}$$

在 $0 \leq t \leq 52$ ， p' 从0.486增至0.744，也就是说无约束解超出了 $u \leq 0.6$ 的范围，实际上当 $t \geq 33.8$ 时， $u > 0.6$ ，作汉氏函数

$$H = -\frac{3}{2} p^2 + 202p - 5,000u^2 - 10,200u + 200pu - 5240 + \lambda u,$$

其它最优化必要条件都与前相同，不过要使 H 在约束(2)下极大化，可运用 Kuhn-Tucker 法则。作

$$J = H - \mu(u - 0.6) \quad (4)$$

Kuhn-Tucker 条件为

$$\frac{\partial J}{\partial u} = -10,000u - 10,200 + 200p + \lambda - \mu = 0 \quad (5)$$

$$\mu(u - 0.6) = 0 \quad (6)$$

$$u \leq 0.6, \mu \geq 0 \quad (7)$$

(6)式意味着 $\mu=0$ 或 $u=0.6$ 至少二者必居其一。

(i) 令 $u=0.6$, 即 $p'(t)=0.6$, 积分得

$$p(t)=0.6t+k \quad (8)$$

由于 $p(t_0)=70$, 所以 $k=70$. 但若 $p(t_1)=100$, 则 $k=68.8$, 显然斜率为0.6的直线不联结 $p=70, t=0$; $p=100, t=52$. 联结这两点的直线斜率为0.577.

(ii) 若 $\mu=0$, 则(5)变成 $\frac{\partial H}{\partial u}=0$, 相应的解就是无约束最优解, 它在 $33.8 \leq t \leq 52$ 超出了条件(2), 可见最优解有隅角点, 它在开始时刻沿无约束最优曲线运动, 在最后时刻, 沿 $p(t)=0.6t+68.8$ 运动, 联立无约束最优解(3)和(8), (其中 $k=68.8$), 则有

$$15.391e^{0.0173t}-12.724e^{-0.0173t}+67.333=0.6t+68.8 \quad (9)$$

解出 $t^*=11.3$, 其时, $p(t^*)=75.58$. 于是约束最优解为

$$p(t)=\begin{cases} 15.391e^{0.173t}-12.724e^{-0.0173t}+67.333, & 0 \leq t < 11.3 \\ 0.6t+68.8 & 11.3 \leq t \leq 52. \end{cases}$$

分段积分, 可求出总的目标泛函极值为169,000. 请注意, 尽管原来的无约束最优解直到 $t=33.8$ 之后才超出对控制变量的约束, 但考虑价格变化的最终影响, 所以必须在较早时刻放弃原来的最优价格, 才能使价格变动的最终影响不超过约束。

§9 离散型最大值原理

经济计量模型大多数是离散型的, 所以在经济计量模型的基础上进行最优控制主要使用离散型最大值原理。

设有变分问题

$$\text{Max: } I[y]=\int_{x_0}^{x_1} F(x, y, \dot{y})dx \quad (3-113)$$

其中

$$y(x_0)=y_0, \quad y(x_1)=y_1,$$

令 $u=\dot{y}$, 则可写为控制问题

$$\text{Max: } I[y, u]=\int_{x_0}^{x_1} F(x, y, u)dx \quad (3-114)$$

其中

$$y(x_0)=y_0, \quad y(x_1)=y_1, \quad \text{且 } \dot{y}(x)=u(x).$$

我们可将状态方程 $\dot{y}(x)=u(x)$ 当作等式约束。

根据积分的定义, 我们可将(3-114)写成离散的近似式

$$\text{Max: } S[y, u]=\sum_{i=0}^{n-1} F(x_i, y_i, u_i)\Delta x \quad (3-115)$$

其中 y_0 和 y_n 给定, 且

$$\frac{y_{i+1}-y_i}{\Delta x}=u_i, \quad i=0, \dots, n-1$$

利用拉氏乘数法则作拉氏函数

$$L = \sum_{i=0}^{n-1} \left[F(x_i, y_i, u_i) \Delta x - \lambda_i \left(\frac{y_{i+1} - y_i}{\Delta x} - y_i \right) \right] \quad (3-116)$$

最优化的一阶条件为

$$\frac{\partial L}{\partial y_i} = \frac{\partial F}{\partial y_i} \Delta x - \frac{\lambda_{i-1}}{\Delta x} + \frac{\lambda_i}{\Delta x} = 0, \quad (i = 1, \dots, n-1) \quad (3-117a)$$

$$\frac{\partial L}{\partial u_i} = \frac{\partial F}{\partial u_i} \Delta x + \lambda_i = 0, \quad (i = 1, \dots, n-1) \quad (3-117b)$$

$$\frac{\partial L}{\partial \lambda_i} = - \left(\frac{y_{i+1} - y_i}{\Delta x} - y_i \right) = 0. \quad (i = 1, \dots, n-1) \quad (3-117c)$$

定义汉氏函数 H :

$$H(x_i, y_i, u_i, \lambda_i) = \sum_{i=0}^{n-1} [F(x_i, y_i, u_i) \Delta x + \lambda_i u_i] \quad (3-118)$$

于是

$$L = H - \sum_{i=0}^{n-1} \lambda_i \frac{y_{i+1} - y_i}{\Delta x}$$

且

$$\frac{\partial L}{\partial y_i} = \frac{\partial H}{\partial y_i} - \frac{\lambda_{i-1}}{\Delta x} + \frac{\lambda_i}{\Delta x}, \quad \frac{\partial L}{\partial u_i} = \frac{\partial H}{\partial u_i},$$

$$\frac{\partial L}{\partial \lambda_i} = \frac{\partial H}{\partial \lambda_i} = \frac{y_{i+1}}{\Delta x} - \frac{y_i}{\Delta x},$$

于是必要条件(3-117)可改写为

$$\frac{\partial H}{\partial y_i} = \frac{\lambda_{i-1} - \lambda_i}{\Delta x} \quad \text{或} \quad - \frac{\partial H}{\partial y_i} = \frac{\lambda_i - \lambda_{i-1}}{\Delta x} \quad (3-119a)$$

$$\frac{\partial H}{\partial u_i} = 0 \quad (3-119b)$$

$$\frac{\partial H}{\partial \lambda_i} = \frac{y_{i+1} - y_i}{\Delta x} \quad (3-119c)$$

(3-119a) 和 (3-119c) 是两个给定 y_i 初始值和终端值的差分方程, 而 (3-119b) 意味着 $u_i (i = 0, \dots, n-1)$ 使汉氏函数的值最大化。

对于自由右边界问题, y_n 不定, 则意味着(3-117a)中的 $\frac{\partial L}{\partial y_i}$ 中的 i 不仅包含 $i = 1, \dots, n-1$, 且包含 $i = n$, 不过由于 y_n, λ_n 在(3-115)的 F 中都没有, 所以 $\frac{\partial F}{\partial y_n} = 0$, $\lambda_n = 0$, 所以从(3-117a)可推出

$$- \frac{\lambda_{n-1}}{\Delta x} = 0 \quad (3-120)$$

这是右端点条件, 而左端点为给定的 y_0 。

第四章 经济管理中的最优控制问题

§1 最优经济增长和最优积累率

早在经济控制论诞生以前的1928年, Frank Ramsey 在 *Economic Journal* 上发表了一篇“储蓄的数学理论”, 此文用动态最优化方法解决了一种较简单的最优经济增长问题。最优经济增长问题一直是经济控制论中讨论的主要问题, 这方面积累的文献已多不胜数。Ramsey 关于最优经济增长的论文目前被当作经济控制论最早的文献被大家所引用。

最优经济增长问题就是我们常说的积累与消费的最优比例问题, 当前消费太多, 则积累太少这会使未来的消费不能迅速增加, 若积累太多, 虽可使未来消费较快增长, 但却使当前消费减少, 我们的问题是在某一时期适当安排各年积累消费比例, 以使总消费额最大化。这个问题最常见的形式还要引入效用函数, 这个效用函数一般设计成消费额之边际效用递减, 其意义是, 若某年消费额很多, 则继续增加消费额, 效用不大, 而当消费额很小时, 增加一点消费, 其效用较大。也有的模型将效用或消费额全部贴现, 要求某一时期内之贴现值最大; 还有的模型引入技术进步, 资本折旧, 人口增长等因素。

我们在此讨论将消费总额最大化作为目标泛函的最优经济增长问题。

设 t 年国民收入为 $Y(t)$, 固定资产为 $K(t)$, 职工人数为 $L(t)$, 则有 Coob—Douglas 生产函数

$$Y = AK^\alpha L^{1-\alpha}, \quad 0 < \alpha < 1 \quad (4-1)$$

设人均消费额为 $y = \frac{Y}{L}$, 人均固定资产为 $k = \frac{K}{L}$, 于是上式可改写为

$$\frac{Y}{L} = A \frac{K^\alpha}{L} \frac{L^{1-\alpha}}{L},$$

即

$$\frac{Y}{L} = A \left(\frac{K}{L} \right)^\alpha$$

$$y = Ak^\alpha. \quad (4-2)$$

又设 t 时刻的投资为 $I(t)$, 且固定资产折旧率为 μ , 于是我们有

$$I(t) = \dot{K}(t) + \mu K(t), \quad (4-3)$$

令人均投资额为 $i = \frac{I}{L}$, 则

$$i = \frac{\dot{K}}{L} + \mu k \quad (4-4)$$

$\therefore$

$$\dot{k} = \frac{d}{dt} \frac{K}{L} = (\dot{K}L + K\dot{L})L^{-2}$$

$$\dot{L} = \frac{K}{L} - k \frac{L}{L}, \quad (4-5)$$

若 t 时刻职工人数为初始时刻职工人数 L_0 和 t 的函数 $L(t) = L_0 e^{nt}$, 即职工人数增长率为

$$\frac{\dot{L}}{L} = n,$$

代入(4-5), 则

$$\dot{k} = \frac{K}{L} - nk$$

即

$$\frac{K}{L} = \dot{k} + nk. \quad (4-6)$$

将上式代入(4-4), 则

$$\dot{i} = \dot{k} + (\mu + n)k. \quad (4-7)$$

人均收入 $y(t)$ 为人均消费 $c(t)$ 和人均积累 $i(t)$ 之和, 即

$$y(t) = c(t) + i(t). \quad (4-8)$$

将(4-2)和(4-7)代入, 则

$$Ak^a = c(t) + \dot{k}(t) + (\mu + n)k(t)$$

即

$$\dot{k} = Ak^a - c - (\mu + n)k \quad (4-9)$$

这就是系统之状态方程, 其中 c 为控制变量, k 为状态变量, 我们将此状态方程作为约束条件。在此约束下, 目标泛函为在 $0-T$ 年内使人均消费总额极大化, 即

$$\text{Max: } \int_0^T c(t) dt \quad (4-10)$$

且 $k(0) = k_0$, $k(T)$ 不定, $\bar{c} \leq c(t) \leq y(t)$ 。

$\bar{c}$ 为人均消费水平的下界。

汉氏函数为

$$\begin{aligned} H &= c + \lambda [Ak^a - c - (\mu + n)k] \\ &= (1 - \lambda)c + [\lambda Ak^a - \lambda(\mu + n)k] \end{aligned} \quad (4-11)$$

H 是控制变量 c 的线性函数, 这是所谓 bang-bang 控制问题。从(4-11)不难看出, 要使 H 对 c 极大化, 则有如下三种情况:

$$\begin{array}{lll} \text{若 } \lambda > 1, & \text{则 } c^* = \bar{c} \\ \text{若 } \lambda = 1, & \text{则 } c^* = c(t) \\ \text{若 } \lambda < 1, & \text{则 } c^* = y(t) \end{array} \quad (4-12)$$

根据最大值原理, 必要条件还有

$$-H_k = \dot{\lambda}, \quad H_\lambda = \dot{k} \quad (4-13)$$

$$\therefore H_\lambda = -\lambda[\alpha A k^{\alpha-1} - (\mu+n)] \quad (4-14)$$

$$H_\lambda = A k^\alpha - c - (\mu+n)k \quad (4-15)$$

$$\therefore \dot{\lambda} = -\lambda[\alpha A k^{\alpha-1} - (\mu+n)] \quad (4-16)$$

$$\dot{k} = A k^\alpha - c - (\mu+n)k \quad (4-17)$$

当系统达到均衡时, $\dot{k}=0, \dot{\lambda}=0$, 代入(4-16)和(4-17), 则

$$\alpha A k^{*\alpha-1} = \mu+n$$

$$\text{即} \quad k^* = \left(\frac{\mu+n}{\alpha A} \right)^{\frac{1}{\alpha-1}} \quad (4-18)$$

$$\begin{aligned} \text{且} \quad c^* &= A k^{*\alpha} - (\mu+n)k^* \\ &= A k^{*\alpha} - \alpha A k^{*\alpha} \\ &= (1-\alpha) A k^{*\alpha} \end{aligned} \quad (4-19)$$

$\therefore k^* \neq 0, \alpha \neq 1, \therefore c^* \neq 0$, 又 $\therefore c^* \neq y(t) = A k^{*\alpha}$,

$\therefore \lambda$ 不大于 1, 且不小于 1, 即

$$\lambda = 1$$

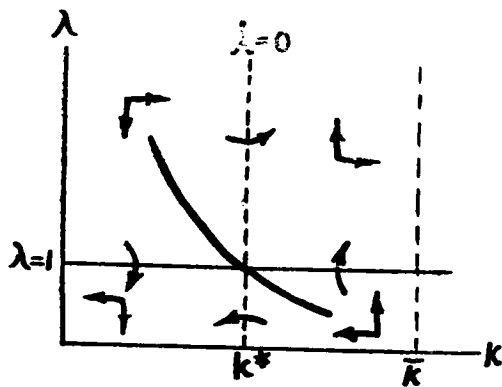


图 4-1

在 (λ, k) 相空间作图 4-1, 当 $\frac{dy}{dk} = \alpha A k^{\alpha-1} = \mu+n$ 时, 即 $k=k^*$, 则 $\dot{\lambda}=0$, 若 $k < k^*$, 则 $\frac{dy}{dk} > \mu+n$, 考虑(4-16), 不难看出此时 $\dot{\lambda} < 0$, λ 递减。若 $\lambda < 1$, 根据(4-12), 我们有 $c^* = y(t) = A k^\alpha$, 代入(4-17), 则

$$\dot{k} = A k^\alpha - (\mu+n)k$$

$\therefore$ 上式小于 0, 所以 k 渐减, 因此在相图 $\lambda=1$ 的下部, $k=k^*$ 左部, 相轨迹运动方向是左下方。

当 $\lambda > 1$ 时, $c^* = 0$, 且

$$\dot{k} = A k^\alpha - (\mu+n)k,$$

由于总收入 $y = A k^\alpha$ 大于 $(\mu+n)k$, 所以 $\dot{k} > 0$, 因此 k 渐增。于是在相图 4-1 的 $\lambda > 1$ 区域, 有向右的水平箭头。

$\because k=k^*$ 时, $\dot{\lambda}=0$, $\therefore$ 在相图中虚线右边, $\dot{\lambda}>0$, 有平行于 λ 轴向上的箭头; 在虚线左边, $\dot{\lambda}<0$, 有平行于 λ 轴向下的箭头。

我们利用相图判断了解的特性, 下面我们讨论初始条件不同情况对解的影响。

(i) 若人均固定资产之初始值 $k_0 < k^*$, 则 $\frac{dy}{dk} > (\mu+n)$, $\therefore \dot{\lambda} < 0$ 。由于 k_0 在 k^* 左边, 所以 k 将向右运动, 从 k_0 趋于 k^* , 又由于 $\dot{\lambda} < 0$, 所以 λ 渐减, 它向下运动, 从相图 4—1 不难看出, k 向右, λ 向下, 则是 $k < k^*, \lambda > 1$ 之区域。根据 (4—12), $\lambda > 1$ 时, $c = \bar{c}$, 消费从最低水平 $\bar{c}$ 起始。这时资本以其最大速率

$$\dot{k} = Ak^\alpha - (\mu+n)k \quad (4-21)$$

增长。利用初始条件 $k(0) = k_0$ 和端点条件 $k(t^*) = k^*$, 我们可用龙格—库塔法求上列微分方程的数值解。此法的计算公式为

$$k_{n+1} = k_n + \frac{1}{h} (a_1 + 2a_2 + 2a_3 + a_4) \quad (4-22)$$

取步长 $h = \frac{1}{2}$ 年, 则

$$\begin{aligned} a_1 &= \frac{1}{2} [Ak_n^\alpha - (\mu+n)k_n], \\ a_2 &= \frac{1}{2} \left[A \left(k_n + \frac{a_1}{2} \right)^\alpha - (\mu+n) \left(k_n + \frac{a_1}{2} \right) \right], \\ a_3 &= \frac{1}{2} \left[A \left(k_n + \frac{a_2}{2} \right)^\alpha - (\mu+n) \left(k_n + \frac{a_2}{2} \right) \right], \\ a_4 &= \frac{1}{2} [A(k_n + a_3) - (\mu+n)(k_n + a_3)]. \end{aligned}$$

由于 k^* 已由 (4—18) 给出, 所以用以上迭代公式, 计算到 $k_{n+1} = k^*$ 时, 则 t^* 亦近似解出。 $k < k^*$ 时的时间轨线如图 (4—2) 所示。

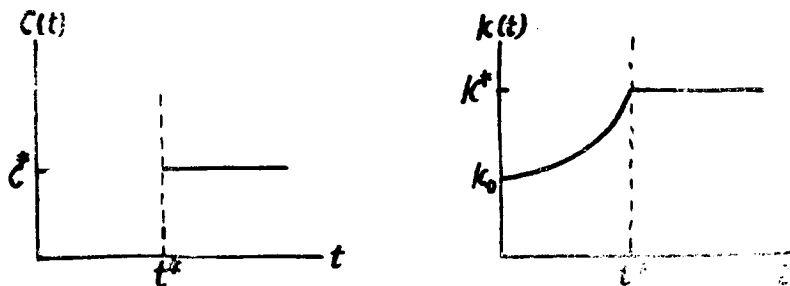


图 4—2

当 $k = k^*$ 时, 根据 (4—20), $\lambda = 1$, 于是消费转换为 (4—19) 给出的常数水平

$$c^* = (1 - \alpha) A k^{*\alpha}. \quad (4-23)$$

最优消费水平和最优积累水平如图4-2所示。

设最优积累率为 α , 则

$$\alpha = \frac{i}{c+i} = \frac{k}{c+k} \quad (4-24)$$

于是我们根据以上求解过程制定最优积累率计划为

$$\alpha^* = \begin{cases} \frac{k(t)}{c+k(t)}, & t < t^* \\ \frac{k^*}{c^*+k^*}, & t > t^* \end{cases} \quad (4-25)$$

其中 k^* , c^* 由(4-18), (4-19)给出, $k(t)$ 由(4-21)确定, 加上端点条件, 用数值法求解(4-21)不但能求出 $k(t)$ 之时间轨线 (在 $0 < t < t^*$), 且能确定 t^* 之近似值。

(ii) 若初始时刻人均固定资产 $k_0 > k^*$, 则

$$\frac{dy}{dk} < \mu + n, \quad \therefore \dot{\lambda} > 0,$$

λ 随时间递增。此种情况下相图 4-1 中的 k 在 k^* 之右, k 从 k_0 向左朝 k^* 运动, λ 从下向上运动, 所以系统应处于 $\lambda < 1, k > k^*$ 区域, 根据(4-12), 当 $\lambda < 1$ 时, $c(t)$ 在 $c = y(t)$ 起始, 这是 c 之最大值, 见图4-3. c 保持在这一水平直到 $t = t^*$. 也就是说, 如果以前的固定资产 k 超过了均衡水平 k^* , 则将收入全用于消费, 让固定资产通过折旧尽快从 k_0 降至 k^* .

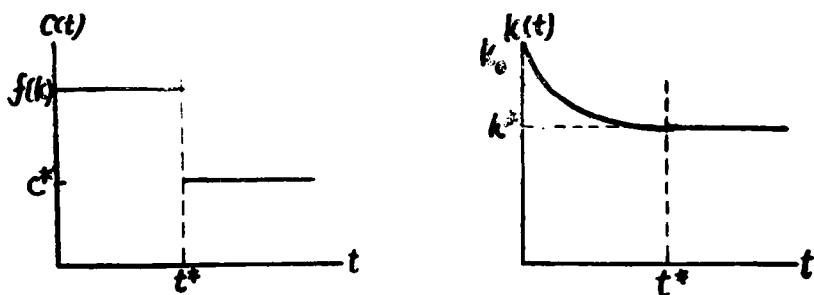


图 4-3

将 $c(t) = y(t) = A k^\alpha$ 代入(4-17), 则得到

$$\dot{k} = -(\mu + n)k \quad (4-26)$$

解之得

$$k^* = k_0 e^{-(\mu+n)t^*} \quad (4-27)$$

从中求出

$$t^* = \frac{\ln k^* - \ln k_0}{\mu + n} \quad (4-28)$$

k^* 由(4-18)给出。(4-28)给出了从完全不投资到投资的最优转换时刻。由于积累率

$$\begin{aligned}\alpha(t) &= \frac{i(t)}{i(t) + c(t)} \\ &= \frac{k(t)}{k(t) + c(t)},\end{aligned}$$

当 $t < t^*$ 时, $k(t)$ 由(4-21)给出, 所以

$$\alpha(t) = \frac{Ak^x - (\mu + n)k}{Ak^x - (\mu + n)k + \bar{c}}, \quad (4-29)$$

其中 k 之时间轨线由迭代公式(4-22)加上端点条件近似确定。

当 $t > t^*$ 时, k 由(4-26)给出, 其中 k 是(4-26)之解, 即

$$k = k_0 e^{-(\mu+n)t}. \quad (4-30)$$

所以最优积累率为

$$\begin{aligned}\alpha(t) &= \frac{k^*}{k^* + c^*} \\ &= \frac{-(\mu + n)k^*}{-(\mu + n)k^* + c^*} \\ &= \frac{-(\mu + n)k_0 e^{-(\mu+n)t^*}}{-(\mu + n)k_0 e^{-(\mu+n)t^*} + c^*}\end{aligned} \quad (4-31)$$

于是我们有最优积累率计划:

$$\alpha^* = \begin{cases} \frac{Ak^x - (\mu + n)k}{Ak^x - (\mu + n)k + \bar{c}} & t < t^*, \\ \frac{-(\mu + n)k_0 e^{-(\mu+n)t^*}}{-(\mu + n)k_0 e^{-(\mu+n)t^*} + c^*} & t > t^*. \end{cases} \quad (4-32)$$

以上宏观经济最优增长控制问题可以推广到企业管理问题中。我们将 $c(t)$ 改为企业分配的工资、奖金、福利基金, 将 $i(t)$ 改为企业的投资量, $k(t)$ 改为企业的固定资产, μ 仍为折旧率, n 为职工增长率, 于是我们就可将以上模型改为企业最优投资计划模型。这方面的企业管理动态模型已有很多文献讨论, 例如 Solow[1971], Lesourne[1973], Hochman[1973], Jacquemin[1971, 1972], Gaskins[1971], Kamien[1971], Ruff[1969], 等等。众多的模型中有的以利润最大化为目标泛函, 有的以销售额最大化为目标泛函, 有的模型加入了对借款及劳力供应的约束条件, 有的模型还考虑到劳动力与固定资产之间的替代, 组合问题, 很多文献都着重讨论怎样调节控制生产增长率和最优规模, 以使目标泛函极大化。

§2 最优库存控制

设某厂的生产量为 $u(t)$, 它是企业可以调节的控制变量。市场对此厂产品的需求量 $D(t)$, $D(t)$ 是种季节性商品, 它是时间 t 之周期波动函数

$$D = [1.2 + 0.5 \sin(30t + 30)]e^{0.1t} \quad (4-33)$$

又设库存量为 $x(t)$, $x(t)$ 与 $u(t)$, $D(t)$ 的关系为

$$\dot{x}(t) = u(t) - D(t) \quad (4-34)$$

这就是状态方程, 其初始条件为

$$x(0) = x_0 \quad (3-34')$$

需求量 $D(t)$ 与生产量 $u(t)$ 之差额平方越大, 则经济损失越大, 设其权重系数为 c_1 , 当 $D(t) > u(t)$ 时, 这种费用主要是缺货的损失, 当 $D(t) < u(t)$ 时, 这种费用主要是处理过多产品的代价。设库存量的所欲水平为 $\bar{x}(t)$, 我们也可以设 $\bar{x}$ 为一常数或等于 0. $x(t)$ 与所欲水平的差额之平方越大, 则费用也越大, 其权重系数设为 c_2 , 当 $x(t) > \bar{x}(t)$ 时, 这种费用主要是库存费用, 当 $x(t) < \bar{x}(t)$ 时, 费用主要是缺少库存担的缺货风险代价。我们的目的是要求总费用在 $0 \sim T$ 时期内最小化, 即目标泛函为

$$\text{Min: } \int_0^T \{c_1[u(t) - D(t)]^2 + c_2[x(t) - \bar{x}(t)]^2\} dt, \quad (4-35)$$

约束条件就是状态方程(4-34)及初始条件(4-34')。

作汉氏函数

$$H = c_1(u - D)^2 + c_2(x - \bar{x})^2 + \lambda(u - D), \quad (4-36)$$

必要条件为

$$-\frac{\partial H}{\partial x} = \dot{\lambda} = 2c_2(x - \bar{x}) \quad (4-37a)$$

$$\frac{\partial H}{\partial \lambda} = \dot{x} = u - D \quad (4-37b)$$

$$\frac{\partial H}{\partial u} = 2c_1(u - D) + \lambda \quad (4-37c)$$

从(4-37c)可解出

$$u = D + \frac{\lambda}{2c_1} \quad (4-38)$$

代入(4-37b), 则

$$\begin{aligned} \dot{x} &= D + \frac{\lambda}{2c_1} - D \\ &= \frac{\lambda}{2c_1} \end{aligned}$$

联立(4-39)和(4-37a)

$$\begin{cases} \dot{x} = \frac{\lambda}{2c_1} \\ \dot{\lambda} = 2c_2(x - \bar{x}) \end{cases} \quad (4-40)$$

我们可视上式为--状态方程,若所欲库存水平 $\bar{x}=0$, 则状态方程为

$$\begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2c_1} \\ 2e_1 & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} \quad (4-41)$$

设其解为

$$x = k_1 e^{\frac{1}{c_1}t} + k_2 e^{c_2 t} \quad (4-42)$$

$$\lambda = k'_1 e^{\frac{1}{c_1}t} + k'_2 e^{c_2 t}$$

将其对 t 求导数, 则,

$$\dot{x} = \frac{k_1}{c_1} e^{\frac{1}{c_1}t} + c_2 k_2 e^{c_2 t} = \frac{\lambda}{2c_1} = \frac{1}{2c_1} (k'_1 e^{\frac{1}{c_1}t} + k'_2 e^{c_2 t})$$

$$\dot{\lambda} = \frac{k'_1}{c_1} e^{\frac{1}{c_1}t} + c_2 k'_2 e^{c_2 t} = 2c_2 x = 2c_2 (k_1 e^{\frac{1}{c_1}t} + k_2 e^{c_2 t})$$

比较等号两边同类项的系数, 则可解出

$$k_1 = \frac{1}{2} k'_1 \quad (4-43)$$

$$k_2 = \frac{1}{2} k'_2$$

于是通解可写为

$$\begin{cases} x = k_1 e^{\frac{1}{c_1}t} + k_2 e^{c_2 t} \\ \lambda = 2k_1 e^{\frac{1}{c_1}t} + 2k_2 e^{c_2 t} \end{cases} \quad (4-44)$$

常数 k_1, k_2 由初始条件和右端点条件 $\lambda(T)=0$ 确定。 $\lambda(T)=0$ 意味着在终端时刻 T , 库存变化率的影子价格为 0, 即不管库存量怎样调节, 都不能使目标泛函极值再减少一点。

将(4-44)代入(4-38), 则最优生产计划为

$$u = D + \frac{1}{c_1} (k_1 e^{\frac{1}{c_1}t} + k_2 e^{c_2 t}) \quad (4-45)$$

本节讨论的二次费用函数的控制问题在经济控制中有特别重要的意义, 一方面这种问题可以采用所谓**线性决策法则**, 这类问题又叫**线性二次问题**; 另一方面二次泛函目标在随机控制条件下的最优解等价于确定性最优解, 我们将在下一章讨论这一著名的“必然等价性原理”。这两个特点使二次费用泛函问题的控制法则特别适于应用。

我们首先作一种假想, 即动态影子价格可表为库存量 $x(t)$ 的线性函数, 而其参数为时间 t 的函数, 即

$$\lambda(t) = a(t) + b(t)x(t) \quad (4-46)$$

然后我们再来小心地证明这一假定是否成立。将上式对 t 求导数, 则有

$$\dot{\lambda}(t) = \dot{a}(t) + \dot{b}(t)x(t) + b(t)\dot{x}(t) \quad (4-47)$$

将(4-37a)和(4-39)代入, 我们得到

$$2c_2(x - \bar{x}) = \dot{a} + \dot{b}x + b\frac{\lambda}{2c_1} \quad (4-48)$$

将(4-46)代替上式中的 λ , 且合并 x 的同类项, 我们求得

$$\dot{a} + \frac{ab}{2c_1} + 2c_2\bar{x} - \left(2c_2 - \frac{b^2}{2c_1} - \dot{b}\right)x = 0.$$

当 t 为不同值时, x 不管为何值, 上式都要成立, 则要求

$$\dot{a} + \frac{ab}{c_1} + 2c_2\bar{x} = 0 \quad (4-50a)$$

$$\dot{b} + \frac{b^2}{2c_1} - 2c_1 = 0 \quad (4-50b)$$

由于 $\lambda(T) = 0$, 所以还有边界条件

$$a(T) = 0, \quad b(T) = 0.$$

(4-50b) 的 b 仅依赖于常数 c_1, c_2 , 它与 $x(t), u(t), \lambda(t)$ 都无关, 而 (4-50a) 中的 a 依赖于 b 和常数 c_1, c_2 以及给定的所欲库存水平 $\bar{x}(t)$, a 亦与 $x(t), u(t), \lambda(t)$ 无关, 可见, 我们在(4-47)中对于 $a(t), b(t)$ 的假定成立。由于(4-38)中的最优生产速率调节计划是 λ 的线性函数, 将(4-46)代入, 则

$$u = D + \frac{1}{2c_1}(a + bx) \quad (4-51)$$

即最优生产速率根据两个因素调节, 一个是需求, 另一个是当时的库存水平 $x(t)$ 。

(4-50)为可变量分离微分方程, 解之得

$$\frac{db}{dt} = \frac{(2c_1)^2 - b^2}{2c_1}$$

$$\int \frac{db}{(2c_1)^2 - b^2} = \int \frac{dt}{2c_1}$$

上式左边为

$$\begin{aligned} & \int \frac{(2c_1 - b) \cdot 2(2c_1)db}{2(2c_1)(2c_1 + b)(2c_1 - b)^2} \\ &= \int \frac{(2c_1 - b)(2c_1 - b + 2c_1 + b)db}{2(2c_1)(2c_1 + b)(2c_1 - b)^2} \\ &= \frac{1}{2(2c_1)} \int d\left(\ln \frac{2c_1 + b}{2c_1 - b}\right) \end{aligned}$$

$$= \frac{1}{2(2c_1)} \ln \frac{2c_1+b}{2c_1-b} \quad (4-52)$$

根据导数法则

$$\frac{d}{dx}(\text{Arth } x) = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (\text{Arth 为反双曲正切})$$

$$\begin{aligned} \therefore \frac{1}{2c_1} d \text{Arth } \frac{b}{2c_1} &= \frac{1}{2c_1} \left[\frac{1}{2} d \ln \frac{2c_1+b}{2c_1-b} \right] \\ &= \frac{1}{2(2c_1)} d \left(\ln \frac{2c_1+b}{2c_1-b} \right) \end{aligned}$$

$\therefore$ (4-52)可改写为

$$\begin{aligned} &\frac{1}{2(2c_1)} \int d \left(\ln \frac{2c_1+b}{2c_1-b} \right) \\ &= \frac{1}{2c_1} \int d \text{Arth } \frac{b}{2c_1} \\ &= \frac{1}{2c_1} \text{Arth } \frac{b}{2c_1} \end{aligned} \quad (4-53)$$

于是(4-50b)之解为

$$\begin{aligned} \pm \frac{1}{2c_1} \text{Arth } \frac{b}{2c_1} &= \frac{t+C}{2c_1} \\ \frac{b}{2c_1} &= \text{th}(\pm t+C) \\ b &= 2c_1 \text{th}(\pm t+C) \end{aligned} \quad (4-54)$$

上式中的 $\pm$ 号, 当 $|b| < |2c_1|$ 时, 正号成立, 当 $|b| > |2c_1|$ 时, 负号成立。常数 C 由 $b(T)=0$ 解出。将 b 代入(4-50a), 由于是一阶线性微分方程, 也不难解出 $a(t)$ 。

当所欲的库存水平 $\bar{x}(t)$ 不为0且不为常数时, 我们就碰到解非齐次线性微分方程组(4-40)的问题, 用一般的方法很难解出这类问题, 而利用上述线性决策法则我们可以解得 $a(t), b(t)$, 将 $a(t), b(t)$ 代入(4-16)则可解出 λ , 再将 $\lambda(t)$ 代入(4-38): $u = D + \frac{\lambda}{2c_1}$, 于是可解最优产量变化率计划, 这就是线性决策法则解决复杂的最优控制问题之功效。

运筹学和数理经济学中最著名的库存问题就是所谓 Welson 批量公式。这个公式中的需求为不变常数, 而最优库存量也是不变常数, 即

$$x^* = \sqrt{\frac{2c_3 D}{c_1 T}}$$

其中 x^* 为最优库存量, c_1 为单位产品单位时间的库存费用, c_3 为每批产品的订货费用, D 为总需求量, T 为所考虑的时间长度。 x^*, c_1, c_3, D, T 都是常数, Welson 批量公式的主要思路是考虑到库存问题中有两种互相矛盾的费用, 一种是定货费用, 它只与定货次数有

关,另一种是库存费用,它正比于库存量,定货次数越多,总的定货费用越多,而库存费用越少,而定货次数少,则总的定货费用少,但库存费用上升。最优化就是兼顾这两方面以使订货费用与库存费用的总和最小。

若需求随时间变化,则最优库存量也会是时间的函数,很多人研究了这类动态最优库存问题。大多数作者都采用了离散模型。这方面最著名的有 Scarf 1960 年的文献及 Modigliani 和 Hohn 1955 年的文献等等。我们在此讨论一个需求和库存随时间连续变化,而订货是间断进行的动态库存模型,通过对这个模型的讨论我们可以初步了解经济控制中著名的**两点边值问题**,以及连续模型与离散模型互相变换的技巧,在此基础上我们可以了解最优控制的数值解法的基本思路。

设每次定货费用为 K ,定货次数为 N ,则总的订货费用为 KN 。库存费用为 $c(t)$,库存量为 $x(t)$,则 $0-T$ 时期的总库存费用为

$$\int_0^T c(t)x(t)dt.$$

为了将间断的定货与连续的需求、库存联结起来,我们引入一些新的符号,设两次相邻的订货时间为 t_i, t_{i-1} ,且两次订货之间的时间间隔为

$$\tau_i \equiv t_i - t_{i-1}, \quad i = 1, \dots, N \quad (4-55)$$

其中 $\tau_1 = t_1, \tau_N = T - t_{N-1}$,这意味着变量 $\tau_1, \dots, \tau_N$ 必须满足

$$\tau_i \geq 0, \quad \sum_{i=1}^N \tau_i = T \quad (4-56)$$

为了将离散的 t_i, t_{i-1} 与连续的表示式 t 挂上勾,我们令

$$t = t_{i-1} + s\tau_i, \quad t_{i-1} \leq t \leq t_i \quad (4-57)$$

然后引入函数

$$x_i(s) = x(t) = x(t_{i-1} + s\tau_i) \quad (4-58)$$

$$t_{i-1} \leq t \leq t_i, \quad i = 1, \dots, N$$

将 $x_i(s)$ 对 s 求导数,则

$$\frac{dx_i(s)}{ds} = \tau_i \dot{x}, \quad 0 \leq s \leq 1, \quad i = 1, \dots, N \quad (4-59)$$

$$\text{且,} \quad x_i(0) = x_{i-1}(1) + u_{i-1} \quad i = 1, \dots, N \quad (4-60)$$

$$\text{即} \quad x(t_{i-1}) = x(t_{i-2} + \tau_{i-1}) + u_{i-1}$$

$$\text{其中} \quad x_0(1) = x_0 = 0, u_{i-1} \text{ 为 } i-1 \text{ 阶段的阶跃定货量。}$$

(4-58), (4-59) 将离散型 i 阶段的需求量 $x_i(s)$ 与连续型 t 时刻需求量 $x(t)$ 联系起来了,而(4-60)意味着 $i-1$ 次订货时的库存量等于原有库存量与订货量之和。

设 D 为累积需求量,则 i 阶段 s 时刻的库存量 $x_i(s)$ 为累积订货量 $\sum_{j=1}^{i-1} u_j$ 与 i 阶段 s

时刻的累积需求量 $D_i(t_{i-1} + s\tau_i)$ 之差, 即

$$x_i(s) = \sum_{j=0}^{i-1} u_j - D(t_{i-1} + s\tau_i) \quad (4-61)$$

$$0 \leq s \leq 1, i = 1, \dots, N$$

对上式我们有约束

$$x_i(s) \geq 0, \text{ 即 } \sum_{j=0}^{i-1} u_j \geq D(t_{i-1} + s\tau_i), i = 1, \dots, N \quad (4-62)$$

也就是说累积需求量不能超过累积定货量, 即不允许出现缺货。

目标泛函为

$$\begin{aligned} \text{Min: } & KN + \int_0^T c(t)x(t)dt \\ &= KN + \int_0^{t_1} c(t)x(t)dt + \int_{t_1}^{t_2} c(t)x(t)dt + \dots + \int_{t_{N-1}}^T c(t)x(t)dt \\ &= KN + \sum_{i=1}^N \int_0^1 c(t_{i-1} + s\tau_i)x_i(s)ds \end{aligned} \quad (4-63)$$

若库存费用 $c(t)$ 为常数 c , 且将(4-61)代入, 则上式可改写为

$$\begin{aligned} & KN + \sum_{i=1}^N \int_0^1 cx_i(s)ds \\ &= KN + \sum_{i=1}^N \int_0^1 c \left[\sum_{j=0}^{i-1} u_j - D(t_{i-1} + s\tau_i) \right] ds \\ &= KN + \sum_{i=1}^N \sum_{j=0}^{i-1} u_j \int_{t_{i-1}}^{t_i} c dt - \int_0^T cD(t)dt \end{aligned} \quad (4-63')$$

由于(4-61), (4-62)要求对于所有 $i = 1, \dots, N$ 成立, D 又是单调不减的函数, 所以我们将(4-62)改写成更为苛刻的约束

$$\sum_{j=0}^{i-1} u_j \geq D(t_i), \quad i = 1, \dots, N \quad (4-64)$$

u_j 和 D 都不小于 0, 在这种条件下, 只有当下式成立时, 目标泛函(4-63')才能最小化,

$$\sum_{j=0}^{i-1} u_j = D(t_i), \quad i = 1, \dots, N \quad (4-65)$$

利用上式可推出

$$\begin{aligned} u_i &= \sum_{j=0}^i u_j - \sum_{j=0}^{i-1} u_j = D(t_{i+1}) - D(t_i) \\ u_0 &= D(t_1) \end{aligned} \quad (4-66)$$

这就是给定 $t_i (i = 1, \dots, N)$ 时, 计算订货量的公式。

将(4—65)代入(4—63'), 则目标泛函为

$$KN + \sum_{i=1}^N D(t_i) \int_{t_{i-1}}^{t_i} c dt - \int_0^T c D(t) dt \quad (4-67)$$

将(4—67)对 t_i 求导数且令其等于 0, 由于 $t = t_{i-1} + s\tau$ 于是我们有

$$d(t_i) \int_{t_{i-1}}^{t_i} c dt = c [D(t_{i-1}) - D(t_i)], \quad i = 1, \dots, N-1$$

其中 $d(t_i) = D'(t_i)$ 为需求速率, 根据初始假定 $t_0 = 0, t_N = T$, 则

$$\begin{aligned} d(t_i) [t_i - t_{i-1}] &= D(t_{i+1}) - D(t_i) \\ i &= 1, \dots, N-1 \end{aligned} \quad (4-68)$$

从上式可求得订货时间的递推计算公式

$$t_{i-1} = t_i - \frac{D(t_{i+1}) + D(t_i)}{d(t_i)}, \quad i = 1, \dots, N-1, \quad (4-69)$$

一般最优订时间都是用离散数字解法, 若我们给定最初和最后一次订货时间, 但却不知道这之间的定货时间, 这类问题用两点边值问题的数字解法非常有效。

我们可首先选定一个 N , 然后按下列步骤进行计算:

(1) 估计 t_{N-1} 的值;

(2) 利用 $t_N = T$ 及给定的 t_{N-1} 值从(4—69)计算 t_{N-2} ;

(3) 不断进行迭代直到算出 t_0 。

(4) 若算得的 t_0 与给定 N 时的 t_0 在允许的误差范围内相等, 则停止计算, 否则改变假定的 t_{N-1} , 再重新计算, 直至求得满意的 t_0 为止。

(5) 利用(4—55)计算

$$\tau_i = t_i - t_{i-1}$$

(6) 利用(4—66)计算

$$u_i = D(t_{i+1}) - D(t_i), \quad u_0 = D(t_1);$$

(7) 将 u_j 值和 τ_i 值代入(4—61), 计算

$$x_i(s) = \sum_{j=0}^{i-1} u_j - D(t_{i-1} + s\tau_i) = x(t)$$

(8) 将算出的 $t_i, \tau_i, x_i(s)$ 及给定的 K, N 代入目标泛函(4—63), 算出目标值;

(9) 改变给定的 N , 重复(1)—(8)的程序, 然后比较各次计算的目标值, 只要某个 N , 使得 $N-1$ 和 $N+1$ 的目标值都高于 N 的目标值, 则 N 和相应的 t_i 就是最优解。

我们用一个例子来说明这种数值解法:

例 某种季节性商品之需求模型为

$$d(t) = (A\pi/2T) \sin(\pi t/T) \quad (1)$$

其累积需求函数为

$$D(t) = \frac{1}{2}A[1 - \cos(\pi t/T)] \quad (2)$$

终端时刻 T 的累积需求量为 A , 单位库存费用为 c , 每次订货费用为 K , 某种具体的商品对于上列参数我们有具体取值:

$$A = 10,000, T = 1, K = 100, c = 0.5$$

代入(1), (2), 则

$$d(t) = 5,500\pi \sin(\pi t) \quad (3)$$

$$D(t) = 5,000[1 - \cos(\pi t)] \quad (4)$$

若需求速率保持在 $A = 10,000$ 不变, 则用 Wilson 批量公式可算出最优定货批量为

$$x^* = \sqrt{\frac{2KA}{cT}} = 2,000 \quad (5)$$

$$N^* = 5 \quad (6)$$

这时最小总费用为 1,000。我们将这种静态最优解 $N^* = 5$ 作为动态问题的第一个试算解。

(1) 可令 $t_4 = 0.8$, 然后利用迭代公式(4—69)计算 t_3, t_2, t_1, t_0 , 算出 $t_0 \neq 0, N \neq 5$ 。

(2) 改变 t_4 值, 直到用迭代公式算得 $t_0 = 0, N = 5$ 为止, 这时算得

$$t_0 = 0$$

$$t_1 = 0.210$$

$$t_2 = 0.374$$

$$t_3 = 0.529$$

$$t_4 = 0.694$$

(3) 利用(4—55)计算 τ_i 值, 得

$$\tau_1 = t_1 - t_0 = 0.210$$

$$\tau_2 = t_2 - t_1 = 0.164$$

$$\tau_3 = t_3 - t_2 = 0.155 \quad (8)$$

$$\tau_4 = t_4 - t_3 = 0.165$$

$$\tau_5 = t_5 - t_4 = 0.306$$

(4) 利用(4—66)可算得各次订货量为

$$u_0 = 1050$$

$$u_1 = 2022$$

$$u_2 = 2376 \quad (9)$$

$$u_3 = 2418$$

$$u_4 = 2134$$

(5) 将 u_j 值和 τ_i 值代入(4—61), 计算

$$x(t) = x_i(s) \quad (10)$$

(6) 将 $K = 100$, $N = 5$, t_i , τ_i , $x_i(s)$ 代入(4—63)算出目标泛函值为 963.46。

(7) 令 $N = 3$, $N = 4$, 重复上列程序, 算得目标泛函值都大于 963.46, 于是求得 $N^* = 5$ 为最优订货次数; 最优订货时间由(7)式给出, 最优订货量由(9)式给出。

§3 最优设备更新和维修计划

对最优设备更新问题的最早文献是 Taylor 的 1923 年著作, 其次 Hotelling 1925 年, Preinrei 1940 年 Bain 1939 年, Bellman 1955 年, Liljeblad 和 Teborgh 1949 年等文献进一步对这一问题的研究作出了贡献。

设某一设备的质量水平为 $x(t)$, 在没有维修的情况下, 质量水平以一比率 α 下降, 又设维修投资为 $u(t)$, 则用货币量表示的设备质量 x 之变化率为

$$\begin{aligned} \dot{x} &= u(t) - \alpha x(t) \\ x(0) &= x_0 \end{aligned} \quad (4-70)$$

上式中的 u 为控制变量, x 为状态变量。

机器运转的收益正比于机器维护的质量水平 $x(t)$, 设其为 $V(t)$, 则

$$V(t) = vx(t) \quad (4-71)$$

若设备在 t 时刻的残值为

$$S[x(t)] = kx(t) \quad (4-72)$$

其中 k 为比例系数。于是目标泛函为

$$\text{Max: } \int_0^T [vx(t) - u(t)] e^{-rt} dt + kx(T)e^{-rT} \quad (4-73)$$

上式第二项未在积分号内, 这里的目标泛函不象上文中的目标泛函是一个积分, 而是一个积分项与一般项组成, 这类问题可写成更一般的形式

$$\text{Max: } I = \int_0^T F(x, u, t) dt + K[x(T)] \quad (4-74)$$

$$\text{其中} \quad x(0) = x_0, \quad \dot{x} = f(x, u, t) \quad (4-74')$$

我们可定义一个新的状态变量

$$\begin{aligned} \dot{y}(t) &= F(x, u, t) \\ y(0) &= 0 \end{aligned} \quad (4-75)$$

于是(4-74)可写为

$$\begin{aligned}\text{Max } I &= y(T) + K[x(T)] \\ &= G[x(T)]\end{aligned}\quad (4-76)$$

若(4-74)中的 $K=0$, 这就是前文中讨论的一般目标泛函形式, 这类问题叫 Lagrange 问题; 若 $F=0$, 则叫 Mayer 问题, 若 F, K 都不为 0, 则叫 Bolza 问题, (4-75), (4-76)给出了将 Bolza 问题中的 Lagrange 部分变换成 Mayer 问题的方法。

现在我们来讨论 Mayer 问题的最大值原理公式, 若 u^*, x^* 使 $G[x(T)]$ 极大化, 且从约束(4-74'), 则我们作拉氏函数

$$L = G[x(T)] + \int_0^T \lambda(t) [\dot{x} - f(x, u, t)] dt \quad (4-77)$$

其中含影子价格 $\lambda(t)$ 的项之所以要积分, 主要是将其化成与 $G[x(T)]$ 一样, 只与终端时刻 T 有关。最优化的条件为

$$\begin{aligned}G[x(T)] + \int_0^T \lambda(t) [\dot{x} - f(x, u, t)] dt \\ \leq G[x^*(T)] + \int_0^T \lambda(t) [\dot{x}^* - f(x^*, u^*, t)] dt\end{aligned}\quad (4-78)$$

利用部分积分上式不等号左边可改写为

$$\begin{aligned}G[x(T)] + \int_0^T \lambda \dot{x} dt - \int_0^T \lambda f dt \\ = G[x(T)] + [\lambda x]_0^T - \int_0^T x \dot{\lambda} dt - \int_0^T \lambda f dt \\ = G[x(T)] + [\lambda(T)x(T)] - [\lambda(0)x_0] - \int_0^T x \dot{\lambda} dt - \int_0^T \lambda f dt\end{aligned}\quad (4-79)$$

(4-78)不等号右边也可利用分部积分改写为

$$\begin{aligned}G[x^*(T)] + \int_0^T \lambda \dot{x}^* dt - \int_0^T \lambda f^* dt \\ = G[x^*(T)] + [\lambda x^*]_0^T - \int_0^T x^* \dot{\lambda} dt - \int_0^T \lambda f^* dt \\ = G[x^*(T)] + \lambda(T)x^*(T) - \lambda(0)x_0 - \int_0^T x^* \dot{\lambda} dt - \int_0^T \lambda f^* dt\end{aligned}\quad (4-80)$$

引入汉氏函数

$$H(x, u, \lambda, t) \equiv \lambda f(x, u, t) \quad (4-81)$$

且将(4-79), (4-80)代入(4-78), 则我们有

$$\begin{aligned}G[x(T)] + \lambda(T)x(T) - \int_0^T x \dot{\lambda} dt - \int_0^T H(x, u, \lambda, t) dt \\ \leq G[x^*(T)] + \lambda(T)x^*(T) - \int_0^T x^* \dot{\lambda} dt - \int_0^T H(x^*, u^*, \lambda, t) dt\end{aligned}\quad (4-82)$$

令 $x = x^* + \Delta x$, $u = u^* + \Delta u$, 则上式可改写为

$$\begin{aligned} & G[x^*(T) + \Delta x(T)] - G[x^*(T)] + \lambda(T)\Delta x(T) \\ & - \int_0^T \Delta x \dot{\lambda} dt - \int_0^T [H(x^* + \Delta x, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t)] dt \geq 0 \end{aligned} \quad (4-83)$$

上式略加变形则为

$$\begin{aligned} & G[x^*(T) + \Delta x(T)] - G[x^*(T)] + \lambda(T)\Delta x(T) \\ & - \int_0^T \Delta x \dot{\lambda} dt - \int_0^T [H(x^* + \Delta x, u^* + \Delta u, \lambda, t) - H(x^*, u^* + \Delta u, \lambda, t)] dt \\ & - \int_0^T [H(x^*, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t)] dt \geq 0 \end{aligned}$$

上式同乘以 $\frac{\Delta x}{\Delta x}$ 或 $\frac{\Delta u}{\Delta u}$, 则有

$$\begin{aligned} & \frac{\Delta x(T)}{\Delta x(T)} \{G[x^*(T) + \Delta x(T)] - G[x^*(T)]\} + \lambda(T)\Delta x - \int_0^T \Delta x \dot{\lambda} dt \\ & - \int_0^T \frac{\Delta x}{\Delta x} [H(x^* + \Delta x, u^* + \Delta u, \lambda, t) - H(x^*, u^* + \Delta u, \lambda, t)] dt \\ & - \int_0^T \frac{\Delta u}{\Delta u} [H(x^*, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t)] dt \geq 0 \end{aligned}$$

令 $\Delta x, \Delta u \rightarrow 0$, 取极限, 则

$$\begin{aligned} & \left\{ \frac{\partial}{\partial x} G[x^*(T)] + \lambda(T) \right\} \Delta x - \int_0^T \left(\dot{\lambda} + \frac{\partial}{\partial x} H \right) \Delta x dt \\ & - \int_0^T \frac{\partial}{\partial u} H \Delta u dt \geq 0 \end{aligned} \quad (4-84)$$

如果经济系统处在最优经济效果状态, 则不管朝什么方向调动 $\Delta u, \Delta x$, 都不能使目标泛函进一步增加, 即上式在 $\Delta u, \Delta x$ 为正, 为负, 或大或小时都应成立, 要满足这一点, 则必须有下列三式成立

$$\begin{cases} -\frac{\partial H}{\partial x} = \dot{\lambda}(t) & (4-85a) \\ \frac{\partial H}{\partial u} = 0 & (4-85b) \\ \lambda(T) = -\frac{\partial}{\partial x} G[x^*(T)] & (4-85c) \end{cases}$$

回顾拉格朗日问题的极大值原理公式(3-86a), (3-86b), (3-86c), 将其与上式对照, 除了(3-86c)与(4-85c)不同外, 其它都完全相同, 不过我们从状态方程(4-74')和汉氏函数定义(4-81)不难推出

$$H = \lambda f(x, u, t), \quad \because \dot{x} = f(x, u, t)$$

$$\therefore \dot{x} = \frac{\partial H}{\partial \lambda} \quad (4-86)$$

这就与(3-86c)完全一致。

现在我们回过头来将上列程式用来解决最优更新维修计划问题。

利用目标泛函(4-73)和状态方程(4-70)作拉氏函数,

$$L = \int_0^T v(x-u)e^{-rt} dt - kx(T)e^{-rT} + \int_0^T \lambda(t) (\dot{x} + ax - u)e^{-rt} dt \quad (4-87)$$

将上式最右端一项分部积分可得

$$\begin{aligned} & \int_0^T \lambda (\dot{x} + ax - u)e^{-rt} dt \\ &= \int_0^T \lambda \dot{x} e^{-rt} dt + \int_0^T \lambda ax e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt \\ &= \lambda x e^{-rt} \Big|_0^T - \int_0^T (\dot{\lambda} e^{-rt} - r\lambda e^{-rt}) x dt \\ & \quad + \int_0^T a\lambda x e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt \\ &= \lambda x e^{-rt} \Big|_0^T + \int_0^T (a+r)\lambda - \dot{\lambda}] x e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt \end{aligned} \quad (4-88)$$

将上式代入(4-87), 则拉氏函数可写为

$$\begin{aligned} L &= \int_0^T (vx - u)e^{-rt} dt - kx(T)e^{-rT} + \lambda x e^{-rt} \Big|_0^T \\ & \quad + \int_0^T [(a+r)\lambda - \dot{\lambda}] x e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt \end{aligned} \quad (4-89)$$

将上式对 x 求导数, 且令其等于 0, 则

$$\int_0^T v e^{-rt} dt + \int_0^T [(a+r)\lambda - \dot{\lambda}] e^{-rt} dt = 0$$

$$\int_0^T [v + (a+r)\lambda - \dot{\lambda}] e^{-rt} dt = 0$$

$$v + (a+r)\lambda - \dot{\lambda} = 0 \quad (4-90)$$

根据影子价格 $\lambda(t)$ 的经济意义, 不难理解, 当 $t=T$ 时, 状态 $x(t)$ 对目标泛函的贡献正是其终端状态的残值比率, 即

$$\lambda(T) = \frac{dS}{dx} = k \quad (4-91)$$

联立(4-90), (4-91), 可解得

$$\begin{aligned}
 \lambda(t) &= e^{\int(\alpha+r)dt} \left[- \int v e^{\int(\alpha+r)dt} dt + C \right] \\
 &= e^{(\alpha+r)t} \left[-v \int e^{-(\alpha+r)dt} dt + C \right] \\
 &= e^{(\alpha+r)t} \left[\frac{v}{\alpha+r} e^{-(\alpha+r)t} + C \right] \\
 &= \frac{v}{\alpha+r} + C e^{(\alpha+r)t} \quad (4-92)
 \end{aligned}$$

将 $t=T$, $\lambda(T)=k$ 代入, 则

$$\begin{aligned}
 k &= \frac{v}{\alpha+r} + C e^{(\alpha+r)T} \\
 C &= \left(k - \frac{v}{\alpha+r} \right) e^{-(\alpha+r)T} \quad (4-93)
 \end{aligned}$$

所以通解为

$$\lambda(t) = \frac{v}{\alpha+r} + \left(k - \frac{v}{\alpha+r} \right) e^{(\alpha+r)(t-T)} \quad (4-94)$$

或写成

$$\lambda(t) = k e^{(\alpha+r)(t-T)} + \frac{v}{\alpha+r} [1 - e^{(\alpha+r)(t-T)}] \quad (4-94')$$

从(4-87)可看出 L 是 u 的线性函数, 所以我们有 bang—bang 控制法则:

$$u(t) = \begin{cases} 0, & \text{当 } \lambda(t) < 1 \\ \bar{u}, & \text{当 } \lambda(t) > 1 \\ \text{不定}, & \text{当 } \lambda(t) = 1 \end{cases} \quad (4-95)$$

其中 $\bar{u}$ 为 $u(t)$ 之上界。

一般而言, 我们有下列不等式:

$$k \leq 1 \leq \frac{v}{\alpha+r} \quad (4-96)$$

此式意味着残值总是不大于其质量决定的原值, 而收益率总是不小于折旧率加上利率。这是合理运行的必要条件。

从(4-94)不难导出

$$\lambda(0) = \frac{v}{\alpha+r} + \left(k - \frac{v}{\alpha+r} \right) e^{-(\alpha+r)T} \quad (4-97)$$

令 $\lambda(0)=1$, 则

$$\begin{aligned}
1 &= \frac{v}{\alpha+r} + \left(k - \frac{v}{\alpha+r}\right) e^{-(\alpha+r)T} \\
\alpha+r &= v + [k(\alpha+r) - v]e^{-(\alpha+r)T} \\
\frac{v - (\alpha+r)}{v - k(\alpha+r)} &= e^{-(\alpha+r)T} \\
\ln \frac{v - k(\alpha+r)}{v - (\alpha+r)} &= (\alpha+r)T \quad (4-98)
\end{aligned}$$

所以, 当 $T(\alpha+r) \geq \ln \frac{v - k(\alpha+r)}{v - (\alpha+r)}$ 时, $\lambda(0) \geq 1$

$$\text{当 } T(\alpha+r) \leq \ln \frac{v - k(\alpha+r)}{v - (\alpha+r)} \text{ 时, } \lambda(0) \leq 1 \quad (4-99)$$

根据(4-91), $\lambda(T) = k$, 又根据(4-96), $k \leq 1$,

$$\therefore \lambda(T) \leq 1 \quad (4-100)$$

$\therefore$ 当 $\lambda(0) \leq 1$ 时, 则在 $0 \sim T$ 时域, $\lambda(t)$ 都小于 1, 最优控制 $u(t) = 0, 0 \leq t \leq T$.

若 $\lambda(0) \geq 1$, 则在 $0 \sim T$ 时域, $\lambda(t)$ 从大于 1 逐渐减少到 1, 然后小于 1, 其中有一时刻 $t = \theta$, 使得 $\lambda(\theta) = 1$.

$$\text{当 } 0 \leq t \leq \theta \text{ 时, } u(t) = \bar{u} \quad (4-101)$$

$$\text{当 } \theta \leq t \leq T \text{ 时, } u(t) = 0$$

令(4-94)中的 $t = \theta$, $\lambda(t) = 1$, 则有

$$1 = \frac{v}{\alpha+r} + \left(k - \frac{v}{\alpha+r}\right) e^{(\alpha+r)(\theta-T)},$$

从中可解出

$$\theta = \frac{1}{\alpha+r} \ln \frac{v}{v - k(\alpha+r)} + T \quad (4-102)$$

§4 最优销售计划

我们在广告上花一些钱, 可以使消费者了解商品, 因而增加对商品的购买, 消费者对商品的了解是企业的资本, 大多数文献将这种“资本”称为“商誉”。这种资本与广告支出有一定的函数关系, 我们可以将广告支出视为“投资”, 而将商誉视为固定资产。若没有广告投资, 则商誉这种固定资产会逐渐“折旧”, 即商誉会逐渐下降。

设 t 时刻的广告支出为 $u(t)$,

t 时刻的累积商誉为 $x(t)$,

则系统的状态方程为

$$\dot{x} = u(t) - \alpha x(t) \quad (4-103)$$

$$x(0) = x_0$$

其中 α 为没有广告支出时商誉的下跌比率。

这个方程最早是由 Nerlove 和 Arrow 在其 1962 年文献中提出，而 Telser 于其 1962 年文献中提出了一些用经济计量学估计状态方程参数的方法及商誉的度量方法。

单位时间内企业的毛利是当时商誉，当时广告费用和价格的函数，设需求函数为

$$D = Ax^\alpha p^\beta$$

其中 A, α, β 为用经济计量学算出的参数， p 为价格，总的毛利为

$$R = (p - c)D \quad (4-104)$$

其中 c 为单位产品的生产费用，从 R 中除掉广告费用后的利润为

$$U = R - u \quad (4-105)$$

目标泛函为

$$\max_x: I = \int_0^T (R - u)e^{-rt} dt + x(T)e^{-rT} \quad (4-106)$$

即要求 $0 \sim T$ 时域内的总利润和最终时刻 T 的商誉之现值极大化。

与上节类似，我们作拉氏函数

$$L = \int_0^T (R - u)e^{-rt} dt + x(T)e^{-rT} - \int_0^T \lambda(t)(\dot{x} + \alpha x - u)e^{-rt} dt \quad (4-107)$$

将上式最右端一项展开分部积分得

$$\begin{aligned} & \int_0^T \lambda(t)(\dot{x} + \alpha x - u)e^{-rt} dt \\ &= \int_0^T \lambda \dot{x} e^{-rt} dt + \int_0^T \alpha \lambda x e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt \\ &= \lambda x e^{-rt} \Big|_0^T - \int_0^T (\dot{\lambda} e^{-rt} - r \lambda e^{-rt}) x dt + \int_0^T \alpha \lambda x e^{-rt} dt \\ & \quad - \int_0^T \lambda u e^{-rt} dt \\ &= \lambda x e^{-rt} \Big|_0^T - \int_0^T [\dot{\lambda} - (\alpha + r)\lambda] x e^{-rt} dt - \int_0^T \lambda u e^{-rt} dt. \end{aligned} \quad (4-108)$$

将上式代入 (4-107)，则

$$\begin{aligned} L &= \int_0^T (k - u)e^{-rt} dt + x(T)e^{-rT} - \lambda x e^{-rt} \Big|_0^T \\ & \quad + \int_0^T [\dot{\lambda} - (\alpha + r)\lambda] x e^{-rt} dt + \int_0^T \lambda u e^{-rt} dt, \end{aligned} \quad (4-109)$$

将上式对 x 求导数，且令其等于 0，则我们有

$$\frac{\partial L}{\partial x} = \int_0^T R'(x)e^{-rt}dt + \int_0^T [\dot{\lambda} - (\alpha + r)\lambda]e^{-rt}dt = 0,$$

即

$$R'(x) + \dot{\lambda} - (\alpha + r)\lambda = 0,$$

$$R'(x) = (\alpha + r)\lambda - \dot{\lambda} \quad (4-110a)$$

必要条件还有

$$\dot{x} = u - \alpha x,$$

$$x(0) = x_0, \quad (4-110b)$$

$$\lambda(T) = 0,$$

将(4-109)对 u 求导数, 且令其等于 0, 则

$$\frac{\partial L}{\partial u} = \int_0^T [R'(u) - 1]e^{-rt}dt + \int_0^T \lambda e^{-rt}dt = 0,$$

即

$$R'(u) + \lambda = 1, \quad (4-110c)$$

将(4-109)对 p 求导数, 且令其等于 0, 则

$$\frac{\partial L}{\partial p} = \int_0^T R'(p)e^{-rt}dt = 0$$

即

$$R'(p) = 0 \quad (4-110d)$$

由于(4-110c)中的 λ 恒为正, λ 为商誉的影子价格, 商誉的变化主要得益于广告投资, 所以 λ 为总的目标泛函对广告投资的边际价值, 由于拉氏函数中含影子价格的项前面为负号, 所以 λ 实际上是广告的边际代价(边际费用), 而 $R'(u)$ 为广告投资的瞬间边际收益, (4-110c)说明在最优状态, 广告投资的瞬间边际收益小于广告投资的总边际费用。Palda 在其1964年文献中报道实际经验研究证实了上述结论。

根据(4-104)

$$R = (p - c)D,$$

代入(4-110c), 则

$$(p - c)D'(u) + \lambda = 1. \quad (4-111)$$

从(4-110d)可导出

$$R'(p) = \frac{\partial}{\partial p} [(p - c)D] = 0$$

即

$$D + (p - c)D'(p) = 0 \quad (4-112)$$

从(4-111)和(4-112)可导出

$$\frac{\frac{\partial D}{\partial u}}{\frac{\partial D}{\partial p}} = \frac{\lambda - 1}{D}$$

即

$$\frac{\frac{\partial D}{\partial u} \frac{u}{D}}{\frac{\partial D}{\partial p} \frac{p}{D}} = \frac{(\lambda - 1)u}{pD}, \quad \frac{\frac{\partial D}{\partial u} \frac{u}{D}}{(\lambda - 1) \frac{\partial D}{\partial p} \frac{p}{D}} = \frac{n}{R} \quad (4-113)$$

也就是说广告支出占销售毛利的比例正比于需求量对广告投资的弹性，反比于需求量对价格的弹性，此式由 Jacquemin 于其1972年的两篇文献中推出，Jacquemin 是在 Dorfman, Steiner 之1957年文献和 Nerlove, Arrow 1962年文献的基础上发展的。

从 (4-111), (4-112) 还可推出

$$\frac{\partial D / \partial u}{\partial D / \partial p} = - \frac{1 - \lambda}{D} = - \frac{\partial R / \partial u}{D}$$

即

$$\frac{\frac{\partial D}{\partial u} \frac{u}{D}}{- \frac{\partial D}{\partial p} \frac{p}{D}} = \frac{\frac{\partial R}{\partial u} u}{Dp} = \frac{\partial R}{\partial u} \frac{u}{R} \quad (4-114)$$

也就是说需求的广告弹性与价格弹性之比率等于利润对广告支出的弹性时，可取得最优经济效果，此式由本书作者导出，权称之为**最优销售律**。

Jacquemin 还从 (4-112) 推出

$$\frac{x}{pD} = \frac{\partial D / \partial x (x/D)}{- \frac{\partial D}{\partial p} \frac{p}{D} [(\alpha + r)\lambda - \dot{\lambda}]} \quad (4-115)$$

若价格取不变常数，则最优化的条件为

$$\frac{\partial R}{\partial x} - (\alpha + r)\lambda + \frac{d\lambda}{dt} = 0 \quad (4-116a)$$

$$\frac{\partial R}{\partial u} + \lambda(t) = 1 \quad (4-116b)$$

$$\dot{x} = u - \alpha x \quad (4-116c)$$

$$x(0) = x_0, \quad \lambda(T) = 0 \quad (4-116d)$$

由于价格取常数，所以需求量只受当前广告支出和商誉影响，而商誉是过去广告支出的累积后果，设需求函数为

$$D = B_1 x^\alpha + B_2 u^\beta \quad 0 < \alpha < 1, \quad 0 < \beta < 1 \quad (4-117)$$

而

$$R = (p - c)D - u$$

$$= (p-c)(B_1x^\alpha + B_2u^\beta) - u \quad (4-18)$$

$$\frac{\partial R}{\partial x} = \alpha(p-c)B_1x^{\alpha-1} \quad (4-119a)$$

$$\frac{\partial R}{\partial u} = \beta(p-c)B_2u^{\beta-1} - 1 \quad (4-119b)$$

将上两式分别代入(4-116a), (4-116b), 则最优化条件可改写为

$$\alpha(p-c)B_1x^{\alpha-1} - (\alpha+r)\dot{\lambda} = 0,$$

$$\dot{\lambda} = (\alpha+r)\lambda - \alpha(p-c)B_1x^{\alpha-1} \quad (4-120a)$$

$$\beta(p-c)B_2u^{\beta-1} - 1 + \lambda = 1,$$

即

$$u = \left[\frac{2-\lambda}{\beta(p-c)B_2} \right]^{\frac{1}{\beta-1}}, \quad (4-120b)$$

$$\lambda = -\beta(p-c)B_2u^{\beta-1} + 2. \quad (4-120c)$$

将(4-120b)代入(4-116c), 可得

$$\dot{x} = \left[\frac{2-\lambda}{\beta(p-c)B_2} \right]^{\frac{1}{\beta-1}} - \alpha x \quad (4-120d)$$

$$x(0) = x_0, \quad \lambda(T) = 0.$$

联立(4-120a)(4-120d), 则得到一阶微分方程组

$$\dot{\lambda} = (\alpha+r)\lambda - \alpha(p-c)B_1x^{\alpha-1} \quad (4-120a)$$

$$\dot{x} = \left[\frac{2-\lambda}{\beta(p-c)B_2} \right]^{\frac{1}{\beta-1}} - \alpha x \quad (4-120d)$$

初始条件为

$$x(0) = x_0, \quad \lambda(0) = -\beta(p-c)B_2[u(0)]^{\beta-1} + 2 \quad (4-120e)$$

由于 $\lambda(0)$ 由控制 $u(0)$ 决定, 所以 $\lambda(0)$ 为决策变量。当

$$\dot{\lambda} = 0, \quad \dot{x} = 0$$

时, 我们有

$$(\alpha+r)\lambda = \alpha(p-c)B_1x^{\alpha-1}$$

即

$$x = \left[\frac{(\alpha+r)\lambda}{\alpha(p-c)B_1} \right]^{\frac{1}{\alpha-1}}, \quad (4-121)$$

及

$$x = \frac{1}{a} \left[\frac{2-\lambda}{\beta(p-c)B_2} \right]^{\frac{1}{a-1}} \quad (4-122)$$

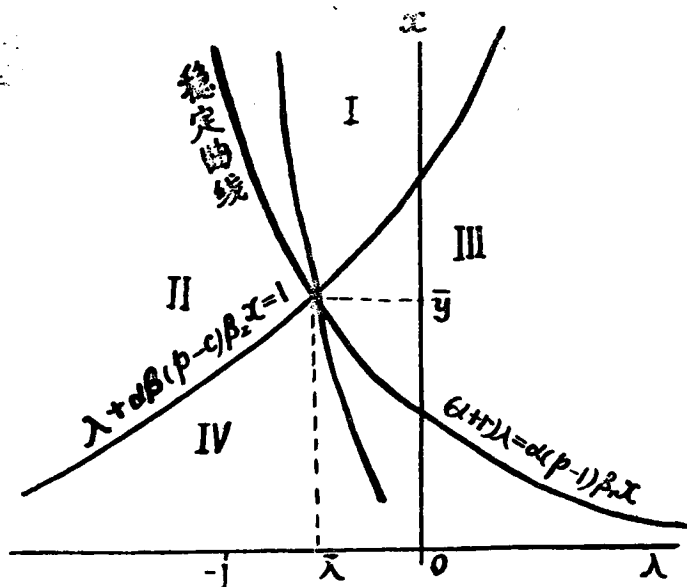


图 4-4

$$\therefore 0 < \alpha < 1, 0 < \beta < 1,$$

$$\therefore (4-121) \text{ 中的 } \frac{dx}{d\lambda} < 0,$$

$$(4-122) \text{ 中的 } \frac{dx}{d\lambda} > 0.$$

我们在 $\lambda-x$ 相平面作出这两条曲线，它们将相平面分为 4 个区域 i, ii, iii, iv, 如图 4-4 所示，我们用 x_1, λ_1 表示与 (4-121) [即与 (4-120a)] 的曲线比较之点坐标，用 x_2, λ_2 表示与 (4-122) [即与 (4-120d)] 的曲线比较之点坐标，则我们对相平面 4 个区域可作如下分析：

$$(i) \quad \because x_1 < \left[\frac{(a+r)\lambda_1}{\alpha(p-c)B_1} \right]^{\frac{1}{a-1}} \quad \therefore \dot{\lambda} > 0, \lambda \text{ 上升},$$

$$\because x_2 > \frac{1}{a} \left[\frac{a-\lambda_2}{\beta(p-c)_2 B_2} \right]^{\frac{1}{a-1}} \quad \therefore \dot{x} < 0, x \text{ 下降},$$

$$(ii) \quad \text{令 } f_1(\lambda_1) = \left[\frac{(a+r)\lambda_1}{\alpha(p-c)B_1} \right]^{\frac{1}{a-1}}$$

$$f_2(\lambda_2) = \frac{1}{\alpha} \left[\frac{\alpha - \lambda_2}{\beta(p-c)B_2} \right]^{\frac{1}{s-1}},$$

则分析可简化为

$$(i) \quad \therefore x_1 > f_1(\lambda_1) \quad \therefore \dot{\lambda} < 0, \lambda \text{ 下降},$$

$$\therefore x_2 > f_2(\lambda_2) \quad \therefore \dot{x} < 0, x \text{ 下降},$$

$$(iii) \quad \therefore x_1 < f_1(\lambda_1) \quad \therefore \dot{\lambda} > 0, \lambda \text{ 上升}$$

$$\therefore x_2 < f_2(\lambda_2) \quad \therefore \dot{x} > 0, x \text{ 上升},$$

$$(iv) \quad \therefore x_1 < f_1(\lambda_1) \quad \therefore \dot{\lambda} < 0, \lambda \text{ 下降}$$

$$\therefore x_2 < f_2(\lambda_2) \quad \therefore \dot{x} > 0, x \text{ 上升}.$$

引入

$$z = x - \bar{x}, \quad q = \lambda - \bar{\lambda} \quad (4-123)$$

其中 $\bar{x}$, $\bar{\lambda}$ 为 x, λ 之稳定值, 将上式代入微分方程组 (4-120a), (4-120d), 我们可得到

$$\begin{cases} \dot{q} = (\alpha + r)(q + \bar{\lambda}) - \alpha(p-c)B_1(z + \bar{x})^{\alpha-1} & (4-124a) \\ \dot{z} = \left[\frac{2 - (q + \bar{\lambda})}{\beta(p-c)B_2} \right]^{\frac{1}{s-1}} - \alpha(z + \bar{x}) & (4-124b) \end{cases}$$

当 $q=0, z=0$ 时, 系统处在均衡状态, 即 $\lambda = \bar{\lambda}, x = \bar{x}$, 且系统稳定下来, 因为此时

$$\dot{q} = 0, \quad \dot{z} = 0 \quad (4-125)$$

将上式代入 (4-124), 且令 $q=0, z=0$, 则

$$\begin{cases} 0 = (\alpha + r)\bar{\lambda} - (p-c)B_1\bar{x}^{\alpha-1} \\ 0 = \left[\frac{2 - \bar{\lambda}}{\beta(p-c)B_2} \right]^{\frac{1}{s-1}} - \alpha\bar{x} \end{cases}$$

从上式解得

$$\bar{\lambda} = \frac{\alpha(p-c)B_1\bar{x}^{\alpha-1}}{\alpha + r} \quad (4-126a)$$

$$\bar{x} = \frac{1}{\alpha} \left[\frac{2 - \bar{\lambda}}{\beta(p-c)B_2} \right]^{\frac{1}{s-1}} \quad (4-126b)$$

即

$$\bar{x} = \frac{1}{\alpha} \left[\frac{2 - \alpha(p-c)B_1 \bar{x}^{\alpha-1} / (\alpha+r)}{B(p-c)B_2} \right]^{\frac{1}{\alpha-1}} \quad (4-126c)$$

从中可解出 $\bar{x}$ ，代回(4-126a)，则可解出 $\bar{\lambda}$ 。

我们可以适当地选择 $u(0)$ ，以使 $\lambda(0)$ 尽量接近 $\bar{\lambda}$ 。根据(4-120b)

$$u(0) = \left[\frac{2 - \lambda(0)}{\beta(p-c)B_2} \right]^{\frac{1}{\alpha-1}} \quad (4-127)$$

令 $\lambda(0) \rightarrow \bar{\lambda}$ ，则

$$u(0) \rightarrow \left[\frac{2 - \bar{\lambda}}{\beta(p-c)B_2} \right]^{\frac{1}{\alpha-1}} \quad (4-128)$$

其中的 $\bar{\lambda}$ 可从方程组(4-126)用数值解法求得。

现在我们将动态模型与 Dorfman, Sleiner 1957 年文献中的静态模型比较一下。

静态模型的目标函数为

$$I = (p-c)D(u, p) - u \quad (4-129)$$

最优化的必要条件为

$$\begin{cases} D + (p-c) \frac{\partial D}{\partial p} = 0 \end{cases} \quad (4-130a)$$

$$\begin{cases} (p-c) \frac{\partial D}{\partial u} - 1 = 0 \end{cases} \quad (4-130b)$$

根据上两式不难导出

$$\frac{u}{pD} = \frac{\frac{\partial D}{\partial u} \frac{u}{D}}{-\frac{\partial D}{\partial p} \frac{p}{D}} \quad (4-131)$$

即需求对广告支出的弹性与需求的价格弹性之比等于广告支出占销售额的比重。将(4-131)与(4-113)比较，可以看出，除了(4-113)多一个 $(1-\lambda)$ 外，两式完全一样。也就是说静态情况下需求的广告弹性为

$$\frac{\partial D}{\partial u} \frac{u}{D}$$

而动态模型中需求的广告弹性要加上一个因子 $\frac{1}{1-\lambda}$ ，因而变成

$$\frac{\partial D}{\partial u} \frac{u}{D} \frac{1}{1-\lambda} \quad (4-132)$$

此式叫做长期广告的需求弹性。

§ 5 最优销售、研究计划和最优产品寿命

此节的内容主要是 Näslund 和 Sellstedt 1972 年文献中的研究成果。

产品的销售计划不但与价格和广告投资有关，而且与新产品的研制有关，若产品换代较慢，则产品的销售寿命较长，则支出较多费用做广告以促进对商品的需求是合理的；另一方面，若产品销售寿命较长，则 RD（研究与发展）投资就不必太高，因为开发新产品对企业来说并不十分重要。

较高的 RD 投资有两方面的作用，一方面这将缩短原有产品之销售寿命，在这种情况下广告投资不能太多，反之，若在广告方面大量投资，则产品之最优寿命会加长，而产品寿命又是广告投资决策必须考虑的一个重要因素。另一方面，较高的 RD 投资将使新产品的效用较高，而效用越高，销售寿命越长，由于存在这些交互影响的关系，厂商必须统筹兼顾，制定出通盘的推销和研制计划，以同时确定产品的最优产品寿命，最优广告水平和 RD 投资。为此，必须首先了解产品的生命周期规律。

一种产品的销售量随时间变化具有共同的周期规律，新产品面世时，叫萌发期，随后销售量增长得越来越快，是谓增长期，此后增长速度越来越慢，逐渐接近销售量的最大极限，是谓成熟期；当销售量达到最大极限且开始下降时，则进入饱和期；当销售量迅速下降时，则是衰老期，也有文献将生命周期按累积销售量分为萌发期，增长期、成熟期，衰老期，其图解参见图4—7。

根据 Pessemier 1966年文献报道，概率论中的 Weibull 分布密度函数与实际销售量时间序列拟合得很好，他根据企业的报表将很多企业的销售量数据做出经济计量模型，其中有20种商品的销量时间序列与 Weibull 分布密度函数拟合得很好，根据零售商的报表，不按企业分类，而按商品种数分类，也有20多种商品销量数据与这种模型拟合得很好。不过数据必须按产品形式和商标分类，而不能按商品用途分类。Weibull 分布密度函数为

$$f(t) = \frac{\beta}{\alpha} t^{\beta-1} e^{-\frac{t^\beta}{\alpha}} \quad (4-133a)$$

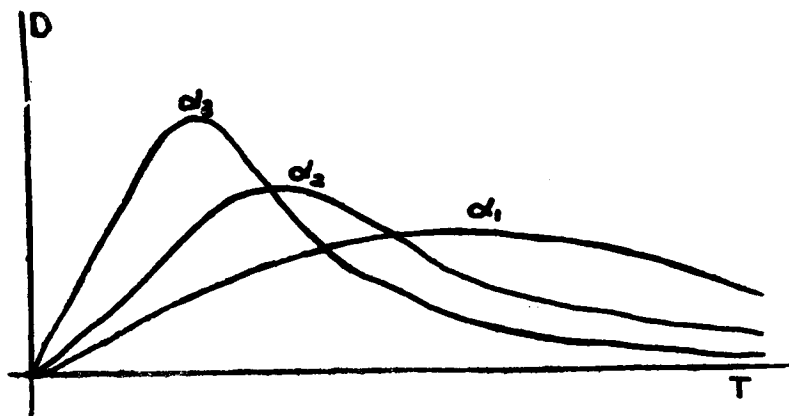


图 4—5

其中参数 $\alpha > 0$, $\beta > 0$ 。 α 为尺度参数, 它反映了特定密度曲线之陡峭或平缓程度, α 越小则密度曲线越平缓, α 越大则密度曲线越陡峭。 β 为形状参数, $\beta = 1$ 时, 曲线为纯粹指数型的, β 越大, 则峰状越显著, 我们令 f 为销售量, t 为时间, 则 (4—133) 为销售周期模型, 图4—5是 α 为不同值时的销售周期模型, 其中 $\alpha_3 > \alpha_2 > \alpha_1$ 。图 4—7 是 β 为不同值时的销售周期模型, 其中 $\beta_3 > \beta_2 > \beta_1$, 将 $f(t)$ 对 t 积分, 则得到 Weibull 分布函数, 它

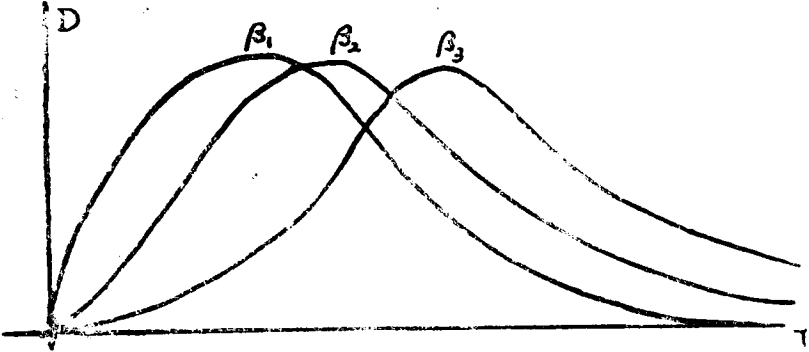


图 4—6

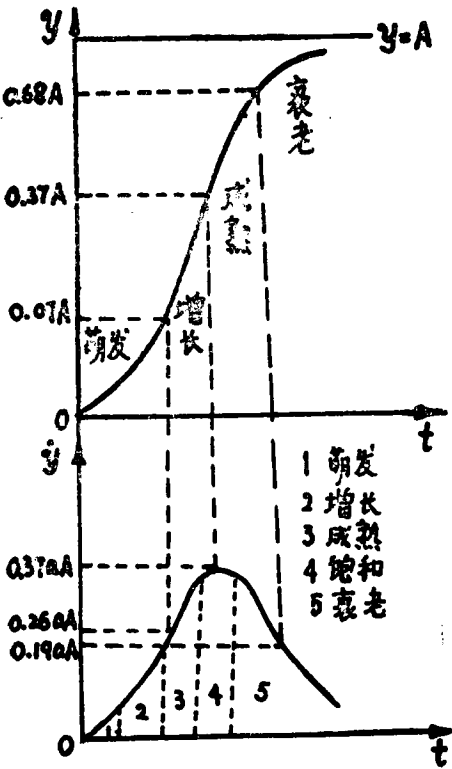


图 4—7

就是累积销售量模型, 设 $F(t)$ 为累积销售量, 则

$$F(t) = \int_0^t f(t) dt = 1 - e^{-\frac{t^\alpha}{\alpha}} \quad (4-133b)$$

适于描述耐用商品销售周期的另一种函数是著名的 Gompertz 曲线, 利用这种曲线描述的销售量模型为

$$f(t) = ay(b - \ln y) \quad (4-134a)$$

将其对时间积分则得到耐用商品拥有量模型

$$F(t) = e^{b - e^{-at}} \quad (4-134b)$$

其图形见图4—7。饱和拥有量为 $A = e^b$, 最大销售量为 $0.37 aA$, 左拐点销量为 $0.26 aA$, 右拐点销量为 $0.19 aA$ 。这类模型经过适当变换, 可以变成线性形式, 所以将数据适当处理后, 可以进行一元线性回归, 用经济计量学算出参数 a, b 。为使模型更具适应性, 我们还可将 (4—134) 改写成

$$f(t) = ace^{b - a^t - ae^{-at}} \quad (4-135a)$$

$$F(t) = e^{b - ae^{-at}} \quad (4-135b)$$

这个模型更便于进行一元线性回归。

我们可以视价格为常数，因而与广告有关的销售利润正比于 (4-135a) 或 (4-133a)，设比例系数为 π ，又设销售额为 $x = f(t)$ 于是利润可写为

$$\pi x = \pi f(t) = \pi \frac{\beta}{\alpha} t^{\beta-1} e^{-\frac{t^\beta}{\alpha}} \quad (4-136a)$$

或

$$\pi x = \pi f(t) = \pi a c e^{b-at} - c e^{-at} \quad (4-136b)$$

我们沿用上节的分析，设 x 为累积商誉，由于商誉正比于销售量，只要适当地选择度量单位，我们可以将其用销售量表示。

设生命周期为 T ，它是决策变量，每隔一时间长度 T ，有一新产品投产，又设试制新产品的费用为 $K(T)$ ，它由研制费用，一次性投资及创办费用等构成。

设有 n 种新产品，共在 n 个阶段研制成，每个阶段的周期长度为 T ，每个阶段的 RD 投资为 K ，将其全部贴现后则为

$$\begin{aligned} & K e^{-rT} + K e^{-2rT} + \dots + K e^{-(n-1)rT} \\ & = K [e^{-rT} + e^{-2rT} + \dots + e^{-(n-1)rT}] \end{aligned} \quad (4-137)$$

令

$$S = e^{-rT} + e^{-2rT} + \dots + e^{-(n-1)rT}$$

则

$$e^{-rT} S = e^{-2rT} + e^{-3rT} + \dots + e^{-nrT}$$

$\therefore$

$$(1 - e^{-rT}) S = e^{-rT} - e^{-nrT}$$

$$S = \frac{e^{-rT} - e^{-nrT}}{1 - e^{-rT}} \quad (4-138)$$

如果我们考虑的是长远计划，则可令 $n \rightarrow \infty$ ，对 S 取极限

$$\lim_{n \rightarrow \infty} S = \frac{e^{-rT}}{1 - e^{-rT}} \quad (4-139)$$

代回 (4-137)，则无穷时域内，RD 总费用之现值为

$$K \sum_{n=1}^{\infty} e^{-nrT} = K \frac{e^{-rT}}{1 - e^{-rT}} \quad (4-140)$$

设每一时刻不做广告的利润为 R_1 ，做广告增加的利润为 πx ，广告支出为 u ，于是长期目标泛函为

$$\begin{aligned} \max: I &= \int_0^{\infty} (R_1 + \pi x - u) e^{-rt} dt - K \sum_{n=1}^{\infty} e^{-nrT}, \\ &= \sum_{n=1}^{\infty} \int_0^T (R_1 + \pi x - u) e^{-nrT} dt - K \frac{e^{-rT}}{1 - e^{-rT}} \\ &= \frac{1}{1 - e^{-rT}} \int_0^T (R_1 + \pi x - u) e^{-rt} dt - K \frac{e^{-rT}}{1 - e^{-rT}}. \end{aligned} \quad (4-141)$$

根据 Vidale 和 Wolfe 于其 1957 年文献中提出的一个模型，销售额与广告支出之状态方

$$\dot{x} = \pi u - \frac{\pi u x}{m} - \alpha x \quad (4-142)$$

$$x(0) = x_0$$

π 可解释为广告的经济效果系数, 其中的 u 满足不等式约束

$$0 \leq u \leq \bar{u}. \quad (4-143)$$

m 为商誉即销售量的上限, 若 m 极大, 则 (4-142) 变成

$$\dot{x} = \pi u - \alpha x \quad (4-144)$$

$$x(0) = x_0$$

作拉氏函数

$$L = \frac{1}{1 - e^{-rT}} \left[\int_0^T (R_1 + \pi x - u) e^{-rt} dt - K e^{-rT} \right] - \int_0^T \lambda (\dot{x} + \alpha x - \pi u) dt, \quad (4-145)$$

分部积分, 我们得到

$$L = \frac{1}{1 - e^{-rT}} \left[\int_0^T (R_1 + \pi x - u) e^{-rt} dt - K e^{-rT} \right] + \int_0^T x \dot{\lambda} dt - \int_0^T \alpha \lambda x dt + \int_0^T \pi \lambda u dt - \lambda(T)x(T) + \lambda(0)x_0. \quad (4-146)$$

对 x 求导数且令其等于 0, 则

$$\int_0^T \frac{\pi e^{-rt}}{1 - e^{-rT}} dt + \int_0^T \dot{\lambda} dt - \int_0^T \alpha \lambda dt = 0, \\ \pi \frac{e^{-rt}}{1 - e^{-rT}} + \dot{\lambda} = \alpha \lambda, \quad (4-147)$$

加上边界条件

$$\lambda(T) = 0 \quad (4-148)$$

从 (4-147), (4-148) 解出,

$$\lambda = \frac{\pi [e^{-(r+\alpha)T+\alpha t} - e^{rt}]}{(1 - e^{-rT})(\alpha + r)} \quad (4-149)$$

令 $\lambda = 1$, 则上式变成

$$\frac{(1 - e^{-rT})(\alpha + r)}{\pi} - e^{-(r+\alpha)T} e^{\alpha t} + e^{-rt} = 0 \quad (4-150)$$

用牛顿近似法或用计算机编个小程序进行试凑计算, 可解出满足上式的 t^* 。

由于 (4-145) 中的 L 是 u 之线性函数, 则

$$u(t) = \begin{cases} \bar{u} & t < t^*, \text{ 即 } \lambda > 1 \\ 0 & t > t^*, \text{ 即 } \lambda < 1 \\ \text{不定 } t = t^*, \text{ 即 } \lambda = 1 \end{cases} \quad (4-151)$$

从 (4-144), 我们可解出:

$$\begin{aligned} x &= e^{-\alpha t} \left[\int \pi u e^{\alpha t} dt + C \right] \\ x &= e^{-\alpha t} \left[\frac{\pi u}{\alpha} e^{\alpha t} + C \right] \end{aligned}$$

将 $t=0, x(0)=x_0$ 代入, 解出 C :

$$C = x_0 - \frac{\pi u}{\alpha}$$

于是通解为

$$x = \frac{\pi u}{\alpha} + \left(x_0 - \frac{\pi u}{\alpha} \right) e^{-\alpha t} \quad (4-152)$$

再将 (4-141) 中的 I 对 T 求导数, 且令其等于 0, 则

$$\begin{aligned} \frac{dI}{dT} &= - \frac{r e^{-rT}}{(1 - e^{-rT})^2} \int_0^T (R_1 + \pi x - u) e^{-rt} dt \\ &\quad + \frac{[R_1(T) + \pi x(T) - u(T)] e^{-rT}}{1 - e^{-rT}} - K'(T) \frac{e^{-rT}}{1 - e^{-rT}} \\ &\quad - K(T) \left[\frac{-r e^{-2rT}}{(1 - e^{-rT})^2} + \frac{-r e^{-rT}}{1 - e^{-rT}} \right] = 0 \\ \frac{-r}{1 - e^{-rT}} \int_0^T (R_1 + \pi x - u) e^{-rt} dt + R_1(T) + \pi x(T) - u(T) \\ &\quad - K'(T) + K(T) \left(\frac{r e^{-rT}}{1 - e^{-rT}} + r \right) = 0 \\ R_1(T) &= rI(T) + K'(T) - rK(T) - \pi x(T) + u(T) \end{aligned} \quad (4-153)$$

当 $\lambda = 1$ 时, $u = \bar{u}$, 根据 (4-152), 这时

$$x = \frac{\pi \bar{u}}{\alpha} + \left(x_0 - \frac{\pi \bar{u}}{\alpha} \right) e^{-\alpha T} \quad (4-154)$$

代入 (4-153), 则最优化的条件为

$$\begin{aligned} R_1(T) &= rI(T) + K'(T) - rK(T) - \pi \left[\frac{\pi \bar{u}}{\alpha} + \left(x_0 - \frac{\pi \bar{u}}{\alpha} \right) e^{-\alpha T} \right] \\ &\quad + u(T) \end{aligned} \quad (4-155)$$

(4-153), (4-155) 中的 $R_1(T) + \pi x(T)$ 为总利润, $rI(T)$ 为总利息, $K'(T)$ 为 RD 投资之边际费用, 一般 $K(T)$ 是 K 的递减函数, 因此 $K'(T) < 0$, $rK(T)$ 为用于 RD 的利息。

Toffler 于其1970年文献中提出, 通讯和运输效率提高后, Weibull 分布的参数 α 将增加, 这意味着生命周期的曲线左移。这将使我们的最优 RD 投资增加而广告最优支出减少。

Mickwitz 于其1959年文献中以及 Lambin 于其1970年文献中研究了产品销售的不同阶段, 厂商采用各种销售手段的经济效果。Mickwitz 在其1955年文献中指出, 在萌发阶段, 产品质量的 RD 是最重要的推销手段, 在增长阶段这种手段仍十分有效。在增长和成熟阶段, 因为生产同类商品的厂商增多, 因此企业面临的主要问题是使人们购买本企业商标的产品而不购买竞争商标之产品, 依 Mickwitz 之见, 这一阶段, 价格和广告是最重要的销售手段, 在饱和阶段, 质量更上一层楼是最重要的销售手段, 在衰老阶段, 寻求新的消费者的广告是重要的, 但其收效也不会十分显著, 这些见解与我们的分析相结合, 则广告支出最好进行到 t^* 时刻, 一般最优寿命 $T > t^*$, 因此 $u(T) = 0$, 所以最优化的条件 (4-153) 中的 $u(T)$ 将消失。

现在我们讨论状态方程取 Vidale 和 Wolfe 提出的 (4-142) 形式时的最优化问题。这时状态方程为

$$\begin{aligned} \dot{x} &= \pi u \frac{\pi u x}{m} - \alpha x \\ x(0) &= x_0 \quad x \leq m, \quad 0 \leq u \leq \bar{u} \end{aligned} \quad (4-153)$$

设目标泛函为

$$\text{Max}; I = \int_0^{\infty} (\alpha x - u) e^{-rt} dt \quad (4-154)$$

作拉氏函数

$$\begin{aligned} L &= \int_0^{+\infty} (\alpha x - u) e^{-rt} dt \\ &\quad - \int_0^{+\infty} \lambda_1 \left[\dot{x} + \alpha x - \pi u + \frac{\pi u x}{m} \right] e^{-rt} dt - \int_0^{+\infty} \lambda_2 (x - m) e^{-rt} dt \end{aligned} \quad (4-155)$$

分部积分得:

$$\begin{aligned} L &= \int_0^{+\infty} (\alpha x - u) e^{-rt} dt - \int_0^{+\infty} x \left(-\dot{\lambda}_1 + \alpha \lambda_1 + \frac{\pi u}{m} \lambda_1 - r \lambda_1 \right) e^{-rt} dt \\ &\quad + \int_0^{+\infty} \lambda_1 \pi u e^{-rt} dt - \int_0^{+\infty} \lambda_2 (x - m) e^{-rt} dt - \lambda_1 x e^{-rt} \Big|_0^T \end{aligned} \quad (4-156)$$

将 L 对 x 求导数, 令其等于0, 则

$$\begin{aligned} &\int_0^{+\infty} \alpha e^{-rt} dt - \int_0^{+\infty} \left(-\dot{\lambda}_1 + \alpha \lambda_1 + \frac{\pi u}{m} \lambda_1 - r \lambda_1 \right) e^{-rt} dt \\ &\quad - \int_0^{+\infty} \lambda_2 e^{-rt} dt = 0 \\ &\alpha + \dot{\lambda}_1 - (\alpha + r) \lambda_1 - \frac{\pi u}{m} \lambda_1 - \lambda_2 = 0 \end{aligned}$$

即

$$\dot{\lambda} - \left(\alpha + r + \frac{\pi u}{m} \right) \lambda_1 - \lambda_2 + a = 0. \quad (4-157)$$

当系统处于稳定状态时, $\dot{x} = 0$, $\dot{\lambda} = 0$, 则从 (4-153) 和 (4-157) 分别可解得

$$\alpha x = \pi u \left(1 - \frac{x}{m} \right) \quad (4-158a)$$

$$\left(\alpha + r + \frac{\pi u}{m} \right) \lambda_1 + \lambda_2 = a \quad (4-158b)$$

若 m 很大, 则 $x < m$, $\lambda_2 = 0$, 从 (4-158) 可解出

$$\pi u = \frac{\alpha x}{1 - \frac{x}{m}}$$

代入 (4-158b)

$$\left[\alpha + r + \frac{\alpha x}{m-x} \right] \lambda_1 = a. \quad (4-159)$$

由于解为 bang-bang 控制, 所以

$$u = \begin{cases} 0, & \text{当 } \lambda_1 < \frac{\alpha(m-x)}{(\alpha+r)m-rx} \\ \bar{u}, & \text{当 } \lambda_1 > \frac{\alpha(m-x)}{(\alpha+r)m-rx} \\ \text{不定}, & \text{当 } \lambda_1 = \frac{\alpha(m-x)}{(\alpha+r)m-rx} \end{cases} \quad (4-160)$$

令 (4-159) 中的 $\lambda_1 = 1$, 则可解出

$$x = \frac{m(a - \alpha - r)}{a - r} \quad (4-161)$$

则它和 $y_1, \dots, y_m$ 仍是间接相关的, 即使只与和 $y_1, \dots, y_m$ 间接相关的状态变量之间存在耦合, 则仍与 $y_1, \dots, y_m$ 间接相关。

(iii) 若在化为最简形式后, $x_1, \dots, x_n$ 中的某些变量, 和 $y_1, \dots, y_m$ 中的任一个, 既无直接相关, 又无间接相关, 则这些状态变量与 $y_1, \dots, y_m$ 不相关。

显然, 与 $y_1, \dots, y_m$ 直接或间接相关的状态变量才是能观测的, 而和 $y_1, \dots, y_m$ 不相关的状态变量是不能观测的, 不能观测意味着不能通过 y 识别 x 。

据此我们有能观性的定义:

若下列线性系统

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \quad (5-7)$$

对于 $t_1 > t_0$, 且根据 $[t_0, t_1]$ 的观测值 $y(t)$, $t \in [t_0, t_1]$, 能唯一地确定系统在 t_0 时刻的任意初始状态 x_0 , 则称系统在 $[t_0, t_1]$ 上是状态完全能观测的。

例 设有线性时不变系统

$$\begin{cases} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -4 & 5 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} -5 \\ 1 \end{pmatrix} u \\ y = [1 \ -1] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + u \end{cases}$$

先考察其能控性, 引入非奇异阵

$$P = \begin{pmatrix} 1 & -5 \\ 1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} \frac{1}{6} & \frac{5}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{pmatrix},$$

则可将原系统化为最简形式

$$\begin{aligned} \begin{pmatrix} \dot{x}'_1 \\ \dot{x}'_2 \end{pmatrix} &= P^{-1}AP \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} + P^{-1}Bu \\ &= \begin{pmatrix} \frac{1}{6} & \frac{5}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} -4 & 5 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \\ &\quad + \begin{pmatrix} \frac{1}{6} & \frac{5}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} -5 \\ 1 \end{pmatrix} u = \begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \end{aligned}$$

其标量式为

$$\begin{cases} \dot{x}'_1 = x'_1 \\ \dot{x}'_2 = -5x'_2 + u \end{cases}$$

其中 x'_2 是与 u 相关的，故能控， x'_1 和 u 不相关，故不能控，因此此系统的状态不完全能控。

再考察能观测性。同样经非奇异变换，有最简形式

$$\begin{aligned} \begin{pmatrix} \dot{x}'_1 \\ \dot{x}'_2 \end{pmatrix} &= P^{-1}AP \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \\ y &= CP \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = [1 \ -1] \begin{pmatrix} 1 & -5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \\ &= [0 \ -6] \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}, \end{aligned}$$

将其展开，则为

$$\begin{cases} \dot{x}'_1 = x'_1 \\ \dot{x}'_2 = -5x'_2 \\ y = -6x'_2 \end{cases}$$

显然 x'_2 是和 y 直接相关的，所以是能观测的；而 x'_1 是和 y 无直接或间接相关，从而是不能观测的，因此系统的状态是不完全能观测的。

§ 2 随机经济系统

经济系统的状态方程往往不是确定型的，例如我们经常碰到如下的离散随机线性经济计量模型

$$y_t = ay_{t-1} + b + u_t \quad (5-8)$$

对于任何整数 t ，误差项 u_t 是一个随机变量，其均值为 0，方差为 v ，这种差分方程又叫一阶自回归过程。

方程(5-8)之解与确定型方程之解不同，它不是时间的确定性函数，方程不再严格地给出 y_t 之时间轨迹，但是它给出了对于时间 t 的解之概率分布，若 u_t 遵从正态分布，那么给定了 y_0 后， $y_1, \dots, y_k$ (对任何 k) 的联合分布是多元正态的，因为正态随机变量的线性组合是联合正态的。在这种情况下，我们只要研究均值、方差和协方差就能掌握所需的信息。

给定 y_0 后，为导出 y_t 的均值，我们可将

$$y_{i-1} = ay_{i-2} + b + u_{i-1}$$

$$y_{t-2} = ay_{t-3} + b + u_{t-2}$$

.....

不断代入前一个方程, 最后得到

$$\begin{aligned} y_t &= b + ab + a^2b + \cdots + a^{t-1}b + a^t y_0 + u_t + au_{t-1} + \cdots + a^{t-1}u_1 \\ &= b + (1-a)^{-1}(1-a^t) + a^t y_0 + u_t + au_{t-1} + \cdots + a^{t-1}u_1 \end{aligned} \quad (5-9)$$

其中 $a \neq 1$, 对上式等号两边取期望, 由于 $Eu_j = 0$, 所以

$$Ey_t = \overline{y_t} = b(1-a)^{-1}(1-a^t) + a^t y_0 \quad (5-10)$$

对(5-8)等号两边取期望, 我们也可得到

$$Ey = \overline{y_t} = a\overline{y_{t-1}} + b \quad (5-11)$$

其齐次解为 $a^t y_0$, 其特解为 $b(1-a)^{-1}(1-a^t)$, 所以上式之解就是(5-10)。

为求 y_t 的方差及 y_t 与 y_{t-k} 之协方差, 我们可以将(5-9)等号两边减去均值 $\overline{y_t}$, 并令

$$y_t^* = y_t - \overline{y_t},$$

于是

$$y_t^* = u_t + au_{t-1} + a^2u_{t-2} + \cdots + a^{t-1}u_1 \quad (5-12)$$

类似地

$$y_{t-k}^* = u_{t-k} + au_{t-k-1} + a^2u_{t-k-2} + \cdots + a^{t-k-1}u_1, \quad (5-13)$$

将(5-12)乘(5-13), 并对其取期望, 由于对于 $i \neq j$, $Eu_i u_j = 0$, 于是得到协方差:

$$\begin{aligned} \gamma(t, k) &\equiv Ey_t^* y_{t-k}^* = E(u_t + au_{t-1} + \cdots + a^{t-1}u_1) \\ &\quad \times (u_{t-k} + au_{t-k-1} + \cdots + a^{k+2(t-k-1)}u_1^2) \\ &= va^k(1 + a^2 + a^4 + \cdots + a^{2(t-k-1)}) \\ &= v(1-a^2)^{-1}(1-a^{2(t-k)})a^k \end{aligned} \quad (5-14)$$

上式中的 $k=0$ 时, 则表示 y_t 的方差。(5-14)叫作时间序列 y_t 的**自协方差函数**, 因为它表示 y_t 与它的滞后 k 期的值之间的协方差。

如果 $|a| < 1$, 则

$$\lim_{t \rightarrow \infty} \overline{y_t} = \overline{y} = b(1-a)^{-1} \quad (5-15)$$

$$\lim_{t \rightarrow \infty} \gamma(t, k) = \gamma_k = v(1-a^2)^{-1}a^k \quad (5-16)$$

如果时间序列的均值和自协方差函数不随时间变化, 就叫做**弱平稳的**, 若所有 y_t 和 y_{t+k} 的任何子集的联合分布不变, 这个序列称为**平稳的**。平稳包含着弱平稳。一个时间序列

处于稳态或均衡当且仅当它是平稳的。对于确定性经济系统而言，稳态意味着观察值不变，对于随机时间序列，稳态意味着概率分布不变。若时间序列的自协方差函数随时间的流逝而趋近于一个不变常数，则这个序列是**协方差平稳的**。这个自协方差函数仍然是 y_t 与 y_{t-k} 之间的时滞 k 的函数。

协方差平稳的系统就是满足(5—16)的系统。我们考虑其中的自协方差函数 γ_k ，当 $k=0$ 时， $\gamma_0 = v(1-a^2)^{-1}$ 是时间序列的方差，它是时间序列与其均值离差的测度。随机扰动 u_t 的方差 v 越大，方差 γ_0 就越大。

为了消除测度单位不同的影响，我们常常用**相关系数**描述 y_t 与 y_{t-k} 之间的相关。相关系数等于协方差除以两变量标准差之积。若标准差相等，它是协方差与两变量中一个变量方差的比率。时间序列的**自相关函数**是它的自协方差函数与它的方差之比。

对于模型(5—8)，在协方差平稳的条件下，自相关函数为

$$\rho_k = \frac{\gamma_k}{\gamma_0} = a^k \quad (5-17)$$

由于 $0 < a < 1$ ， $a=0.9$ 就相当大了。如果 a 的值很大，这意味着 y_t 与 y_{t-1} 之间的差别不很大。若 a 很小，例如 $a=0.05$ ，那么 y_t 几乎等于随机扰动 u_t 。 y_t 与 y_{t-1} 只有0.05的相关系数， y_t 与 y_{t-2} 只有0.0025的相关系数。

如果 a 为较大的负值，例如 $a=-0.9$ ，则 y_t 有两单位时间长度的周期，过程是振荡的。可见自协方差函数的周期波动可以反映时间序列本身的波动。

以上讨论可以推广到由线性随机差分方程组构成的多个时间序列的情形：

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_m y_{t-m} + b + u_t \quad (5-18)$$

其中 y_t 是时间序列矢量， A_i 是实系数矩阵， b 是截距矢量， u_t 是随机矢量，其均值为0，协方差阵为 V 。当 $t \neq s$ 时， u_t 与 u_s 是统计独立的。(5—18)可改写成状态方程表达式：

$$\begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-m+1} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & \cdots & A_m \\ I & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m} \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (5-19)$$

令

$$Y_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-m+1} \end{pmatrix}, \quad Y_{t-1} = \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m} \end{pmatrix}, \quad B = \begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad U = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

则(5—19)可缩写为

$$Y_t = AY_{t-1} + B + U_t \quad (5-20)$$

将(5—20)逐次迭代消去含时滞的 Y ，得到

$$Y_t = B + AB + A^2B + \cdots + A^{t-1}B + A^t Y_0 + U_t + AU_{t-1} + \cdots + A^{t-1}U_1 \quad (5-21)$$

对上式两边取数学期望，则可得到均值函数

$$EY_t = \bar{Y}_t = B + AB + A^2B + \cdots + A^{t-1}B + A^tY_0. \quad (5-22)$$

另外, 我们可将(5-20)两边直接取数学期望, 求得均值函数

$$\bar{Y}_t = A\bar{Y}_{t-1} + B \quad (5-23)$$

上式的解就是(5-22), 因为 $EY_0 = Y_0$.

当且仅当矩阵 A 的特征根的绝对值小于 1, (5-22)随 t 增大而趋于稳态。为证明此点, 可象第一章§6那样, 用线性变换 P, P^{-1} 将 A 矩阵规范化, 即令

$$A = PDP^{-1}$$

其中 D 是由 A 之特征根组成的对角阵, 矩阵 B 的列是相应于特征根的特征矢量。据此可将(5-23)改写成

$$\begin{aligned} \bar{Y}_t &= B + PDP^{-1}B + PD^2P^{-1}B + \cdots + PD^{t-1}P^{-1}B + PD^tP^{-1}Y_0 \\ &= P(I + D + D^2 + \cdots + D^{t-1})P^{-1}B + PD^tP^{-1}Y_0 \end{aligned} \quad (5-24)$$

括弧内矩阵之和为对角阵, 其对角元为

$$1 + \lambda_i + \lambda_i^2 + \cdots + \lambda_i^{t-1}$$

当且仅当所有 λ_i 的绝对值都小于 1, 则

$$\begin{aligned} \lim_{t \rightarrow \infty} (1 + \lambda_i + \cdots + \lambda_i^{t-1}) &= \lim_{t \rightarrow \infty} (1 - \lambda_i)^{-1} (1 - \lambda_i^t) \\ &= (1 - \lambda_i)^{-1}, \end{aligned}$$

即 $\bar{Y}_t$ 的稳态为

$$\lim_{t \rightarrow \infty} \bar{Y}_t = P(I - D)^{-1}P^{-1}B \quad (5-25)$$

∴

$$P(I - D)P^{-1} = I - A$$

∴

$$[P(I - D)P^{-1}]^{-1} = (I - A)^{-1}$$

即

$$P(I - D)^{-1}P^{-1} = (I - A)^{-1}$$

∴

$$\lim_{t \rightarrow \infty} \bar{Y}_t = (I - A)^{-1}B \quad (5-26)$$

考察了均值后, 我们再来讨论 y_{it} 和 $y_{jt, t-k}$ 的方差和协方差。当 k 为不同值时 y_{it} 和 $y_{jt, t-k}$ 之间的协方差表示两个时间序列在不同时点的联系程度, 并称为**互协方差**。若 y_{it} 和 y_{jt} 的协方差很大, 在同一时刻观察到的两个时间序列就有密切联系。为消除测度单位不同造成的不可比性, 我们通常用**互相关系数**描述相关程度。互相关系数是互协方差与 $\sqrt{\text{var } y_{it} \cdot \text{var } y_{jt, t-k}}$ 之商, 即与两个时间序列标准差乘积之商。为了研究互协方差, 有必要引入矩阵

$$\Gamma(t, k) = [\text{cov}(y_{it}, y_{jt, t-k})] = [\gamma_{ij}(t, k)] \quad (5-27)$$

上式中的对角线元素是各个时间序列的自协方差。从矢量的角度而言, 我们也可称之为时间

序列矢量 Y_t 的自协方差矩阵。

为求自协方差矩阵(5—27)，先定义

$$Y_t^* = Y_t - \bar{Y}_t \quad (5-28)$$

从(5—20)减去(5—23)可得

$$Y_t^* = AY_{t-1}^* + U_t \quad (5-28)$$

这意味着

$$Y_t^* = U_t + AU_{t-1} + A^2U_{t-2} + \cdots + A^{t-1}U_1 \quad (5-29)$$

其中设 $Y_0 = \bar{Y}_0$ 为常数矢量，因此 $Y_0^* = 0$ 。当 $k \geq 0$ 时，自协方差矩阵 $\Gamma(t, k)$ 为

$$\begin{aligned} E(Y_t^* Y_{t-k}^{*'}) &= E(U_t + AU_{t-1} + \cdots + A^{t-1}U_1) \\ &\quad \times (U'_{t-k} + U'_{t-k-1}A' + \cdots + U'_1A'^{t-k-1}) \\ &= E(A^k U_{t-k} U'_{t-k} + A^{k+1} U_{t-k-1} U'_{t-k-1} A' + \cdots + A^{t-1} U_1 U'_1 A'^{t-k-1}) \\ &= A^k V + A^{k+1} V A' + \cdots + A^{t-1} V A'^{t-k-1}, \quad (k \geq 0) \end{aligned} \quad (5-30)$$

其中用到关系式 $EU_t U'_s = 0$, $(t \neq s)$, $EU_t U'_t = V$ 。

当 $t \rightarrow \infty$ 时，自协方差矩阵(5—30)对任何非负整数 k 趋于一个极限的充要条件是 A 的所有特征根的绝对值都小于1。这个定理证明如下。

$$\begin{aligned} \Gamma(t, k) &= E(Y_t^* Y_{t-k}^{*'}) \\ &= A^k [V + AVA' + A^2VA'^2 + \cdots + A^{t-k-1}VA'^{t-k-1}] \\ &= A^k [V + PDP^{-1}VP'^{-1}DP' + \cdots + PD^{t-k-1}P^{-1}VP'^{-1}D^{t-k-1}P'] \\ &= A^k [P(P^{-1}VP'^{-1} + DP^{-1}VP'^{-1}D + \cdots \\ &\quad + D^{t-k-1}P^{-1}VP'^{-1}D^{t-k-1})P'] \\ &= A^k [P(W + DW'D + D^2W'D^2 + \cdots + D^{t-k-1}WD^{t-k-1})P'] \end{aligned} \quad (5-31)$$

$$\text{其中} \quad W = P^{-1}VP'^{-1}, \text{ 即 } PWP' = V \quad (5-32)$$

(5—31)最后一行括弧内矩阵的 $i-j$ 元素是

$$\begin{aligned} w_{ij} &= w_{ij}\lambda_i\lambda_j + w_{ij}(\lambda_i\lambda_j)^2 + \cdots + w_{ij}(\lambda_i\lambda_j)^{t-k-1} \\ &= w_{ij}(1 - \lambda_i\lambda_j)^{-1}(1 - \lambda_i^t\lambda_j^{t-k}) \end{aligned} \quad (5-33)$$

当且仅当所有特征根的绝对值都小于1时，这个矩阵的元素趋于极限

$$w_{ij}(1 - \lambda_i\lambda_j)^{-1} \quad (5-34)$$

利用(5—31)和(5—33)，自协方差矩阵可记为

$$\Gamma(t, k) = A^k \{P[w_{ij}(1 - \lambda_i\lambda_j)^{-1}(1 - \lambda_i^t\lambda_j^{t-k})]P^{-1}\}, \quad (k \geq 0) \quad (5-35)$$

其中 w_{ij} 由(5—32)定义, 括弧中的式子表示矩阵的 $i-j$ 元素。当 $k=0$ 时, 则同时的变量 $Y_{1t}, \dots, Y_{nt}$ 的协方差矩阵

$$\Gamma(t, 0) = P[w_{ij}(1 - \lambda_i \lambda_j)^{-1}(1 - \lambda_i^t \lambda_j^t)]P^{-1} \quad (5-36)$$

(5—35)说明, 自协方差矩阵满足关于变量 t 和 k 的矩阵差分方程, 即

$$\Gamma(t, k) = A\Gamma(t-1, k-1) = \dots = A^k\Gamma(t-k, 0), \quad (k \geq 0) \quad (5-37)$$

上式可用来从(5—36)给出的 $\Gamma(t-k, 0)$ 来计算 $\Gamma(t, k)$ 。

当且仅当 A 的所有特征根之绝对值都小于1时, 自协方差矩阵随 t 的增加而趋于一个极限。这种多元时间序列是平稳的, 稳态时自协方差函数可以写为

$$\lim_{t \rightarrow \infty} \Gamma(t, k) \equiv \Gamma_k = A^k \Gamma_0 = A^k \{P[w_{ij}(1 - \lambda_i \lambda_j)^{-1}]P^{-1}\}, \quad (k \geq 0), \quad (5-38)$$

从(5—31)的第一行还可导出

$$\Gamma(t, 0) = V + AVA' + \dots + A^{t-1}VA'^{t-1} \quad (5-39)$$

对协方差平稳时间序列矢量的稳态, 如果 A 的特征根之绝对值都小于1, 则上式右边的项将随 $t \rightarrow \infty$ 而越来越小, 则有

$$\lim_{t \rightarrow \infty} \Gamma(t, 0) = \Gamma_0 = V + AVA' + A^2VA'^2 + \dots \quad (5-40)$$

我们可以利用其收敛性来计算 Γ_0 , 进而计算 Γ_k 。

若随机变量或随机矢量是时间的函数 $x(t)$, 则我们讨论的是一个**随机过程**, 离散模型描述的随机过程叫**随机序列**

实际中常碰到的是正态随机序列, 如果随机序列 $\{x_k\}$ 满足如下条件

$$E\{x_k\} = a_k, \quad \text{var } x_k = R_k, \quad \text{cov}(x_k, x_l) = R_{kl} \quad (k \neq l)$$

则叫正态随机序列。对于正态随机序列, 序列中任意有限个矢量之线性组合构成的随机矢量, 仍然具有正态分布密度。

据此可引入白噪声序列和有色噪声序列概念。如果随机序列 $\{x_k\}$ 中各个时刻的随机矢量间是两两不相关的, 即成立如下的关系式

$$\text{cov}(x_k, x_l) = 0, \quad (k \neq l), \quad \text{var } x_k = R_k \quad k = 1, 2, \dots,$$

则称此随机序列为**白噪声序列**, 否则称为**有色噪声序列**。

如果一个白噪声序列同时又是正态随机序列, 则称此序列为**正态白噪声序列**(或称**高斯白噪声**)。正态白噪声序列之充要条件是各时刻随机矢量间的两两不相关性等价于相互独立性。

§ 3 滤波理论

若某个变量 a 之第 k 次观测值为 x_k , 每次观测的干扰即噪声为 u_k , 则我们有

$$x_k = a + u_k, \quad k = 1, \dots, n \quad (5-41)$$

u_k 互相独立, 且服从同一正态分布 $N(0, \sigma^2)$ 。

我们可以用

$$\bar{x}_n = (x_1 + \cdots + x_n)/n \quad (5-42)$$

来估计 a 值, 精度将比用 x_k 估计 a 高得多, 因为 $\bar{x}_n - a$ 服从均值为 0, 方差为 $\frac{\sigma^2}{n}$ 的正态分布, 随着观测次数 n 的加大, 这个精度越来越高, 取 n 次观测的平均值来估计 a , 这就是一种滤波运算, 它可以滤掉噪声 u_k 的影响, 使估计误差减少。

若又增加了第 $n+1$ 个观测值 x_{n+1} , 我们可以用下式估计 a

$$\bar{x}_{n+1} = (x_1 + \cdots + x_n + x_{n+1})/(n+1). \quad (5-43)$$

进行这种计算必须保存原来 n 个数据, n 越大, 要保存的数据越多。这相当于维纳、柯尔莫哥洛夫的滤波思想。

我们也可以改用下式估计 a

$$\begin{aligned} \bar{x}_{n+1} &= (n\bar{x}_n + x_{n+1})/(n+1) \\ &= [(n+1)\bar{x}_n + x_{n+1} - \bar{x}_n]/(n+1) \\ &= \bar{x}_n + \frac{x_{n+1} - \bar{x}_n}{n+1} \end{aligned} \quad (5-44)$$

用这种算法, 计算机只要保存以前的一个估计值 $\bar{x}_n$, 测出 x_{n+1} 后, 可马上算出 $\bar{x}_{n+1}$ 。式中 $x_{n+1} - \bar{x}_n$ 是新数据与原估计值 $\bar{x}_n$ 相差的值, 叫做**新息**, $\frac{1}{n+1}$ 叫**增益因子**。以上是一种递推滤波运算, 它就是卡尔曼, 布西的滤波思想。

我们首先考虑一维滤波的算法。我们要根据对 a 的两个测量量 x_1, x_2 对 a 进行估计, x_1, x_2 的数学期望 $Ex_1 = Ex_2 = a$, 方差为

$$\begin{aligned} Dx_1 &= E(x_1 - a)^2 = \sigma_1^2 \\ Dx_2 &= E(x_2 - a)^2 = \sigma_2^2 \end{aligned} \quad (5-45)$$

估计算式为

$$\hat{x} = (1-w)x_1 + wx_2, \quad 0 \leq w \leq 1 \quad (5-46)$$

$$\begin{aligned} \text{显然, } E\hat{x} &= E[(1-w)x_1 + wx_2] \\ &= (1-w)x_1 + wEx_2 = a \end{aligned} \quad (5-47)$$

因此 $\hat{x}$ 是 a 的无偏估计量。我们的问题是怎样选取 w 之值以使 $\hat{x}$ 的方差

$$D\hat{x} = E(\hat{x} - a)^2 \quad (5-48)$$

极小化。

$$\begin{aligned}
 \therefore \quad (\hat{x} - a)^2 &= [(1-w)x_1 + wx_2 - a]^2 \\
 &= [(1-w)(x_1 - a) + w(x_2 - a)]^2 \\
 &= (1-w)^2(x_1 - a)^2 + w^2(x_2 - a)^2 + 2(1-w)w(x_1 - a)(x_2 - a)
 \end{aligned}$$

且 $(x_1 - a)$ 与 $(x_2 - a)$ 相互独立

$$\begin{aligned}
 \therefore \quad \hat{\sigma}_2^2 &= E(\hat{x} - a)^2 \\
 &= E(1-w)^2(x_1 - a)^2 + Ew^2(x_2 - a)^2 \\
 &= (1-w)^2\sigma_1^2 + w^2\sigma_2^2
 \end{aligned} \tag{5-49}$$

将上式对 w 求导数且令其等于 0，则

$$\begin{aligned}
 \frac{d\hat{\sigma}^2}{dw} &= -2(1-w)\sigma_1^2 + 2w\sigma_2^2 = 0 \\
 w &= \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}
 \end{aligned} \tag{5-50a}$$

$$1 - w = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{5-50b}$$

将上两式代入(5-46)，则得到 a 的线性最小方差估计

$$\begin{aligned}
 \hat{x} &= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}x_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}x_2 \\
 &= x_1 + w(x_2 - x_1)
 \end{aligned} \tag{5-51}$$

代入(5-49)，可算出 $\hat{x}$ 之方差

$$\hat{\sigma}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = (1-w)\sigma_1^2 \tag{5-52}$$

总结上述推导，可得出--维滤波算式

$$\hat{x} = x_1 + w(x_2 - x_1) \tag{5-53a}$$

$$\hat{\sigma}^2 = (1-w)\sigma_1^2 \tag{5-53b}$$

$$w = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \tag{5-53c}$$

(a) 式说明，方差最小的新的最优估计量是在原估计量基础上加上增益部分构成。增益部分是新息 $x_2 - x_1$ 经增益因子 w “放大” 形成。 $\sigma_1 = \sigma_2$ 时， $w = 0.5$ ， $\sigma_1 > \sigma_2$ 时， $w > 0.5$ ，
 则 $\hat{x}$ 靠近 x_2 ，即靠近方差较小的数据；

(b) 式说明新估计量的方差 $\hat{\sigma}^2$ 确实小于原估计量的方差 σ_1^2 ，因为 $0 < 1 - w < 1$ ；

(c) 式给出了增益因子 w 的算法。

例 有两种统计报表算得的国民收入不相同。两种报表分析的结果都无系统偏差，即统计国民收入之数学期望都等于其实际值。已知甲报表和乙报表分析值的方差分别为 0.04 和 0.01，甲报表统计得国民收入 0.19 亿，乙报表统计得 0.20 亿，试用乙报表的统计对甲报表的统计结果进行滤波。

$$\begin{aligned}x_1 &= 0.19, & \sigma_1^2 &= 0.04 \\x_2 &= 0.20, & \sigma_2^2 &= 0.01 \\w &= \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} = \frac{0.04}{0.04 + 0.01} = 0.8, \\ \hat{x} &= x_1 + w(x_2 - x_1) \\ &= 0.19 + 0.8(0.20 - 0.19) = 0.198.\end{aligned}$$

下面我们考虑多维滤波问题。设 X_1, X_2 是 n 维矢量 X^* 的两个独立无偏估计量，

$$EX_1 = EX_2 = X^*.$$

两个估计量的方差阵分别为

$$\begin{aligned}D(X_1) &= E(X_1 - X^*)(X_1 - X^*)' = V_1 \\D(X_2) &= E(X_2 - X^*)(X_2 - X^*)' = V_2\end{aligned}\tag{5-54}$$

V_1, V_2 的行列式不为 0，因而是两个正定方阵。我们的问题是选择非负 n 阶方阵 W ， $0 \leq W \leq I$ ，使新的估计量

$$\hat{X} = (I - W)X_1 + WX_2\tag{5-55}$$

的方差阵 $D\hat{X}$ 尽可能小。许多实际问题中，两个估计量的维数可能不同，一个是 X^* 的无偏估计量 X_1 ，另一个是 X^* 的线性变换 $Y = HX^*$ 的无偏估计量 Y_2 ，且二者是独立的，即 $EX_1 = X^*, EY_2 = Y^* = HX^*$ ，或

$$\begin{aligned}X_1 &= X^* + U_1 & EU_1 &= 0 \\Y_2 &= HX^* + U_2 & EU_2 &= 0\end{aligned}\tag{5-56}$$

X_1 和 X^* 都是 n 维矢量， U_1 是 n 维噪声矢量， Y_2 是 m 维矢量， H 是 $m \times n$ 矩阵，叫观测矩阵， U_2 是 m 维观测噪声矢量，方差阵 $DU_2 = R$ 。

这种滤波的问题是要通过观测值 Y_2 来改进 X^* 的初步估计 X_1 。

设 X_1 的方差阵

$$DX_1 = E(X_1 - X^*)(X_1 - X^*)' = V_1,$$

可以设想有 X_2 使 $Y_2 = HX_2$ 。 Y_2 既是 HX^* 的无偏估计，所以 X_2 也是 X^* 的一个无偏估计，我们应取新估计量

$$\begin{aligned}
\hat{X} &= X_1 + W(X_2 - X_1) \\
&= X_1 + KH(X_2 - X_1) \\
&= X_1 + K(Y_2 - HX_1)
\end{aligned} \tag{5-57}$$

最小方差滤波是选择 K 使 $D\hat{X}$ 最小化。我们有

$$\begin{aligned}
D\hat{X} &= E(\hat{X} - X^*)(\hat{X} - X^*)' \\
&= E(X_1 + KY_2 - KHX_1 - X^*)(X_1 + KY_2 - KHX_1 - X^*)' \\
\therefore \quad &X_1 + KY_2 - KHX_1 - X^* \\
&= (X_1 - X^*) + K(HX^* + U_2) - KHX_1 \\
&= (X_1 - X^*)(I - KH) + KU_2 \\
\therefore \quad &(X_1 + KY_2 - KHX_1 - X^*)' \\
&= [(I - KH)(X_1 - X^*) + KU_2]' \\
&= (X_1 - X^*)'(I - KH)' + U_2'K' \\
\therefore \quad D\hat{X} &= E[(I - KH)(X_1 - X^*) + KU_2][(X_1 - X^*)'(I - KH)' + U_2'K'] \\
&= E\{(I - KH)(X_1 - X^*)(X_1 - X^*)'(I - KH)' + KU_2U_2'K' \\
&\quad + (I - KH)(X_1 - X^*)U_2'K' + KU_2(X_1 - X^*)'(I - KH)'\} \\
&= (I - KH)V_1(I - KH)' + KKK' \\
&= V_1 - KHV_1 - V_1H'K' + KHV_1H'K' + KKK'
\end{aligned}$$

将 $D\hat{X}$ 对 K 求导数, 令其等于 0, 解出,

$$K = V_1H'(HV_1H' + R)^{-1} \tag{5-58a}$$

代回(5-57), 则可解出

$$\hat{X} = X_1 + K(Y_2 - HX_1) \tag{5-58b}$$

将(5-58a)代回 $D\hat{X}$ 的表达式, 则有

$$\begin{aligned}
D\hat{X} &= V_1 - KHV_1 - V_1H'K' + K(HV_1H' + R)K' \\
&= V_1 - KHV_1 - V_1H'K' + KK^{-1}V_1H'K' \\
&= (I - KH)V_1
\end{aligned} \tag{5-58c}$$

(5-58a, b, c)就是相当于(5-53a, b, c)的滤波算式。

例 有25年的统计数据 (s_k, y_k) , $k = 1, \dots, 25$. 利用最小二乘法求得储蓄 s 与收入

Y 的经济计量模型为

$$\hat{s} = 18.05 + 0.221y$$

对于这个模型而言, $X_1 = (18.05, 0.221)'$ 是总体回归方程 $\hat{y} = a + by$ 中 $X^* = (a, b)'$ 的估计值。假设

$$E(X_1 - X^*)(X_1 - X^*)' = V_1 = \begin{pmatrix} 0.04 & 0.002 \\ 0.002 & 0.04 \end{pmatrix}.$$

我们又得到第26年的统计数据 $y = 100$, $s = 40$, 且设

$$Y = HX^* + U_2 = a + by + U_2$$

中噪声 U_2 的方差 $R = 0.02$, 这里 $H = (1, y) = (1, 100)$, 根据以上数据, 可利用 (5—58a) 计算 K :

$$\begin{aligned} K &= \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} = V_1 H' (H V_1 H' + R)^{-1} \\ &= \begin{pmatrix} 0.04 & 0.002 \\ 0.002 & 0.04 \end{pmatrix} \begin{pmatrix} 1 \\ 100 \end{pmatrix} \left\{ [1, 100] \begin{pmatrix} 0.04 & 0.002 \\ 0.002 & 0.04 \end{pmatrix} \begin{pmatrix} 1 \\ 100 \end{pmatrix} + 0.02 \right\}^{-1} \\ &= \begin{pmatrix} 0.0006 \\ 0.01 \end{pmatrix} \end{aligned}$$

然后利用 (5—58b) 计算 $\hat{X} = (\hat{a}, \hat{b})'$:

$$\begin{aligned} \hat{X} &= \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = X_1 + K(Y_2 - HX_1) \\ &= \begin{pmatrix} 18.05 \\ 0.221 \end{pmatrix} + \begin{pmatrix} 0.0006 \\ 0.01 \end{pmatrix} \left\{ 40 - [1, 100] \begin{pmatrix} 18.05 \\ 0.221 \end{pmatrix} \right\} \\ &= \begin{pmatrix} 18.04991 \\ 0.2195 \end{pmatrix}. \end{aligned}$$

于是滤波后的经济计量模型为

$$s = 18.04991 + 0.2195y$$

将已知数据代入 (5—58c), 可算出

$$DX = (I - KH)V_1$$

$$= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.0006 \\ 0.01 \end{bmatrix} [1 \quad 100] \right\} \begin{bmatrix} 0.04 & 0.002 \\ 0.002 & 0.04 \end{bmatrix} \\ = \begin{bmatrix} 0.039856 & -0.0004012 \\ -0.0004 & -0.002 \end{bmatrix}$$

从此例中易知 $D\hat{X}$ 比 V_1 要小得多。

以上分析可以推广到动态线性系统的递推滤波，这就是卡尔曼滤波。

设动态系统的状态方程和输出方程分别为

$$\begin{cases} X_k = \phi_{k,k-1}X_{k-1} + B_{k-1}U_{k-1} + \Gamma_{k-1}W_{k-1} & (5-59a) \\ Y_k = HX_k + V_k & (5-59b) \end{cases}$$

其中 X_k 是 n 维状态矢量， U_{k-1} 是 r 维控制矢量， W_{k-1} 是 l 维模型噪声， $\phi_{k,k-1}$ 是 $n \times n$ 转移矩阵， B_{k-1} ， Γ_{k-1} 分别是 $n \times r$ ， $n \times l$ 控制矩阵和噪声矩阵。 Y_k 是 m 维观测矢量， H_k 是 $m \times n$ 观测矩阵， V_k 是 m 维噪声矢量。假设 W_k ， V_k 都是白噪声，

$$\begin{aligned} EW_k &= EV_k = 0 \\ E(W_k W_l') &= 0 & (k \neq l) \\ E(V_k V_l') &= 0 & (k \neq l) \\ E(W_k V_l') &= 0 \\ E(W_k W_k') &= Q_k & (5-60) \\ E(V_k V_k') &= R_k \\ EX_0 &= \mu_0 \\ DX_0 &= P_0. \end{aligned}$$

其中 X_0 为随机初始状态。

先取 $EX_0 = \mu_0$ 为初始状态 X_0 的最优估计量 $\hat{X}_0$ ，使得方差阵最小。设 X_0 与 W_k ， X_0 与 V_k 都不相关。由于 W_0 的均值为 0，所以取 X_1 的初步估计量为

$$X_1'' = \phi_{1,0}\hat{X}_0 + B_0U_0 \quad (5-61)$$

易知，它是 X_1 的无偏估计。由 $k=1$ 时的观测

$$Y_1 = H_1X_1 + V_1 \quad (5-62)$$

按前文介绍的滤波的思路，应取 X_1 的新估计量为

$$\hat{X}_1 = X_1'' + K_1(Y_1 + H_1X_1'') \quad (5-63)$$

易知， $\hat{X}_1$ 应为 X_1 的无偏估计。仿此作 X_2 的初步估计

$$X_2'' = \phi_{2,1} \hat{X}_1 + B_1 U_1 \quad (5-64)$$

改进后的 X_2 估计为

$$\hat{X}_2 = X_2'' + K_2(Y_2 - H_2 X_2'') \quad (5-65)$$

X_2'' 与 $\hat{X}_2$ 均为 X_2 的无偏估计。如此递推下去，我们有一般的递推滤波算式

$$\hat{X}_k = X_k'' + K_k(Y_k - H_k X_k'') \quad (5-66a)$$

$$X_k'' = \phi_{k,k-1} \hat{X}_{k-1} + B_{k-1} U_{k-1} \quad (5-66b)$$

$$K_k = P_k'' H_k' (H_k P_k'' H_k' + R_k)^{-1} \quad (5-66c)$$

$$\begin{aligned} P_k'' &= E(X_k'' - X_k)(X_k'' - X_k)' \\ &= \phi_{k,k-1} P_{k-1} (\phi_{k,k-1})' + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}' \end{aligned} \quad (5-66d)$$

$$P_k = (I - K_k H_k) P_k'' \quad (5-66e)$$

卡尔曼滤波用来处理含有色噪声的问题与处理白噪声模型，计算上并无原则差别，这正是其突出优点。详细的证明需要较多的数学工具，我们只要求学生能运用此种方法。

例 设有下列线性随机状态方程

$$\begin{cases} x_k = 0.5x_{k-1} + w_{k-1} \\ y_k = x_k + v_k \end{cases}$$

模型噪声的期望 $E\{w_k\} = 0$ ，协方差 $\text{cov}\{w_k, w_l\} = 1\delta_{kl}$ 。量测噪声的期望 $E\{v_k\} = 0$ ，协方差 $\text{cov}\{v_k, v_l\} = 2\delta_{kl}$ ，其中

$$\delta_{kl} = \begin{cases} 1 & k=l \\ 0 & k \neq l \end{cases}$$

初始状态之期望 $E\{x_0\} = 0$ ，方差 $\text{var}\{x_0\} = P_0 = 1$ 。

(i) 定出(5-66a—e)之滤波算式。从已知条件知 $\phi_{k,k-1} = 0.5$ ， $B = 0$ ， $\Gamma_{k-1} = 1$ ， $H = 1$ ， $R = 2$ ， $Q = 1$ 。代入(5-66)，则有

$$\begin{cases} \hat{x}_k = x_k'' + K_k(y_k - H_k x_k'') \\ \quad = \phi_{k,k-1} \hat{x}_{k-1} + K_k(y_k - H \phi_{k,k-1} \hat{x}_{k-1}) \\ \quad = 0.5 \hat{x}_{k-1} + K_k(y_k - 0.5 \hat{x}_{k-1}) \\ K_k = P_k'' (P_k'' + 2)^{-1} \\ P_k'' = 0.25 P_{k-1} + 1 \\ P_k = (1 - K_k) P_k'' \end{cases}$$

(ii) 进行递推计算：

当 $k=1$ 时, $y_1=4$,

由 $P_0=1$, $y_1=4$, $\hat{x}_0=0$, 可算出

$$P''=1.25, K_1=\frac{1.25}{3.25}=\frac{5}{13}, \hat{x}_1=\frac{5}{13}(4-0)=\frac{20}{13},$$

$$P_1=\left(1-\frac{5}{13}\right)\times 1.25=\frac{10}{13}.$$

(iii) 进行第二阶段递推计算,

$$k=2, y_2=2$$

由 $P_1=\frac{10}{13}$, $y_2=2$, $\hat{x}_1=\frac{20}{13}$, 可算出

$$P''=0.25\times\frac{10}{13}+1=\frac{31}{26}, K_2=\frac{31}{26}/\frac{83}{26}=\frac{31}{83},$$

$$\hat{x}_2=0.5\times\frac{20}{13}+\frac{31}{83}\left(2-0.5\times\frac{20}{13}\right)=\frac{1326}{1079},$$

$$P_2=\left(1-\frac{31}{83}\right)\frac{31}{26}=\frac{62}{83}.$$

如此等等, 照此步骤可不断算下去。

§4 动态规划与随机最优经济控制

给出一个随机动态经济系统

$$x_{k+1}=f_k(x_k, u_k, w_k), k=0, 1, \dots, N-1 \quad (5-67)$$

其中 x_k 为 $N-1$ 维状态矢量, u_k 为 $N-1$ 维控制矢量, u_k 只在容许控制集合 $U_k(x_k)$ 中取值。随机干扰 w_k 遵从给定概率分布, 最优概率密度 p_k 与 x_k, u_k 有关, 而与前期干扰 $w_{k-1}, \dots, w_0$ 无关。设 u_k 中的容许控制为 μ_k 。给定初始状态 x_0 , 我们的问题是求一个容许控制规律

$$\pi = \{\mu_0, \dots, \mu_{N-1}\}$$

以使下列费用函数极小化, 即

$$\min_{\pi} J_{\pi}(x_0) = E_{w_k} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k[x_k, \mu_k(x_k), w_k] \right\} \quad (5-68)$$

$k=0, \dots, N-1$

且服从状态方程约束

$$x_{k+1}=f_k[x_k, \mu_k(x_k), w_k], k=0, 1, \dots, N-1, \quad (5-69)$$

我们用一个库存控制的例子将上述动态规划模型具体化。设

x_k —— k 阶段初可利用的存货,

h —— k 阶段单位库存费用,

u_k —— k 阶段初始时刻定购的货物, (假设马上到货)

w_k —— k 阶段遵从给定概率分布的需求,

c ——单位定货费用,

p ——单位缺货费用

又设, $w_0, \dots, w_{N-1}$ 为独立随机变量, 过量需求可累积, 只要一到货则可冲销累积的过量需求, 过量需求用负的存货表示。状态方程为

$$x_{k+1} = x_k + u_k - w_k. \quad (5-70)$$

我们的目的是使下列费用函数最小化

$$E \left\{ \sum_{k=0}^{N-1} [cu + p \max(0, w_k - x_k - u_k) + h \max(0, x_k + u_k - w_k)] \right\} \quad (5-71)$$

这里我们假定 N 阶段初的过量需求是损失, 而留下的存货 x_N 没有支出, 即终期费用为 0。我们的目的是要找出一个容许的定货计划 $\{\mu_0, \dots, \mu_{N-1}\}$, 其中 $\mu_k = \mu_k(x_k)$, $k=0, \dots, N-1$, 以使预期费用最小化。

若令前述基本问题(5-68)中的

$$g_N(x_N) = 0$$

$$g_k(x_k, u_k, w_k) = cu_k + p \max(0, w_k - x_k - u_k) + h \max(0, x_k + u_k - w_k),$$

及(5-69)中的

$$f_k(x_k, u_k, w_k) = x_k + u_k - w_k,$$

$$U_k(x_k) = [0, \infty).$$

则库存控制问题就是前述基本问题的一种。

动态规划的算法以贝尔曼最优性原理为基础。

假定 $\{\mu_0^*, \mu_1^*, \dots, \mu_{N-1}^*\}$ 为最优控制法则, 考虑 i 时刻 x_i 状态下的子问题, 我们要使从 i 到 N 的序贯总费用最小, 即

$$\min_i E \left\{ g_N(x_N) + \sum_{k=i}^{N-1} g_k[x_k, \mu_k(x_k), w_k] \right\}, \quad (5-72)$$

则控制法 $\{\mu_i^*, \mu_{i+1}^*, \dots, \mu_{N-1}^*\}$ 对于子问题亦为最优。以此为基础我们可以导出动态规划的算法。我们以库存控制问题为例导出这种算法。

(i) 在库存问题中, 设 $N-1$ 阶段可利用的存货为 x_{N-1} , 最优定货为 $u_{N-1}^* = \mu_{N-1}^*(x_{N-1})$, 它使

$$\min_{w_{N-1}} E \{ cu_{N-1} + h \max(0, x_{N-1} + u_{N-1} - w_{N-1}) + p \max(0, w_{N-1} - x_{N-1} - u_{N-1}) \}$$

又设最后阶段的最优费用为 $J_{N-1}(x_{N-1})$, 则

$$J_{N-1}(x_{N-1}) = \min_{u_{N-1} \geq 0} E \{ cu_{N-1} + h \max(0, x_{N-1} + u_{N-1} - w_{N-1}) +$$

$$+ p \max(0, w_{N-1} - x_{N-1} - u_{N-1}) \} \quad (5-73)$$

J_{N-1} 为存货 x_{N-1} 的函数, 算出 J_{N-1} 的同时, 我们就可确定最优定货计划 $\mu_{N-1}^*(x_{N-1})$, 其中 $\mu_{N-1}^*(x_{N-1}) \geq 0$ 使上式右端对于每一个 x_{N-1} 最小化。

(i i) 在 $N-2$ 阶段, 初始时刻之库存为 x_{N-2} , 定货为 $u_{N-2} = \mu_{N-2}(x_{N-2})$, 它不仅使 $N-2$ 阶段之预期费用最小化, 且使下式最小化

($N-2$ 阶段之预期费用) + ($N-1$ 阶段运用最优计划确定的 $N-1$ 阶段预期费用)
上式就是

$$\begin{aligned} & E_{w_{N-2}} \{ cu_{N-2} + h \max(0, x_{N-2} + u_{N-2} - w_{N-2}) + p \max(0, w_{N-2} - x_{N-2} - u_{N-2}) \} \\ & + E_{w_{N-2}} \{ J_{N-1}(x_{N-1}) \} \end{aligned} \quad (5-74)$$

利用状态方程

$$x_{N-1} = x_{N-2} + u_{N-2} - w_{N-2} \quad (5-75)$$

于是(5-74)最后一项可写为

$$E_{w_{N-2}} \{ J_{N-1}(x_{N-2} + u_{N-2} - w_{N-2}) \} \quad (5-76)$$

于是最后两个阶段的最优费用 $J_{N-2}(x_{N-2})$ 可写为

$$\begin{aligned} J_{N-2}(x_{N-2}) = & \min_{u_{N-2} \geq 0} E_{w_{N-2}} \{ cu_{N-2} + h \max(0, x_{N-2} + u_{N-2} - w_{N-2}) \\ & + p \max(0, w_{N-2} - x_{N-2} - u_{N-2}) + J_{N-1}(x_{N-2} + u_{N-2} - w_{N-2}) \} \end{aligned} \quad (5-77)$$

对于每一个 x_{N-2} 算出 $J_{N-2}(x_{N-2})$, 同时也就确定了最优定货计划 $\mu_{N-2}^*(x_{N-2})$

(iii) k 阶段之初始库存为 x_k , 定货为 u_k , 子目标为使下式最小化:

(k 阶段之期望费用) + ($k+1$ 阶段, ..., $N-1$ 阶段给出相应最优计划时的期望费用)。

令 $J_k(x_k)$ 为 k 阶段最优费用值, 则

$$\begin{aligned} J_k(x_k) = & \min_{u_k \geq 0} E_{w_k} \{ cu_k + h \max(0, x_k + u_k - w_k) \\ & + p \max(0, w_k - x_k - u_k) + J_{k+1}(x_k + u_k - w_k) \} \end{aligned} \quad (5-78)$$

这就是库存问题动态规划之求解方程。

这种算法可写成如下命题

命题 令 $J^*(x_0)$ 为费用函数(5-72)之最优值, 则

$$J^*(x_0) = J_0(x_0)$$

其中函数 J_0 由下列动态规划解法之最后一步给出。这一解法从最后阶段向初始阶段倒推。

$$J_N(x_N) = g_N(x_N) \quad (5-79)$$

$$J_k(x_k) = \inf_{u_k \in U_k(x_k)} E \{g_k(x_k, u_k, w_k) + J_{k-1}[f_k(x_k, u_k, w_k)]\} \\ k = 0, 1, \dots, N-1 \quad (5-80)$$

若 $u_k^* = \mu_k^*(x_k)$ 使(5-80)右边对于每一 x_k 和 k 最小化, 则控制法则 $\{\mu_0^*, \dots, \mu_{N-1}^*\}$ 为最优。

上述命题可证明如下。

w_k 之概率只依赖于 x_k 和 u_k , 而与前期干扰 $w_0, \dots, w_{k-1}$ 无关, 这使我们能将费用 $J^*(x_0)$ 之最优值写为

$$J^*(x_0) = \inf_{\mu_0, \dots, \mu_{N-1}} J(x_0, \mu_0, \dots, \mu_{N-1}) \\ = \inf_{\mu_0, \dots, \mu_{N-1}} [E\{g_0[x_0, \mu_0(x_0), w_0] + E\{g_1[x_1, \mu_1(x_1), w_1] + \dots \\ + E\{g_{N-1}[x_{N-1}, \mu_{N-1}(x_{N-1}), w_{N-1}] + g_N(x_N)\} \dots\}]] \quad (5-81)$$

其中 w_k 的期望是 x_k 和 $\mu_k(x_k)$ 上的期望 ($k = 0, 1, \dots, N-1$)。上述表达式也可记为

$$J^*(x_0) = \inf_{\mu_0} [E\{g_0[x_0, \mu_0(x_0), w_0] + \inf_{\mu_1} [E\{g_1[x_1, \mu_1(x_1), w_1] + \dots \\ + \inf_{\mu_{N-1}} [E\{g_{N-1}[x_{N-1}, \mu_{N-1}(x_{N-1}), w_{N-1}] + g_N(x_N)\}]] \dots]]. \quad (5-82)$$

上述方程中, μ_k 是对于所有 x_k 和 k 使得目标函数最小化, 且 $\mu_k(x_k) \in U_k(x_k)$ 。另外这种最小化也服从状态方程 $x_{k+1} = f_k[x_k, \mu_k(x_k), w_k]$ 。

由于

$$\inf_{\mu \in M} F[x, u(x)] = \inf_{u \in U(x)} F(x, u) \quad (5-83)$$

其中 M 为对于所有 x , 使 $\mu(x) \in U(x)$ 的所有 $\mu(x)$ 之集合。利用(5-82)和状态方程 $x_{k+1} = f_k(x_k, u_k, w_k)$, 且将(5-81)中的 J_k 引入(5-82), 我们得到

$$J^*(x_0) = J_0(x_0). \quad (5-84)$$

显然, 若 $\mu_k^*(x_k)$ 对于每一 x_k 和 k 使(5-81)的右边最小化, 则 $\{\mu_0^*, \dots, \mu_{N-1}^*\}$ 为最优控制法则。

例 若前例的库存问题中, 库存和需求为非负整数变量, 且对存货 $(x_k + u_k)$ 有一上界, 过量需求 $(w_k - x_k - u_k)$ 是损失, 于是存货之状态方程变成

$$x_{k+1} = \max(0, x_k + u_k - w_k) \quad (1)$$

假定存货 $(x_k + u_k)$ 的最大数量为 2 单位; 问题共分为 3 个阶段; 库存费用 h 和定货费用 c 都

为 1 单位, 缺货费用 p 为 3 单位, 需求 w_k 对于所有阶段有同样的概率分布:

$$P(w_k = 0) = 0.1, \quad P(w_k = 1) = 0.7, \quad (2)$$

$$P(w_k = 2) = 0.2;$$

固定费用 $K = 0$, 初始存货 $x_0 = 0$. 我们利用(5-79)和(5-80)之动态规划算法进行递推计算:

$$(i) \quad J_3(x_3) = 0,$$

$$J_k(x_k) = \min_{0 \leq u_k \leq 2 - x_k} E\{u_k + \max(0, x_k + u_k - w_k) + 3\max(0, w_k - x_k - u_k) + J_{k+1}[\max(0, x_k + u_k - w_k)]\} \quad (3)$$

其中 x_k, u_k, w_k 可取值 0, 1, 2; $k = 0, 1, 2$.

(ii) 对于三个可能的状态我们计算 $J_2(x_2)$:

(a) $x = 0$.

$$\begin{aligned} J_2(0) &= \min_{u_2=0,1,2} E\{u_2 + \max(0, u_2 - w_2) + 3\max(0, w_2 - u_2)\} \\ &= \min_{u_2=0,1,2} \{u_2 + 0.1[\max(0, u_2) + 3\max(0, -u_2)] \\ &\quad + 0.7[\max(0, u_2 - 1) + 3\max(0, 1 - u_2)] + 0.2[\max(0, u_2 - 2) \\ &\quad + 3\max(0, 2 - u_2)]\} \end{aligned} \quad (4)$$

对于 u_2 可能的 3 个值, 我们分别计算(4)等号右边的期望值:

$$u_2 = 0: E\{\cdot\} = 0.7 \times 3 \times 1 + 0.2 \times 3 \times 2 = 3.3$$

$$u_2 = 1: E\{\cdot\} = 1 + 0.1 \times 1 + 0.2 \times 3 \times 1 = 1.7$$

$$u_2 = 2: E\{\cdot\} = 2 + 0.1 \times 2 + 0.7 \times 1 = 2.9$$

我们从中选择 $J_2(0) = 1.7, \mu_2^*(0) = 1$.

(b) $x_2 = 1$,

$$\begin{aligned} J_2(1) &= \min_{u_2=0,1} E\{u_2 + \max(0, 1 + u_2 - w_2) + 3\max(0, w_2 - 1 - u_2)\} \\ &= \min_{u_2=0,1} \{u_2 + 0.1[\max(0, 1 + u_2) + 3\max(0, -1 - u_2)] + 0.7[\max(0, u_2) \\ &\quad + 3\max(0, -u_2)] + 0.2[\max(0, u_2 - 1) + 3\max(0, 1 - u_2)]\} \end{aligned} \quad (5)$$

对于 u_2 的可能值

$$u_2 = 0: E\{\cdot\} = 0.1 \times 1 + 0.2 \times 3 \times 1 = 0.7;$$

$$u_2 = 1: E\{\cdot\} = 1 + 0.1 \times 2 + 2 + 0.7 \times 1 = 1.9.$$

因此选择 $J_2(1) = 0.7$, $\mu_2^*(1) = 0$.

(c) $x_2 = 2$, 我们有

$$\begin{aligned} J_2(2) &= E\{\max(0, 2 - w_2) + 3\max(0, w_2 - 2)\} \\ &= 0.1 \times 2 + 0.7 \times 1 = 0.9. \end{aligned} \quad (6)$$

所以 $J_2(2) = 0.9$, $\mu_2^*(2) = 0$.

(iii) 对于三种可能的状态 $x_1 = 0, 1, 2$, 利用前阶段已得到的 $J_2(0)$, $J_2(1)$, $J_2(2)$ 之值计算 $J_1(x_1)$.

$$\begin{aligned} (a) \quad J_1(0) &= \min_{u_1=0,1,2} E\{u_1 + \max(0, u_1 - w_1) + 3\max(0, w_1 - u_1) \\ &\quad + J_2[\max(0, u_1 - w_1)]\} \end{aligned} \quad (7)$$

$$u_1 = 0: E\{\cdot\} = 0.1 \times J_2(0) + 0.7[3 \times 1 + J_2(0)] + 0.2[3 \times 2 + J_2(0)] = 5.0;$$

$$u_1 = 1: E\{\cdot\} = 1 + 0.1[1 + J_2(1)] + 0.7 \times J_2(0) + 0.2[3 \times 1 + J_2(0)] = 3.3;$$

$$u_1 = 2: E\{\cdot\} = 2 + 0.1[2 + J_2(2)] + 0.7[1 + J_2(1)] + 0.2 \times J_2(0) = 3.82;$$

因此 $J_1(0) = 3.3$, $\mu_1^*(0) = 1$;

$$\begin{aligned} (b) \quad J_1(1) &= \min_{u_1=0,1} E\{u_1 + \max(0, 1 + u_1 - w_1) + 3\max(0, w_1 - 1 - u_1) \\ &\quad + J_2[\max(0, 1 + u_1 - w_1)]\} \end{aligned} \quad (8)$$

$$u_1 = 0: E\{\cdot\} = 0.1[1 + J_2(1)] + 0.7 \times J_2(0) + 0.2[3 \times 1 + J_2(0)] = 2.3;$$

$$u_1 = 1: E\{\cdot\} = 1 + 0.1[2 + J_2(2)] + 0.7[1 + J_2(1)] + 0.2 \times J_2(0) = 2.82,$$

所以 $J_1(1) = 2.3$, $\mu_1^*(1) = 0$.

$$\begin{aligned} (c) \quad J_1(2) &= E\{\max(0, 2 - w_1) + 3\max(0, w_1 - 2) + J_2[\max(0, 2 - w_1)]\} \\ &= 0.1[2 + J_2(2)] + 0.7[1 + J_2(1)] + 0.2 \times J_2(0) = 1.82 \end{aligned} \quad (9)$$

所以 $J_1(2) = 1.82$, $\mu_1^*(2) = 0$.

(iv) 由于初始状态 $x_0 = 0$, 所以只要计算

$$\begin{aligned} J_0(0) &= \min_{u_0=0,1,2} E\{u_0 + \max(0, u_0 - w_0) + 3\max(0, w_0 - u_0) \\ &\quad + J_1[\max(0, u_0 - w_0)]\} \end{aligned} \quad (10)$$

$$u_0 = 0: E\{\cdot\} = 0.1 \times J_1(0) + 0.7[3 \times 1 + J_1(0)] + 0.2[3 \times 2 + J_1(0)] = 6.6;$$

$$u_0 = 1: E\{\cdot\} = 1 + 0.1[1 + J_1(1)] + 0.7 \times J_1(0) + 0.2[3 \times 1 + J_1(0)] = 4.9;$$

$$u_0 = 2: E\{\cdot\} = 2 + 0.1[2 + J_1(2)] + 0.7[1 + J_1(1)] + 0.2 \times J_1(0) = 5.352;$$

所以 $J_0(0) = 4.9$, $\mu_0^*(0) = 1$.

若初始状态未知, 则还要计算 $J_0(1)$ 和 $J_0(2)$, 不难验证这种计算可产生

$$\begin{aligned} J_0(1) &= 3.9, & \mu_0^*(1) &= 0 \\ J_0(2) &= 3.352, & \mu_0^*(2) &= 0. \end{aligned}$$

所以最优定货计划是存货为 0 时定货 1 单位, 否则不定货。

§5. 二次线性问题和必然等价原理

设有经济系统之线性状态方程

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \quad k = 0, 1, \dots, N-1. \quad (5-85)$$

我们要求解控制规律 $\{\mu_0^*(x_0), \dots, \mu_{N-1}^*(x_{N-1})\}$ 使二次费用函数最小化:

$$E_{w_k, u_k, N-1} \left\{ x_N' Q_N x_N + \sum_{k=0}^{N-1} (x_k' Q_k x_k + u_k' R_k u_k) \right\} \quad (5-86)$$

其中 x_k 为 n 维矢量, u_k 为 m 维矢量; 矩阵 A_k, B_k, R_k 有适当的维数; Q_k 为对称正半定矩阵, R_k 为对称正定矩阵; 干扰 w_k 为不依赖 x_k, u_k 而遵从给定概率分布之独立随机矢量。矢量 w_k 有 0 均值和有穷二阶矩, 对控制 u_k 无约束。

上述系统要求状态保持靠近原点, 这类二次线性问题是自动控制理论中常见的。二次费用函数对于与所欲状态的较大偏离有很大的惩罚, 而对较小的偏离则有相对小的惩罚。如果干扰 w_k 之均值不是 0, 则二次费用函数可写为更一般的形式:

$$E\{(x_N - \bar{x}_N)' Q_N (x_N - \bar{x}_N) + \sum_{k=0}^{N-1} [(x_k - \bar{x}_k)' Q_k (x_k - \bar{x}_k) + u_k' R_k u_k]\}, \quad (5-87)$$

这一控制系统要求状态保持靠近给定轨线 $(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_N)$ 。

研究(5-86)与(5-87)没有什么差别, 为使书写简明, 我们讨论(5-86)的形式, 读者不难将分析推广到(5-87)的形式。

运用动态规划, 我们有求解的递推算式:

$$J_N(x_N) = x_N' Q_N x_N \quad (5-88)$$

$$J_k(x_k) = \min_{u_k} E\{x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}(A_k x_k + B_k u_k + w_k)\} \quad (5-89)$$

这类问题产生的序贯费用函数 J_k 是二次型的且控制规律为状态之线性函数。将二次式(5-87), (5-89)对 $k = N-1$ 展开, 并利用 $E\{w_{N-1}\} = 0$, 消去

$$E\{w_{N-1}' Q_N (A_{N-1} x_{N-1} + B_{N-1} u_{N-1})\},$$

我们有

$$\begin{aligned} J_{N-1}(x_{N-1}) &= x_{N-1}' Q_{N-1} x_{N-1} + \min_{u_{N-1}} [u_{N-1}' R_{N-1} u_{N-1} + u_{N-1}' B_{N-1}' Q_N B_{N-1} u_{N-1} \\ &\quad + x_{N-1}' A_{N-1}' Q_N A_{N-1} x_{N-1} + 2x_{N-1}' A_{N-1}' Q_N B_{N-1} u_{N-1}] \end{aligned}$$

$$+ E\{w'_{N-1} Q_N w_{N-1}\} \quad (5-90)$$

对 u_{N-1} 求导数, 令其等于 0, 得到

$$(R_{N-1} + B'_{N-1} Q_N B_{N-1}) u_{N-1} = -B'_{N-1} Q_N A_{N-1} x_{N-1} \quad (5-91)$$

左边括弧内的矩阵乘上 u_{N-1} 为正定的, 且 $B'_{N-1} Q_N B_{N-1}$ 为正半定的, 于是最小化控制矢量为

$$u^*_{N-1} = -(R_{N-1} + B'_{N-1} Q_N B_{N-1})^{-1} B'_{N-1} Q_N A_{N-1} x_{N-1} \quad (5-92)$$

将上式代入(5-90), 我们有

$$J_{N-1}(x_{N-1}) = x'_{N-1} K_{N-1} x_{N-1} + E\{w'_{N-1} Q_N w_{N-1}\} \quad (5-93)$$

其中

$$K_{N-1} = A'_{N-1} [Q_N - Q_N B_{N-1} (B'_{N-1} Q_N B_{N-1} + R_{N-1})^{-1} B'_{N-1} Q_N] A_{N-1} + Q_{N-1}$$

矩阵 K_{N-1} 为对称正半定的。

我们可以用完全类似的方式从动态规划算式(5-89)推出 $N-2$ 阶段之最优控制, 以此类推, 我们可得到每一 k 阶段之最优控制规律

$$\mu^*_k(x_k) = L_k x_k \quad (5-94)$$

其中增益矩阵

$$L_k = -(B'_k K_{k+1} B_k + R_k)^{-1} B'_k K_{k+1} A_k \quad (5-95)$$

其中 K_k 由下列递推式给出:

$$K_N = Q_N \quad (5-96)$$

$$K_k = A'_k [K_{k+1} - K_{k+1} B_k (B'_k K_{k+1} B_k + R_k)^{-1} B'_k K_{k+1}] A_k + Q_k \quad (5-97)$$

费用函数之最小值为

$$J_0(x_0) = x'_0 K_0 x_0 + \sum_{k=0}^{N-1} E\{w'_k K_{k+1} w_k\} \quad (5-98)$$

方程(5-97)叫离散 Riccati 方程, 它在现代控制理论中起着十分重要的作用。不少文献对其性质进行了广泛而深入的研究。即当 A_k, B_k, Q_k, R_k 分别为常数矩阵 A, B, Q, R (A 为 $n \times n$ 矩阵, B 为 $n \times m$ 矩阵, A, B 可控, R, Q 可观, Q 为 $n \times n$ 对称正半定矩阵, R 为 $m \times m$ 对称正半定矩阵) 时, 则随着 $k \rightarrow \infty$, 解收敛于稳定的控制规律 $\{\mu^*, \mu^*, \dots, \mu^*\}$ 。根据这个性质, 若碰到线性系统

$$x_{k+1} = Ax_k + Bu_k + w_k, \quad k = 0, 1, \dots, N-1 \quad (5-99)$$

且 N 为一个很大的数, 则我们可以用下列稳定线性控制规律近似控制规律(5-94)

$$\mu^*(x) = Lx \quad (5-100)$$

$$L = -(B'KB + R)^{-1} B'KA \quad (5-101)$$

$$K = A'[K - KB(B'KB + R)^{-1}B'K]A + Q \quad (5-102)$$

根据以上分析, 我们可导出必然等价原理。

设 w_k 与 A, B 无关, 其概率分布给定且具有有穷二阶矩, 我们有动态规划问题:

$$J_N(x_N) = x_N' Q x_N \quad (5-103)$$

$$J_k(x_k) = \min_{u_k} E_{w_k \sim A, B} \{x_k' Q x_k + u_k' R u_k + J_{k+1}(Ax_k + Bu_k + w_k)\} \quad (5-104)$$

类似前述推导, 最优控制规律为

$$\mu_k^*(x_k) = L_k x_k \quad (5-105)$$

其中增益矩阵

$$L_k = -[R + E\{B'K_{k+1}B\}]^{-1}E\{B'K_{k+1}A_k\} \quad (5-106)$$

而 K_k 由下列递推方程给出

$$K_N = Q \quad (5-107)$$

$$K_k = E\{A'K_{k+1}A\} - E\{A'K_{k+1}B\}[R + E\{B'K_{k+1}B\}]^{-1} \\ \times E\{BK_{k+1}A\} + Q_k \quad (5-108)$$

不难看出最优控制 μ^* 与随机矢量 w 之概率分布无关, 仅依赖于 w 之均值, 因此随机控制问题变成了确定性控制问题。通过下面的例子, 更容易看清这一点。

设二次费用函数为

$$\min_w E\{(ax + bu + w)^2\} \quad (5-109)$$

其中 a, b 为给定常数, w 为随机变量, 相应的最优解为

$$u^* = -\frac{a}{b}x - \frac{1}{b}E\{w\} \quad (5-110)$$

u^* 与随机变量 w 之概率分布无关, 仅依赖于均值 $E\{w\}$, 结果最优化问题与已知 $E\{w\}$ 的确定性问题完全等价, 这就是**必然等价原理**。

与前类似的库存控制问题中, 设初始固定费用 $K = 0$, 系统状态方程为

$$x_{k+1} = x_k + u_k - w_k, \quad k = 0, 1, \dots, N-1 \quad (5-111)$$

假定序贯需求 w_k 有界且独立, N 阶段末未满足之需求是损失, 且 N 阶段末留下之库存为 0, 需求 $w_0, \dots, w_{N-1}$ 有不变均值。以下推导稍加改变, 则可以对于没有这些假定推导出同样的结论。

预期费用为

$$E_{w_k} \left\{ \sum_{k=0}^{N-1} [cu_k + p \max(0, w_k - x_k - u_k) + h \max(0, x_k + u_k - w_k)] \right\} \quad (5-112)$$

运用动态规划, 我们有

$$J_N(x_N) = 0 \quad (5-113)$$

$$J_k(x_k) = \min_{u_k \geq 0} [cu_k + L(x_k + u_k) + E\{J_{k+1}(x_k + u_k - w_k)\}] \quad (5-114)$$

其中

$$L(y) = pE\{\max(0, w_k - y)\} + hE\{\max(0, y - w_k)\}. \quad (5-115)$$

对于单一阶段而言, 时刻 $N-1$ 之最优计划为

$$\mu_{N-1}^*(x_{N-1}) = \begin{cases} s_{N-1} - x_{N-1}, & \text{当 } x_{N-1} < s_{N-1} \\ 0 & \text{当 } x_{N-1} \geq s_{N-1} \end{cases} \quad (5-116)$$

根据动态规划方程(5-114), 我们有

$$J_{N-1}(x_{N-1}) = \begin{cases} c(s_{N-1} - x_{N-1}) + L(s_{N-1}) \\ L_{N-1}(x_{N-1}) \end{cases} \quad (5-117)$$

由于 L 为凸的, 且 s_{N-1} 使 $cy + L(y)$ 最小化, 所以 (5-117) 也是凸的。若给定 J_N 为凸的, 我们也能证明 J_{N-1} 的凸性。

同样, 易知, 由于 $c < p$, 则

$$\lim_{|y| \rightarrow \infty} G_k(y) = \infty$$

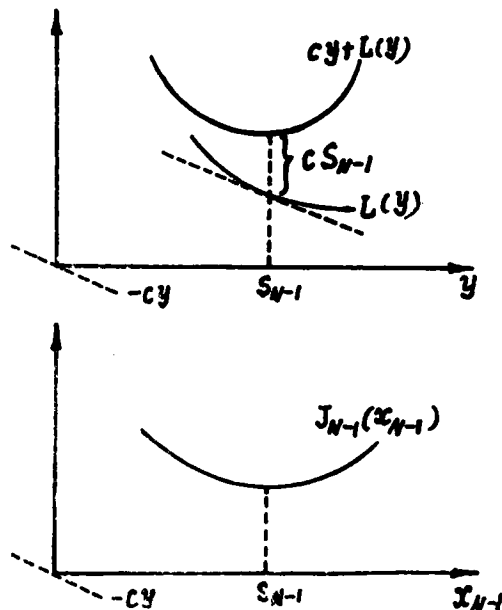


图 5-1

且我们有

$$J_k(x_k) = \begin{cases} c(s_k - x_k) + L(s_k) + E\{J_{k+1}(s_k - w_k)\}, & \text{当 } x_k < s_k, \\ L(x_k) + E\{J_{k+1}(x_k - w_k)\}, & \text{当 } x_k \geq s_k, \end{cases} \quad (5-118)$$

其中 s_k 使 $cy + L(y) + E\{J_{k+1}(y - w)\}$ 最小化。 J_{k+1} 之凸性 [意味着 $L(x) + E\{J_{k+1}(x - w)\}$ 之凸性] 足以用来证明 J_k 之凸性, 因此最优控制规律由数量序列 $\{S_0, S_1, \dots, S_{N-1}\}$ 确定且有如下形式:

$$\mu_k^*(x_k) = \begin{cases} s_k - x_k, & \text{当 } x_k < s_k \\ 0, & \text{当 } x_k \geq s_k \end{cases} \quad (5-119)$$

对于每一 k , 数量 s_k 使下式最小化:

$$G_k(y) = cy + L(y) + E\{J_{k+1}(y - w)\} \quad (5-120)$$

附录 A 函数取极大(小)值的充分条件, 函数的凹性、凸性与集合的凸性、凹性。

为导出函数取极大值的充分条件, 我们将函数

$$y = f(X) \quad (1)$$

在 X^* 展成泰勒级数, 其中 $X = (x_1, \dots, x_n)'$.

$$\begin{aligned} f(X) - f(X^*) &= df(X^*) + \frac{1}{2!} d^2f(X^*) + \frac{1}{3!} d^3f(X^*) \\ &+ \dots + \frac{1}{(n-1)!} d^{(n-1)}f(X^*) + R_n \end{aligned} \quad (2)$$

其中
$$df(X^*) = \frac{\partial f}{\partial x_1} dx_1 \Big|_{x_1^*} + \dots + \frac{\partial f}{\partial x_n} dx_n \Big|_{x_n^*}$$

令
$$\frac{\partial f}{\partial x_i \partial x_j} = f_{ij}, \quad dX = [dx_1, \dots, dx_n]'$$

则
$$d^2f(X^*) = [dX]' \begin{bmatrix} f_{11} & \dots & f_{1n} \\ \dots & \dots & \dots \\ f_{n1} & \dots & f_{nn} \end{bmatrix} [dX]. \quad (3)$$

取极值的必要条件为

$$df(X^*) = 0 \quad (4)$$

由于 $dx_1, \dots, dx_n$ 可取任意不为 0 的微小数值 (可正可负), 所以一阶全微分等于 0 意味着

$$\frac{\partial f}{\partial x_i} = 0, \quad i = 1, \dots, n, \quad (5)$$

取极大值的充分条件为 (2) 左边的

$$f(X) - f(X^*) < 0 \quad (6)$$

$$\therefore df(X^*) = 0$$

$\therefore$ (6) 在二阶近似的条件下意味着

$$d^2f(X^*) < 0 \quad (7)$$

即

$$[dX]' [H] [dX] < 0 \quad (8)$$

其中

$$[H] = \begin{bmatrix} f_{11} & \dots & f_{1n} \\ \dots & \dots & \dots \\ f_{n1} & \dots & f_{nn} \end{bmatrix}$$

$[H]$ 叫 Hessian 矩阵, 由于 $f_{ij} = f_{ji}$, 所以 $[H]$ 为对称矩阵。

由于(3)是二次式, 所以(8)意味着要求二次式之值小于0, 我们先看最简单的二次式:

$$\begin{aligned} & f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2 \\ &= f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + \frac{f_{12}^2}{f_{11}}dx_1^2 + f_{22}dx_2^2 - \frac{f_{12}^2}{f_{11}}dx_1^2 \\ &= f_{11}\left(dx_1 + \frac{f_{12}}{f_{11}}dx_2\right)^2 + \left(\frac{f_{11}f_{22} - f_{12}^2}{f_{11}}\right)dx_2^2 \end{aligned} \quad (9)$$

也就是说我们可以用满秩线性变换将一般二次型化成正规型, 二次式变成了完全平方和的形式, 其值的符号由系数

$$f_{11} \text{ 和 } \frac{f_{11}f_{22} - f_{12}^2}{f_{11}}$$

确定。当 $f_{11} < 0$, $f_{11}f_{22} - f_{12}^2 > 0$ 时, 二次式(9)之值小于0。而 f_{11} 正是函数 $f(x_1, x_2)$ 的 Hessian 矩阵一阶行列式之值, $f_{11}f_{22} - f_{12}^2$ 又正是 $f(x_1, x_2)$ 之 Hessian 矩阵二阶行列式之值。

以上推导可扩展到

$$y(x_1, x_2, \dots, x_n)$$

的情形, 令 $H_1, H_2, \dots, H_n$ 为 Hessian 矩阵各阶子矩阵, 则不难证明

$$\begin{aligned} & [dX]'[H][dX] \\ &= |H_1|(\cdot)^2 + \frac{|H_2|}{|H_1|}(\cdot)^2 + \frac{|H_3|}{|H_2|}(\cdot)^2 + \dots + \frac{|H_n|}{|H_{n-1}|}(\cdot)^2 \end{aligned} \quad (10)$$

所以当 Hessian 矩阵各阶行列式之符号正负相间且一阶行列式为负时, 二次式之值恒小于0, 这时我们定义 Hessian 矩阵负定, 若二次式之值恒小于或等于0, 我们定义 Hessian 矩阵半负定。因此极大化的充分条件为函数之 Hessian 矩阵负定。

同理, 若要求

$$y = f(X)$$

极小化, 则

$$f(X) - f(X^*) > 0 \quad (11)$$

$$\therefore df(X^*) = 0$$

$$\therefore d^2f(X^*) > 0 \quad (12)$$

$$\therefore [dX]'[H][dX] > 0 \quad (13)$$

根据(10), 当 H 之各阶子行列式之值的符号都大于或等于0时, (13)成立。我们定义 Hessian 各阶子行列式之值恒大于0为 Hessian 矩阵正定, 而各阶子行列式之值恒大于或小于0为 Hessian 矩阵半正定。因此函数在 X^* 取极小值的充分条件为其 Hessian 矩阵正定。

下面我们讨论凸集合、凹集合与函数的凹性、凸性以及函数取极值的充分条件之间的

关系。

设有集合 $X', X'' \in S$, 则当

$$X''' = (1-a)X' + aX'', \quad 0 < a < 1 \quad (14)$$

亦使 $X''' \in S$, 则 n 维欧氏空间 R^n 中的点集合 S 是凸性的。如果 S 是凸性的, X', X'', X''' 为此集合中的点, 若

$$f(X''') \geq \min [f(X'), f(X'')] \quad (15)$$

则函数 $f(X)$ 为准凹函数。

设所有 X 中的一部分构成一集合 L_b , 使得

$$f(X) \geq b \quad (16)$$

则只有当 $f(X)$ 为准凹性时, L_b 才是凸的。

证明: (i) 必要条件:

令 $X', X'' \in L_b$, 同时令

$$b = \min [f(X'), f(X'')] \quad (17)$$

则 X' 和 $X'' \in L_b$ 。 $\because L_b$ 为凸的, 则其凸性组合

$$X''' = (1-a)X' + aX'', \quad 0 < a < 1, \quad (18)$$

亦属于 L_b , $\therefore$

$$f(X''') \geq b. \quad (19)$$

根据(17), (19), f 为准凹性的。

(ii) 充分条件

令 L_b 是使 $f(X) \geq b$ 之 X 的集合, 若 $X', X'' \in L_b$, 则

$$f(X') \geq b, \quad f(X'') \geq b.$$

$\therefore f(X)$ 为准凹函数,

$$\therefore f(X''') \geq b \quad (20)$$

$$\therefore X''' \in L_b$$

$$\therefore L_b \text{ 为凸的}$$

函数准凹性的必要和充分条件可以合并叙述如下: 如果 n 维向量 X 之可微分的函数 $f(X)$ 是准凹性的, 则

$$(-1)^k \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1k} & f_1 \\ f_{21} & f_{22} & \cdots & f_{2k} & f_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{k1} & f_{k2} & \cdots & f_{kk} & f_k \\ f_1 & f_2 & \cdots & f_k & 0 \end{vmatrix} \geq 0 \quad k = 2, \dots, n. \quad (21)$$

如果(21)是严格的不等式,则函数是严格准凹性的。如果函数是严格准凹性的,则在约束条件下极大化问题之二阶条件可获得满足。

关于函数满足凹性的条件有如下定理:只有当下式成立时,函数才是凹性的:

$$f[(1-a)X' + aX''] \geq (1-a)f(X') + af(X'') \quad (22)$$

其中 X', X'' 为凸集合的任意两点,且 $0 < a < 1$ 。即 X', X'' 的函数值用 $(1-a), a$ 加权平均得到的线性插值不大于 X', X'' 加权平均值之函数值。显然这种凹性函数才存在极大值。

对于一元函数来说,这意味着函数的斜率是递减的,即二阶导数小于 0。

如果(22)中只有不等号成立,则 $f(X)$ 是严格凹性的,对于线性函数求来说,它是凹性的,但不一定是严格凹性的。而曲线则可能是严格凹性的。

假定 X' 和 X'' 都属于凸性集合,只有当下式成立时,可微函数 $f(X)$ 才是严格凹性的。

$$f(X'') - f(X') < \sum_i \frac{\partial f}{\partial x_i} (x'_i - x''_i) \quad (23)$$

其中偏导数在 X' 点取值。

必要条件可证明如下:

$$(1-a)X' + aX'' = X' + a(X'' - X') \quad (24)$$

$$\text{且} \quad (1-a)f(X') + af(X'') = f(X') + a[f(X'') - f(X')] \quad (25)$$

若 f 为严格凹的,则

$$f[X' + a(X'' - X')] > f(X') + a[f(X'') - f(X')] \quad (26)$$

利用泰勒定理,上式右边可展为一阶近似式

$$f(X') + a \sum_i \frac{\partial f}{\partial x_i} (x'_i - x''_i) \quad (27)$$

式中偏导数是在 X' 和 X'' 之间取值的

$$X' + ba(X'' - X'), \quad 0 \leq b \leq 1 \quad (28)$$

将(27)代入(26)左边,则

$$\sum_i \frac{\partial f}{\partial x_i} (x'_i - x''_i) > f(X'') - f(X'). \quad (29)$$

只有当 $f(x)$ 的 Hessian 矩阵是负定时,可微函数 $f(X)$ 才是严格凹性的。

这一定理的必要条件可证明如下。

利用泰勒定理, $f(X)$ 的二阶近似式可写为

$$f(X+V) = f(X) + \sum_i f_i v_i + \frac{1}{2} V' H V \quad (30)$$

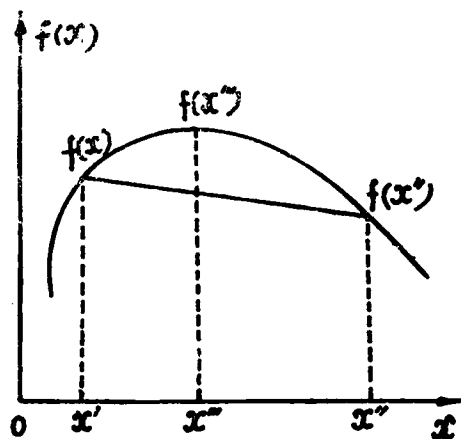


图 A

式中偏微分在 X 点取值, f 的 Hessian 矩阵 H 是在 X 和 $X+V$ 之间的一点取值的, 则

$$f(X+V) - f(X) - \sum_i f_i v_i = \frac{1}{2} V' H V \quad (31)$$

根据(23), 上式左边是负的, 因此 H 为负定的。

由此可见, 严格凹性函数能够满足极大化的充分条件。

附录 B

极大化之 Kuhn-Tucker 条件

假定 $f(X)$ 是一个含有三个变量的函数, 其中 $X = (x_1, x_2, x_3)$, 若要在

$$X \geq 0 \quad (1)$$

的约束下求 $f(X)$ 的极大值, 则分为如下三种情况。

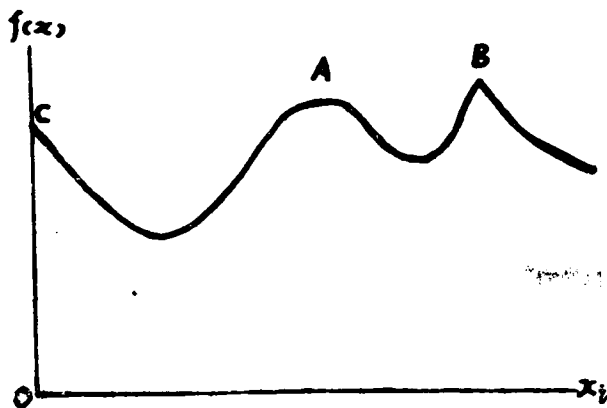


图 B

(i) 图中 A 点是在领域内部且在这一点函数是可微的, 则极大化的条件为

$$\frac{\partial f}{\partial x_i} = 0, \quad i = 1, 2, 3, \quad (2)$$

(ii) 图中 B 点虽在领域中, 但在此点函数是不可微的, 因为右边的微分 $\frac{\partial f^+}{\partial x_i}$ 和左边的微分 $\frac{\partial f^-}{\partial x_i}$ 不同。在这种情况下, 极大化的必要条件通常是

$$\frac{\partial f^+}{\partial x_i} \leq 0, \quad \frac{\partial f^-}{\partial x_i} \geq 0, \quad (3)$$

(iii) 图中 C 点在领域的边界上, 因此, 左边的微分不存在, 在这种情况下, 极大化的必要条件是

$$\frac{\partial f^+}{\partial x_i} \leq 0 \quad (4)$$

将(2)与(4)合并, 则当 f 可微时, 极大化之必要条件是

$$\frac{\partial f}{\partial x_i} = 0, \quad \text{当 } x_i > 0 \text{ 时} \quad (5)$$

$$\frac{\partial f}{\partial x_i} \leq 0, \quad x_i = 0 \text{ 时} \quad (6)$$

(5)与(6)可写成一个式子, 即

$$\frac{\partial f}{\partial x_i} \leq 0, \quad x_i \frac{\partial f}{\partial x_i} = 0 \quad (7)$$

如果我们加上不等式约束条件

$$g^j(X) \geq 0, \quad j = 1, 2, \quad (8)$$

则使 $f(X)$ 极大化问题可能有两种情况:

(i) 若 $g^j(X) = 0$, 则第 j 个约束条件有效, 亦 X 受第 j 个等式约束。

(ii) 若 $g^j(X) > 0$, 则第 j 个限制条件不起作用, 亦即 X 不受第 j 个等式约束, 即不等号成立。

在(i)种情况下,

$$g^j = 0, \text{ 相应的拉氏乘数 } \lambda_j \neq 0 \quad (9)$$

在(ii)种情况下

$$g^j > 0, \text{ 相应的拉氏乘数 } \lambda_j = 0 \quad (10)$$

我们可以将(9)与(10)合并为

$$g^j \geq 0, \quad \lambda_j g^j = 0 \quad (11)$$

对于以上在不等式约束下求非线性函数极大化的问题, 有著名的 Kuhn-Tucker 定理作为求解的基础, 这个定理是:

要求解 X^* 使 $f(X)$ 极大化, ($X = (x_1, \dots, x_m)$), 且服从约束

$$g^j(X) \geq 0, \quad j = 1, \dots, n \quad (12)$$

$$x_i \geq 0, \quad i = 1, \dots, m \quad (13)$$

则充要条件为, 存在非负 m 维矢量 $\lambda = (\lambda_1, \dots, \lambda_m)$ 使得当 $\lambda \geq 0, X \geq 0$ 时,

$$L(X, \lambda^*) \leq L(X^*, \lambda^*) \leq L(X^*, \lambda) \quad (14)$$

其中

$$L(X, \lambda) = f(X) + \sum \lambda_j g^j(X) \quad (15)$$

$L(X, \lambda)$ 叫做核 (kernel), (X^*, λ^*) 叫做鞍点。

鞍点之必要条件为:

要使一对矢量 (X, λ) 成为可微函数 $L(X, \lambda)$ 之鞍点, 其必要条件为

$$x_i > 0 \quad \text{及} \quad \frac{\partial L}{\partial x_i} = 0$$

或

$$x_i = 0 \quad \text{及} \quad \frac{\partial L}{\partial x_i} \leq 0,$$

合并可写作

$$-\frac{\partial L}{\partial x_i} \leq 0, \quad \text{及} \quad \sum_i \frac{\partial L}{\partial x_i} x_i = 0 \quad (16)$$

且要求

$$\lambda_j > 0 \quad \text{及} \quad \frac{\partial L}{\partial \lambda_j} = 0$$

或

$$\lambda_j = 0 \quad \text{及} \quad \frac{\partial L}{\partial \lambda_j} \geq 0$$

合并可写作

$$\frac{\partial L}{\partial \lambda_j} \geq 0, \quad \text{及} \quad \sum_j \lambda_j \frac{\partial L}{\partial \lambda_j} = 0. \quad (17)$$

鞍点之充分条件为:

要使 (X, λ) 成为可微函数 $L(X^*, \lambda^*)$ 之鞍点, 其充分条件是(16), (17)必须成立,

且

$$L(X, \lambda^*) \leq L(X^*, \lambda^*) + \sum_i \frac{\partial L}{\partial x_i} (x_i - x_i^*). \quad (18)$$

$$L(X^*, \lambda) \geq L(X^*, \lambda^*) + \sum_j \frac{\partial L}{\partial \lambda_j} (\lambda_j - \lambda_j^*) \quad (19)$$

若 $\frac{\partial L}{\partial x_i}$ 在 X^* 取值 ($i = 1, \dots, m$), 则(16), (17), (18), (19)则构成了不等式约束下,

非线性函数取极大值的 Kuhn-Thcker 条件, 它是非线性规划的有力工具。而极大化问题等价于鞍点问题。

参 考 文 献

- [1] A. Bensoussan, etc.: *Management Applications of Modern Control Theory* 1974 North-Holland
- [2] D. P. Bertsekas: *Dynamic Programming and Stochastic Control* 1976
- [3] Gregory C. Chow: *Analysis and Control of Dynamic Economic Systems*, 1975, John Wiley Sons.
- [4] M. D. Intriligator: *Mathematical Optimization and Economic Theory*, 1971, Prentice-Hall
- [5] K. C. Kogiku: *An Introduction to Macroeconomic Models*
- [6] K. C. Kogiku: *Microeconomic Models*
- [7] R. E. Miller: *Dynamic Optimigation and Economic Applications*, 1979 McGraw-Hill
- [8] L. K. Timothy: *State Space Analysis: An Introduction* 1968
- [9] O. 兰格: 《经济控制论导论》(杨小凯译) 中国社会科学出版社 1981年