现代控制理论

张洪钺 主编

（第 二 册）

最优控制理论

孟宪仲 编

北京航空學院出版社

现 代 控 制 理 论

张 洪 钺　主编

（第二册）

最优控制理论

孟宪伟　编

北 京 航 空 学 院 出 版 社

内 容 简 介

本书比较全面地介绍了现代控制理论的基本内容，附有大量的例题和习题，切合工程实际，便于自学。全书分为四篇：第一篇线性系统理论，第二篇最优控制理论，第三篇最佳估计理论，第四篇系统辨识。分四册出版，每册一篇。

本书是第二册。主要介绍最优控制中的变分法、庞特里亚金的极大（小）原理、贝尔曼的动态规划以及二次型指标最优线性反馈控制等。

本书主要作为高等院校自动控制专业研究生课的教材，也可供从事自动控制工作的科技人员参考。

现代控制理论 第二册
——最优控制理论
张洪钺 主编 孟宪仲 编
责任编辑 胡益民 宋淑乔 严文璇
北京航空学院出版社出版
新华书店北京发行所发行 各地新华书店经售
昌平振兴胶印厂印装
787×1092 1/16 印张：8.5 字数：214千字
1987年12月第一版 1987年12月第一次印刷 印数：6000册
统一书号：15432·080 定价：1.5元
ISBN 7-81012-027-1/TP·005

目　　录

第二篇　最优控制理论

第二篇　最优控制理论

第一章　绪论

§1—1　最优控制发展簡史

第二次世界大战以后发展起来的自动调节原理，对设计与分析单输入单输出的线性时不变的系统是有效的；然而近代航空及空间技术的发展对控制精度提出了很高的要求，并且被控制的对象是多输入多输出的，参数是时变的。面临这些新的情况，建立在传递函数基础上的自动调节原理就日益显出它的局限性来。这种局限性首先表现在对于时变系统，传递函数根本无法定义，对多输入多输出系统从传递函数概念得出的工程结论往往难于应用。由于工程技术的需要，以状态空间概念为基础的最优控制理论渐渐发展起来。

最优控制理论所要解决的问题是：按照控制对象的动态特性，选择一个容许控制，使得被控对象按照技术要求运转，同时使性能指标达到最优值。从数学方面看，就是求解一类带有约束条件的泛函极值问题，因此这是一个变分学的问题；然而变分理论只是解决容许控制属于开集的一类最优控制问题，而在工程实践中还常遇到容许控制属于闭集的一类最优控制问题，这就要求人们研究新方法。在研究最优控制的方法中，有两种方法最富成效：一种是苏联学者庞特里雅金（*Понтрягин*、*Л.С*）提出的"极大值原理"；另一种是美国学者贝尔曼（*Bellman*，*R*）提出的"动态规划"。

极大值原理是庞特里雅金等人在1956至1958年间逐步创立的，先是推测出极大值原理的结论，随后又提供了一种证明方法。动态规划是贝尔曼在1953年至1958年间逐步创立的，他依据最优性原理发展了变分学中的哈密顿—雅可比（*Hamilton—Jacobi*）理论，构成了动态规划。

由于电子计算机技术的发展，使得设计计算和实时控制有了实际可用的计算工具，为实际应用一些更完善的数学方法提供了工程实现的物质条件，高速度、大容量计算机的应用，一方面使控制理论的工程实现有了可能，另一方面又提出了许多需要解决的理论课题，因此这门学科目前是正在发展的，极其活跃的科学领域之一。

最优控制理论在一些大型的或复杂的控制系统设计中，已经取得了富有成效的实际应用。目前很多大学在自动控制理论课程中已经开始适当增加这方面的内容，而对于自动控制方面的研究生则普遍作为必修课程。

§1—2　最优控制問题的实例

例1—1　月球上的软着陆问题

为了使宇宙飞船在月球表面实现软着陆，要寻求发动机推力的最优控制规律，以便使燃料的消耗为最少。设飞船质量为$m(t)$，高度为$h(t)$，垂直速度为$v(t)$，发动机推力为$u(t)$，月球表面的重力加速度为常数g。设不带燃料的飞船质量为M，初始燃料的总质量为F，初始高度为h_0，初始的垂直速度为v_0，那么飞船的运动方程式可以表示为：

$$\begin{cases} \dot{h}(t)=v(t) \\ \dot{v}(t)=-g+\dfrac{u(t)}{m(t)} \\ \dot{m}(t)=-ku(t) \end{cases} \tag{1—1}$$

上面运动方程式中的k是比例常数。

初始条件

$$\begin{aligned} h(0)&=h_0 \\ v(0)&=v_0 \\ m(0)&=M+F \end{aligned} \tag{1—2}$$

终端条件

$$\begin{aligned} h(t_f)&=0 \\ v(t_f)&=0 \end{aligned} \tag{1—3}$$

约束条件　发动机最大推力为α

$$0\leqslant u(t)\leqslant \alpha \tag{1—4}$$

性能指标是使燃料消耗为最小，也就是使飞船在着陆时的质量为最大，即

$$J=m(t_f) \tag{1—5}$$

达到最大值。

我们的任务是寻求发动机推力的最优控制规律$\hat{u}(t)$，它应满足约束条件（1—4），能使飞船由初始状态（1—2）转移到终端状态（1—3），并且使性能指标（1—5）为极大。

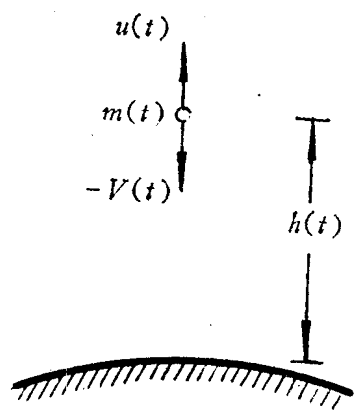

图1—1　月球上的軟着陆

例1—2　拦截问题

设发射一枚拦截器L，其任务是在空中拦截目标M，现在的问题是应当怎样控制拦截器L的运动，才能最快地击中目标。

设拦截器和目标的运动是在同一平面内，又设目标M的运动方程是：

$$\begin{cases} \dot{x}_M(t)=v_{XM}(t) \\ \dot{v}_{XM}(t)=0 \\ \dot{y}_M(t)=v_{yM}(t) \\ \dot{v}_{yM}(t)=-g \end{cases} \tag{1—6}$$

上面运动方程中：(x_M, y_M) 表示目标M的位置坐标，(v_{xM}, v_{yM}) 表示目标M的速度，g 表示重力加速度。

设拦截器的质量为$m(t)$，其发动机的有效喷气速度为C，假设为常数。用$f(t)$ 表示推力的大小，其质量变化率就等于推进剂消耗率，即为

$$\dot{m}(t) = -\frac{1}{C}f(t) \tag{1—7}$$

我们用(x_L, y_L)表示拦截器的坐标，用 (v_{xL}, v_{yL}) 表示其速度，用$\varphi(t)$ 表示拦截器的推力方向与水平线的夹角，如图1—2所示，那么拦截器的运动方程是

$$\begin{cases} \dot{x}_L(t) = v_{XL}(t) \\ \dot{y}_L(t) = v_{yL}(t) \\ \dot{v}_{XL}(t) = \dfrac{f(t)}{m(t)}\cos\varphi(t) \\ \dot{v}_{yL}(t) = \dfrac{f(t)}{m(t)}\sin\varphi(t) - g \\ \dot{m}(t) = -\dfrac{1}{C}f(t) \end{cases} \tag{1—8}$$

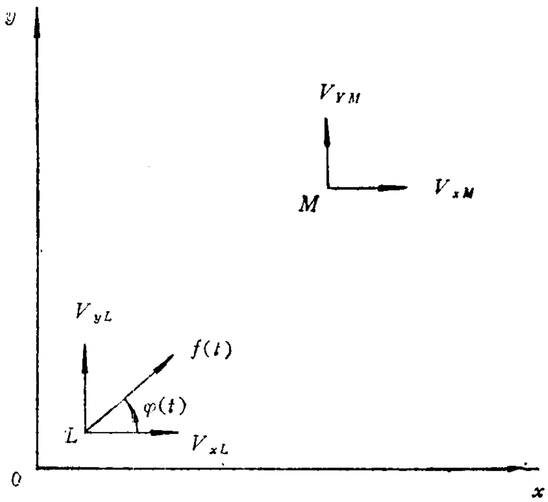

图1—2　拦截问题

为了简化数学表达式，我们取相对运动坐标，使用下列相对运动参数

$$x(t) = x_L(t) - x_M(t)$$
$$y(t) = y_L(t) - y_M(t)$$
$$v_x(t) = v_{xL}(t) - v_{xM}(t)$$
$$v_y(t) = v_{yL}(t) - v_{yM}(t)$$

这样方程（1—6）及（1—8）就合并成下列相对运动方程

$$\begin{cases} \dot{x}(t) = v_x(t) \\ \dot{y}(t) = v_y(t) \\ \dot{v}_x(t) = \dfrac{f(t)}{m(t)}\cos\varphi(t) \\ \dot{v}_y(t) = \dfrac{f(t)}{m(t)}\sin\varphi(t) \\ \dot{m}(t) = -\dfrac{1}{C}f(t) \end{cases} \tag{1—9}$$

推力$f(t)$不可能超出发动机的最大推力F，即$f(t)$ 应满足下列约束条件

$$|f(t)| \leqslant F \tag{I—10}$$

拦截器携带的燃料也是有限制的，用m_0表示拦截器满载燃料时的质量，用m_E表示燃料消耗完以后的质量。

对于最快拦截问题可归纳如下

已知动态系统的运动方程（1—9），及其初始条件

$$
\begin{aligned}
&\text{在}t=t_0\text{时} \qquad x(t_0)=x(0)\\
&y(t_0)=y(0)\\
&v_x(t_0)=v_x(0)\\
&v_y(t_0)=v_y(0)\\
&m(t_0)=m_0
\end{aligned}
\tag{1—11}
$$

现在的任务是寻求拦截器的推力$f(t)$的控制规律和推力方向角$\varphi(t)$的控制规律，以使拦截器从发射到击中目标M的时间最短，即

$$
J(f,\ \varphi)=\int_0^{t_f} dt=\hat{t}_f \tag{1—12}
$$

最小，也就是到达终端状态 $x(\hat{t}_f)=0$，$y(\hat{t}_f)=0$

所需的时间最短。

并满足约束条件（1—10）和下列条件

$$
m(\hat{t}_f)\geqslant m_E
$$

§1—3 最优控制問題的提法

纵观以上最优控制问题的实例，可以将最优控制问题的提法叙述如下

设已知受控系统的状态方程

$$
\dot{\mathbf{X}}(t)=f[\mathbf{X}(t),\ \mathbf{u}(t),\ t]
$$

初始条件 $\qquad \mathbf{X}(t_0)=\mathbf{X}_0$

其中 $\qquad \mathbf{X}(t)\in R^n\ t\in[t_0,\ t_f]$

$$
\mathbf{u}(t)\in\Omega\subseteq R^m
$$

向量函数$f[\mathbf{X}(t),\ \mathbf{u}(t),t]$是$\mathbf{X}(t)$,$\mathbf{u}(t)$和$t$的连续函数，并对$\mathbf{X}(t)$和$t$连续可微。

现在要寻找在区间$[t_0,\ t_f]$中，分段连续的最优控制函数$u(t)$，以便把状态变量$\mathbf{X}(t)$从初始状态$\mathbf{X}_0$转移到终端状态$\mathbf{X}(t_f)$，使$\mathbf{X}(t_f)\in S$并使下列性能指标

$$
J(\mathbf{u})=\phi[\mathbf{X}(t_f),\ t_f]+\int_{t_0}^{t_f} L[\mathbf{X}(t),\ \mathbf{u}(t),\ t]dt
$$

达到极值。

其中$\phi[X(t_f),\ t_f]$称为末值型性能指标，$\int_{t_0}^{t_f} L[\mathbf{X}(t),\ \mathbf{u}(t),\ t]dt$称为积分型性能指标，$\phi$和$L$都是$\mathbf{X}(t)$和$t$的连续可微函数。

综上所述，可见最优控制问题都包括以下四个部分：

1. 动态系统的状态方程

状态方程反映了受控动态系统的运动规律，一般写成向量微分方程的形式，即

$$
\dot{\mathbf{x}}(t)=f[\mathbf{X}(t),\ \mathbf{u}(t),\ t]
$$

其中$\mathbf{X}(t)$表示n维状态向量。

$\mathbf{u}(t)$表示m维控制向量。

2. 状态方程的边界条件

动态系统的运动过程就是在状态空间中，从一个状态到另一个状态的转移，如果把这种转移看成是n维状态空间中点的运动，那么一个动态过程就对应于状态空间中的一条轨线。在初始时刻的初始状态通常是已知的，即$\mathbf{X}(t_0)=\mathbf{X}_0$，而到达终端的时间$t_f$和终端状态$\mathbf{X}(t_f)$则有不同的情况；终端时间$t_f$可以有两种情形：一种是固定的，另一种是自由变动的。终端状态$\mathbf{X}(t_f)$可以是状态空间中一个固定的点，也可以是状态空间中一个变动的点，还可以是$\mathbf{X}(t_f)$中有些分量固定，另一些分量自由，但是无论哪一种情况，都可以用一个目标集加以概括。目标集用S来表示，就是：

$$\mathbf{X}(t_f)\in S$$

如果终端状态不受约束，则可看作是目标集扩展到整个n维空间。

3. 容许控制

对于每一个控制问题来说，控制变量$\mathbf{u}(t)$有一个取值范围，这个取值范围对应于m维控制向量空间$\mathbf{R}^m$中的一个集合Ω，于是$\mathbf{u}(t)$的每一个值对应于集合Ω中的一个元素，凡属集合Ω的控制，称为容许控制。Ω称为控制域。

控制变量中有一类是其变化范围受限制的，例如发动机推力$0\leqslant u(t)\leqslant a$，另一类是其变化范围不受限制的，例如推力方向角$\varphi(t)$可以作360°的变化而不受限制。前一类属于某一闭集，后一类属于某一开集。最优控制一定应当是容许控制，即

$$\mathbf{u}(t)\in\Omega\subseteq\mathbf{R}^m$$

4. 性能指标

在状态空间中，要使状态向量由初始状态$\mathbf{X}(t_0)$转移到终端状态$\mathbf{X}(t_f)$，可以用不同的控制函数来实现，为了衡量控制系统在每一个控制函数作用下工作的好坏，就需要用一个性能指标来判断。性能指标的内容取决于最优控制问题所要完成的任务，因此不同的控制问题就有各不相同的性能指标，为最优控制问题确定恰当的性能指标不仅需要理论知识而且需要实践经验。

在有些参考文献中对性能指标也称为：性能泛函、目标函数、评价函数、价值函数。

第二章　最优控制中的变分法

最优控制问题是在一定的约束条件下，寻求使性能指标达到极大值（或极小值）时的控制函数。当被控制对象的运动特性由向量微分方程来描述，性能指标由泛函来描述时，这就是微分方程约束下求泛函的条件极值问题。研究泛函极值问题的数学方法—变分法，由于近代数学的进展，已经发展成一个完整的数学分支，因此变分法是我们研究最优控制问题的一个重要工具。

在这一章的讨论中，我们假定容许控制的取值集合——控制域Ω是m维空间中的开集。

§2—1　赋范綫性空間

我们首先讨论一个基本概念：距离和距离空间，进一步再引出赋范空间和内积空间。

控制系统的状态可以由观测决定，而观测所得的值总是近似的；我们要求近似值能任意逼近准确值，常常就要在状态之间定出距离量度，以便确切地规定出“任意逼近”的概念。如果把控制系统的状态看作是空间中的点，距离就是表达两点之间远近的一个数。

平时我们谈到两点之间的距离，总是非负的，并且只有当两点重合时，才说两点间的距离是零。谈到两点间的距离时与其顺序无关，比如说：由甲地到乙地是100公里，那么也可以说：由乙地到甲地是100公里。如果有三点A、B、C，那么由两点之间直线段为最短引伸出来的三角形不等式，即三角形两边之和大于第三边，也是距离概念的一个基本属性。

我们可以概括出距离空间的确切定义如下

定义2—1　设X是一个不空的集合，X叫做距离空间（*Metric Space*），是指在X中定义一个双变量的实函数$d(x, y)$，满足下列条件：

(1)　$d(x, y)\geqslant 0$　　$x\in$ X，　$y\in$ X

当且仅当$x=y$时，$d(x, y)=0$

(2)　$d(x, y)=d(y, x)$

(3)　$d(x, y)\leqslant d(x, z)+d(z, y)$，$x, y, z\in$ X

我们可以称$d(x, y)$为“x与y间的距离”，也叫做X上的距离。引进距离概念的主要目的在于刻画“任意逼近”的概念，直观上说，动点x无限逼近固定点x_0，是指距离$d(x, x_0)$趋近于0。

定义2—2　距离空间X中的点列x_n叫做收敛于x_0，或称点列x_n以点x_0为极限，是指

$$d(x_n, x_0)\longrightarrow 0 \qquad (n\to\infty)$$

这时我们写成

$$\lim_{n\to\infty} x_n=x_0 \qquad (n\to\infty)$$

距离空间中点列的收敛，有数学分析中所讨论的那种收敛所具有的属性。

定义2—3　设x_n是距离空间X中的点列，如果对于任意的$\varepsilon>0$皆存在一个足够大的数

N，当m，$n>N$时有

$$d(x_m, x_n)<\varepsilon$$

则称序列x_n为距离空间X的基本序列或称为柯西序列（Cauchy Sequence）。

如果距离空间X的每一个基本序列都收敛于X中的一点，则称距离空间X是完备的距离空间（Complete Metric Space）。

在一距离空间X中，点集

$$S(x_0;r)=\{x\in X|d(x, x_0)<r, r>0\}$$

叫做以点x_0为中心，以r为半径的开球。

如果在上边的公式中，把<换成≤，则相应的点集叫做闭球，记成$\bar{S}(x_0;r)$。

$$\sum(x_0;r)=\{x\in X|d(x, x_0)=r\}$$

叫做以x_0为中心，半径为r的球面。据此引出开集的定义如下：

定义2—4 设M是距离空间X中的子集，点$x\in X$叫做集M的内点，是指有一以x为中心的开球$S(x;\varepsilon)\subset M$。M的所有内点的全体叫做M的内部，记作$Int\,M$。如果M的一切点都是它自己的内点，即$M=Int\,M$，那么M叫做开集。

定义2—5 设M是距离空间X的子集，点$x\in X$叫做M的聚点，是指以x为中心的任何开球中都含有M中一个不等于x的点：

$$[S(x;\varepsilon)\backslash\{x\}]\cap M\neq\phi, \quad \forall\varepsilon>0$$

点$x\in X$叫做M的附着点，是指

$$S(x;\varepsilon)\cap M\neq\phi, \qquad \forall\varepsilon>0$$

M的所有附着点的集叫做M的闭包，记成$\overline{M}$。

距离空间X中的集M叫做闭集，指$M=\overline{M}$。

定义2—6 我们用x表示n维列向量，用R^n表示n维列向量的全体，如果规定经过向量加法和向量乘常数这两种运算之后，仍在R^n之内，那么R^n即称为n维线性空间（Linear Space）。

定义2—7 如果线性空间R^n中每一个元素x都有范数$\|x\|$，并且满足下列范数三公理：

(1) $\|x\|\geqslant 0$ 　　当且仅当$x=0$时$\|x\|=0$

(2) $\|x+y\|\leqslant\|x\|+\|y\|$

(3) $\|\alpha x\|=|\alpha|\cdot\|x\|$ 　　α为任意常数

则称R^n为赋范线性空间（Normed Linear Space）。

定义2—8 在n维线性空间R^n中，两个向量

$$x=[x_1, x_2, \cdots\cdots, x_n]^T$$
$$y=[y_1, y_2\cdots\cdots, y_n]^T$$

我们把它们的数积定义为内积

$$\langle x, y\rangle=\sum_{i=1}^{n}x_iy_i$$

用内积定义范数

$$\|x\|=\langle x, x\rangle^{\frac{1}{2}}=\sqrt{\sum_{i=1}^{n}x_i^2}$$

再用范数定义距离

$$d(x, y)=\|x-y\|=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$$

则 R^n 也称为希尔伯特空间（$Hilbert\ \ Space$）。

现在我们再讨论控制向量函数空间，设有 r 个函数：

$$u_1(t),\ u_2(t),\ \cdots u_r(t)$$

在时间闭区间 $t_0 \leqslant t \leqslant t_f$ 上定义，对于每一个固定的 $t\in[t_0, t_f]$，

$$U(t)=[u_1(t),\ u_2(t),\ \cdots,\ u_r(t)]^T$$

是 r 维空间 R^r 中的一个点。

定义2—9 设 L 表示在 r 维空间中取值并且平方可积的函数组 $U(t)$ 的全体组成的集合，在空间 L 中定义内积：

$$\langle U(t),\ V(t)\rangle=\int_{t_0}^{t_f}U^T(t)V(t)dt=\int_{t_0}^{t_f}\sum_{i=1}^{r}u_i(t)v_i(t)dt$$

用内积定义范数

$$\|U(t)\|=\left[\int_{t_0}^{t_f}\sum_{i=1}^{r}u_i^2(t)dt\right]^{\frac{1}{2}}$$

我们称 L 为控制向量函数空间。

$U(t)$ 是 L 空间中的向量函数，记为 $U(t)\in L$。

在向量函数空间中，我们可以用范数定义距离：

$$d(U,\ V)=\|U(t)-V(t)\|=\left[\int_{t_0}^{t_f}\sum_{i=1}^{r}(u_i-v_i)^2dt\right]^{\frac{1}{2}}$$

所以 L 空间也是赋范线性空间。

定义2—10 向量函数空间 L 中，所谓向量函数序列 U_n 收敛于向量函数 U_0。或者以 U_0 为其极限，是指对于任意指定的实数 $\varepsilon>0$，都存在一个充分大的自然数 N，使得 $m>N$ 时，都有

$$d(U_m,\ U_0)<\varepsilon$$

或

$$\lim_{m\to\infty} U_m=U_0$$

所以向量函数空间 L 也是完备的赋范线性空间。

§2—2 綫性算子及泛函

在最优控制问题中的性能泛函，可以看作是赋范线性空间中某个子集到实数集上的算子。在这一节中我们从线性算子的定义谈到泛函及其变分，以及有关的基本定义和定理。

当讨论控制系统时，系统的状态可以用有限维空间中的向量表示；而为了表达这种系统的运动，就要用到向量之间的依赖关系。我们考察系统的运动时，常把系统在某时刻 t 的状态 y 和在某一初始时刻 t_0 的状态 x 加以比较 这个在 t 时刻的状态 y 可以看作是由初始状态 x 决定的，y 对于 x 的这种依赖关系可以看作是由状态空间 R^n 到它自己之中的变换。

$$\mathbf{y}=J(\mathbf{x})$$

例如当考察飞机在空中飞行时，如果只注意它的方位的变化而忽略它的重心移动，我们常在飞机的机身上，固定一个直交笛卡儿坐标系，例如 ξ 轴沿飞机 机 身的纵轴指向前方，η轴沿飞机的右机翼，指向右方，而ζ轴沿垂直方向指向下方，这样 坐在驾 驶员的位置看上去，正好形成一个右手坐标系。这时飞机的运动可以用在时刻t固定在机身上的坐标系相对于初始时刻t_0固定在机身上的坐标系的方位来表达，这也可以看作是三维空间中的旋转。在这种情况下，$\mathbf{x}$，$\mathbf{y}$ 都是三维向量，而J是三维空间的旋转。

又如在电网络系统中，输入信号$r(t)$与输出信号$c(t)$都是依赖于时间的函数，而系统的作用则是把$r(t)$变换成$c(t)$

$$c=J(r)$$

对于控制系统本身，也可以用一个算子来描述， 这个算子把输入信号 $\mathbf{u}(t)$变成输出信号$\mathbf{y}(t)$

$$\mathbf{y}=J(\mathbf{u})$$

在近似情况下，这种算子常常认为是线性的。 例如波的散射问题： 当入射波是两个波的迭加，即

$$J(\varphi_1+\varphi_2)=J(\varphi_1)+J(\varphi_2)$$

又当入射波的振幅增大λ倍时，出射波的振幅也增大同样倍数，即

$$J(\lambda\varphi)=\lambda J(\varphi)$$

现在我们可以写出线性算子的定义如下。

定义2—11 n维线性空间R^n到m维线性空间 R^m 的算子，是指有一个确定的对应规律

$$\mathbf{y}=J(\mathbf{x})$$

它使每一个 $\mathbf{x}\in R^n$对应一个 $\mathbf{y}\in R^m$

如果这个对应关系满足下列线性条件：

(1) $J(x_1+x_2)=J(x_1)+J(x_2)$ $\qquad \forall x_1,\ x_2\in R^n$

(2) $J(\alpha x)=\alpha J(x)$ $\qquad \forall x\in R^n$

则称算子 J 为线性算子。

在许多问题中，只确定所用的函数空间还是不够的，因为实际所用的算子 J 不必对于空间中任意元都有意义；在所考察的空间R^n上使算子 J 作 用有意义的元的全体叫做 J 的定义域，记作$D(J)$，而$D(J)$在 J 作用下的集合

$$\{\,y\mid y=J(x),\ x\in D(J)\,\}$$

叫做 J 的值域，记作$Z(J)$。

我们还要引入连续算子的概念；由于我们对自然现象的认识总是近似的，因而往往需要判断当对输入的了解足够精确时， 对输出的了解是否也能足够精确； 实际上常有这样的情况：当输入状态相差足够小时，就可以保证输出状态相差小到所希望的程度。抽象出来说，当在某个赋范线性空间R^n中考察算子 J时，就要求当$x_n\to x_0$（$n\to\infty$）时能推出

$$J(x_n)\to J(x_0) \qquad (n\to\infty)$$

这样的算子叫做连续算子。

定义2—12 设 J 是赋范线性空间R^n的子集$D(J)$到赋范线性空间R^m中的算子，如果对于每一个收敛于$x_0\in D(J)$的序列$x_n\in D(J)$

皆有

$$\lim_{n \to \infty} J(x_n) = J(x_0)$$

则称算子J在x_0处连续。

若算子$J(x)$在子集$D(J)$中的每一点都连续，则称算子$J(x)$在$D(J)$中连续。

定义2—13 设$y=J(x)$是n维线性空间R^n中的子集D到m维线性空间R^m的算子，并且，在D中当由点X转到点X＋△X时，y就变成

$$y + \triangle y = J(x + \triangle x)$$

如果存在一个由D到R^m的线性算子k，使

$$J(x+\triangle x) - J(x) = k(\triangle x) + \theta(\|\triangle x\|)$$

上式中$\theta(\|\triangle x\|)$是$\|\triangle x\|$的高阶无穷小量，即

$$\lim_{\|\triangle x\| \to 0} \frac{\theta(\|\triangle x\|)}{\|\triangle x\|} = 0$$

则称线性算子k为$J(x)$在x处的微商。记为$k=J'(x)$

算子k本身也是一个由D到R^m的算子。

定义2—14 设J是赋范线性空间R^n中某个子集D到赋范线性空间R^m上的算子，其微商$J'(x)$在D内每一个点上都存在，如果$J'(x)$在$x=x_0$处的微商也存在，那么把它称做算子$J(x)$在$x=x_0$处的二阶微商。记为

$$(J'(x))' = J''(x)$$

用同样办法可以定义算子$J(x)$在x_0处的n阶微商。

$$J^{(n)}(x_0) = (J^{(n-1)}(x_0))'$$

如果算子$J(x)$在$x=x_0$处的n阶微商存在，则称算子$J(x)$在x_0处n次可微。

如果$x(t)$表示系统的状态，它的积分常具有确定的物理意义。积分是由此函数确定的数，或者说：积分乃是由某一函数空间到数域R中的算子，一般说，由距离线性空间R^n到数域R中的算子叫做R^n上的泛函。

平常的积分是一种简单的泛函，即它是线性的

$$\int_a^b [\alpha\phi(x) + \beta\psi(x)]dx = \alpha\int_a^b \varphi(x)dx + \beta\int_a^b \psi(x)dx$$

一般说，线性空间R^n上的泛函J叫做R^n上的线性泛函，指它满足下列条件：

$$J(\alpha x + \beta y) = \alpha J(x) + \beta J(y) \qquad \forall x, y \in R^n$$
$$\forall \alpha, \beta \in R$$

如果在$C[0, 1]$上考察积分，

$$J(x) = \int_0^1 x(t)dt$$

这个线性泛函$J(x)$是连续的。

一般情况下满足下列条件的线性泛函叫做连续的：

$$\|x_n - x\| \to 0 \ (n \to \infty) \quad \Rightarrow \quad \lim_{n \to \infty} J(x_n) = J(x)$$

实际上，这种泛函是最重要的，因为只有这种泛函，它在任何一点的值才能够用这点附近的泛函值任意逼近。在有穷维线性空间上，任何线性泛函都是连续的。

现在我们可以写出变分的定义如下：

定义2—15 如果赋范线性空间R^n上的泛函$J(x)$，作为算子在x_0处是可微的，则其微分$J'(x)\triangle x$称为泛函$J(x)$在$x=x_0$处的变分。记为$\delta J(x_0, \triangle x)$。
相应的可以定义二阶变分

$$\delta^2 J(x, \triangle x)=J''(x)(\triangle x)^2$$

也可以定义n阶变分

$$\delta^n J(x, \triangle x)=J^n(x)(\triangle x)^n$$

变分在泛函研究中所起的作用，正如微分在函数研究中所起的作用一样，对于函数的微分和泛函的变分，也可以写出另外一个几乎完全相当的定义。我们来考虑当x和$\triangle x$固定，而参变数ε之值改变时的$J(x+\varepsilon\triangle x)$的数值。当$\varepsilon=1$时我们得到增加后的函数值，$J(x+\triangle x)$，而当$\varepsilon=0$时，得到的是函数的初值$J(x)$。函数$J(x+\varepsilon\triangle x)$对$\varepsilon$的导函数在$\varepsilon=0$时就等于函数$J(x)$在点$x$处的微分。

$$\frac{\partial}{\partial\varepsilon}J(x+\varepsilon\triangle x)\Big|_{\varepsilon=0}=J'(x+\varepsilon\triangle x)\triangle x\Big|_{\varepsilon=0}=J'(x)\triangle x$$

对于多变量的函数$Z=f(x_1, x_2\cdots, x_n)$来说，也可用同样的办法来处理，即对ε求

$$f(x_1+\varepsilon\triangle x_1, x_2+\varepsilon\triangle x_2,\cdots,x_n+\varepsilon\triangle x_n)$$

的导函数，然后令$\varepsilon=0$，，而得到函数的微分。

$$\frac{\partial}{\partial\varepsilon}f(x_1+\varepsilon\triangle x_1, x_2+\varepsilon\triangle x_2, \cdots, x_n+\varepsilon\triangle x_n)\Big|_{\varepsilon=0}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\triangle x_i$$

对于泛函，也可以把变分定义为泛函$J(x+\varepsilon\triangle x)$对$\varepsilon$的导函数在$\varepsilon=0$时之值。我们可以说：泛函的变分就是泛函增量中的线性主部。也可以说：如果作为泛函增量中线性主部的变分存在的话，那么另一种意义下的变分，即参变数等于初值时泛函对参变数的导函数也必定存在，于是这两个定义就成为同一的了。

定理2—1 设$J(x)$是赋范线性空间R^n上的泛函，如果在$x=x_0$处$J(x)$是可微的，则其变分为

$$\delta J(x_0, \triangle x)=\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}$$

$$0\leqslant\varepsilon\leqslant1$$

如果在$x=x_0$处，$J(x)$是n次可微的，则其n阶变分为

$$\delta^n J(x_0, \triangle x)=\frac{\partial^n}{\partial\varepsilon^n}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}$$

证明：因为泛函$J(x)$在$x=x_0$处可微，所以

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}&=\lim_{\varepsilon\to0}\frac{J(x_0+\varepsilon\triangle x)-J(x_0)}{\varepsilon}\\&=\lim_{\varepsilon\to0}\frac{\delta J(x_0, \varepsilon\triangle x)+\theta(x_0, \varepsilon\triangle x)}{\varepsilon}\\&=\lim_{\varepsilon\to0}\frac{\varepsilon\delta J(x_0, \triangle x)}{\varepsilon}+\lim_{\varepsilon\to0}\frac{\theta(x_0,\varepsilon\triangle x)}{\varepsilon}\\&=\delta J(x_0, \triangle x)\end{aligned}$$

其中
$$\lim_{\varepsilon\to 0}\frac{\theta(x_0,\ \varepsilon\triangle x)}{\varepsilon}=0$$

如果$J(x)$在$x=x_0$处有二阶变分存在，则

$$\begin{aligned}\delta^2 J(x_0,\ \triangle x)&=J^{(2)}(x_0)(\triangle x)^2\\&=[J'(x_0)\triangle x]'\triangle x\\&=\left[\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}\right]'\triangle x\\&=\frac{\partial}{\partial\varepsilon}\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}\\&=\frac{\partial^2}{\partial\varepsilon^2}J(x_0+\varepsilon\Delta x)\Big|_{\varepsilon=0}\end{aligned}$$

同理，如果$x=x_0$处$J(x)$ n阶可微，则其n阶变分为

$$\delta^n J(x_0,\ \triangle x)=\frac{\partial^n}{\partial\varepsilon^n}J(x_0+\varepsilon\triangle x)\Big|_{\varepsilon=0}$$

§2—3 变分原理

变分原理是变分法的理论基础，我们现在讨论它的主要定义和定理，为今后讨论泛函的极值准备条件。

变分在泛函的研究中所起的作用，类似于微分在函数的研究中所起的作用。我们回忆一下微分的定义，如果函数的增量

$$\triangle f=f(x+\triangle x)-f(x)$$

可以表示为下列形式

$$\triangle f=A(x)\triangle x+\beta(x,\ \triangle x)\triangle x$$

此处$A(x)$不依赖于$\triangle x$，而当$\triangle x\to 0$时，$\beta(x,\ \triangle x)\to 0$，则说函数是可微的，将增量中与$\triangle x$有线性关系的那一部分，即$A(x)\triangle x$，叫做函数的微分，记为$df$，用$\triangle x$除$\triangle f$，并使$\triangle x\to 0$，取极限就得到

$$\lim_{\triangle x\to 0}\frac{\triangle f}{\triangle x}=A(x)=f'(x)$$
$$df=f'(x)\triangle x$$

我们可以说：微分是函数增量中的线性主部；变分则是泛函增量中的线性主部。

如果泛函的增量

$$\triangle J=J[x(t)+\delta x]-J[x(t)]$$

可以表示为下列形式

$$\triangle J=L[x(t)+\delta x]+\beta[x(t),\ \delta x]\,\|\delta x\|$$

此处$L[x(t)+\delta x]$对于δx来说是线性泛函，当$\|\delta x\|\to 0$时，$\beta[x(t),\ \delta x]\to 0$，那么在泛函增量中对于$\delta x$是线性的部分即$L[x(t),\ \delta x]$就叫做泛函的变分，记为$\delta J$。

我们再回忆一下多元函数微分的定义，设函数$f(x_1,x_2,\cdots,x_n)$具有连续的偏微商，则有

$$f(x_1+h_1,x_2+h_2,\cdots,x_n+h_n)-f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\frac{\partial f(x_1,\ x_2\cdots,\ x_n)}{\partial x_i}h_i+\varepsilon$$

此处ε是高阶无穷小 即ε与微变量的绝对值$|h_i|(i=1, 2\cdots n)$中的最大者比较起来是高阶无穷小量。

$\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}h_i$是微变量$h_i(i=1,2,\cdots n)$的线性函数，称为多元函数$f(x_1,x_2,\cdots x_n)$的微分。

对于标量函数的变分，也是其增量的线性主部。设标量函数$H=H[\mathbf{x}(t), \mathbf{u}(t)]$其中$x(t)$是$n$维向量，$u(t)$是$m$维向量。其增量形式是

$$\triangle H=H[\mathbf{x}+\delta\mathbf{x}, \mathbf{u}+\delta\mathbf{u}]-H[\mathbf{x}, \mathbf{u}]$$

其变分，即增量中的线性主部就是

$$\delta H=\sum_{i=1}^{n}\frac{\partial H}{\partial x_i}\delta x_i+\sum_{j=1}^{m}\frac{\partial H}{\partial u_j}\delta u_j=\left[\frac{\partial H}{\partial x}\right]^T\delta\mathbf{x}+\left[\frac{\partial H}{\partial u}\right]^T\delta\mathbf{u}$$

其中

$$\frac{\partial H}{\partial \mathbf{x}}=\begin{pmatrix}\frac{\partial H}{\partial x_1}\\ \frac{\partial H}{\partial x_2}\\ \vdots\\ \frac{\partial H}{\partial x_n}\end{pmatrix}=\nabla_x H \qquad \frac{\partial H}{\partial \mathbf{u}}=\begin{pmatrix}\frac{\partial H}{\partial u_1}\\ \frac{\partial H}{\partial u_2}\\ \vdots\\ \frac{\partial H}{\partial u_m}\end{pmatrix}=\nabla_u H$$

对于积分型泛函的变分

设泛函$J=\int_a^b F[\mathbf{x}(t), \dot{\mathbf{x}}(t), t]dt$

端点条件$\mathbf{x}(a)=x_0$，$\mathbf{x}(b)=x_1$ 取符号$\hat{\mathbf{x}}(t)$表示在上列端点条件下，使泛函J达到极小值的函数。

取

$$\delta\mathbf{x}(a)=0 \qquad \delta\mathbf{x}(b)=0$$

$$0\leqslant\varepsilon\leqslant1$$

$$\mathbf{x}(t)=\hat{\mathbf{x}}(t)+\varepsilon\delta\mathbf{x}(t)$$

$$\dot{\mathbf{x}}(t)=\hat{\dot{\mathbf{x}}}(t)+\varepsilon\delta\dot{\mathbf{x}}(t)$$

泛函的变分等于泛函对参变数ε的导函数在$\varepsilon=0$时的值，即

$$\delta J=\left[\frac{\partial J}{\partial\varepsilon}\right]_{\varepsilon=0}=\int_a^b\left[\frac{\partial F}{\partial x}\delta x(t)+\frac{\partial F}{\partial\dot{x}}\delta\dot{x}(t)\right]dt$$

对于向量微分方程的变分 设向量微分方程$\dot{x}(t)=f[\mathbf{x}(t), u(t)]$其展开式即

$$\begin{vmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \vdots\\ \dot{x}_n(t)\end{vmatrix}=\begin{vmatrix}f_1(x, u)\\ f_2(x, u)\\ \vdots\\ f_n(x, u)\end{vmatrix}$$