变分学

于慎根　编著

BIAN FEN XUE

南开大学出版社

变 分 学

于 慎 根 编著

南开大学出版社

变　分　学

于慎根　编著

南开大学出版社出版
（天津八里台南开大学校内）
新华书店天津发行所发行
河北工学院印刷厂印刷
（天津丁字沽河北工学院内）

1987年12月第1版　　1987年12月第1次印刷
开本：787×1092 1/32　印张：3.75
字数：78千　　　　　　印数：1-5000
ISBN7-310-00111-7/O·22
定价：0.66元

内 容 简 介

本书比较严格地叙述了变分法的原理与应用，共五章，对常用的几个非线性泛函的极值理论作了详细的叙述；对一般泛函的极值理论作了必要的介绍。可供高等院校数学物理专业师生、以工科研究生和科技工作者阅读。

序

泛函极值问题常称为变分问题，处理变分问题的方法又称为变分法．本书的目的是想比较严格地叙述变分法的基本原理，对几个常用的泛函的极值理论作详细的叙述，并对一般泛函的极值理论作初步的介绍．为此，叙述时使用了泛函分析的一些最基本的知识，有了它们才能明确地给出变分学中一些最基本的概念，如变分等．否则，不但不能明确地介绍一般泛函的极值理论，而且对常用的几个非线性泛函的极值理论，也很难作到比较满意的叙述．但另一方面，由于使用了泛函分析的一些最基本的知识，因此会给未接触过泛函分析的读者带来一些困难．不过书中所用的泛函分析知识都是最基本的，而且用到的也不多，只要稍许用点时间，就能很快的掌握所用到的那些知识，对阅读本书不会带来多大的困难．

书中配有一定的例题，并附有各章的习题与习题的答案，有些习题也可当作例题看，以利于自学．

由于水平所限，书中一定有很多错误与缺点，恳请批评与指正．

编　者

1987年12月

目　录

第一章 变分问题

函数极值问题，是科学技术上经常遇到的问题。在古典分析中，对函数极值问题，特别是对可微函数的极值问题，已经作了比较有效的讨论和处理。泛函极值问题，在科学技术特别是在近代科学技术中，更有着重要作用。也需要对泛函极值问题作充分的讨论和处理。泛函极值问题，也称为变分问题，处理这类问题的方法，常称为变分法。

§1.1 变分问题的提出

为了便于理解变分问题，我们从历史上有名的三个变分问题谈起。

1．捷线问题

设有不在同一直线上的两点A，B，在A，B两点上联结着某一曲线C，如果略去重物与线之间的摩擦力，重物沿C由A自由下滑到B所需时间最少，便说曲线C是捷线或最速降线，为了把捷线问题写成数学形式，我们取坐标系如图1—1所示。于曲线C：$y=y(x)$上一任取一点$M(x, y)$设S表示$\overset{\frown}{AM}$弧的弧长，m表示重物的质量，g为引力加速度，v为重体从A下滑到M点时的速度，T为重体由A下滑到B所需的时间，即总降落时间。则重体由A到M失去势能

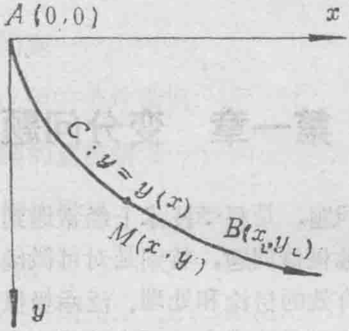

图1－1

mgy，获得动能$\frac{1}{2}mv^2$．由能量守恒定律得到

$$mgy=\frac{1}{2}mv^2,$$

或　　$$v=\sqrt{2gy}\,.$$

于是　　$$\frac{\mathrm{d}s}{\mathrm{d}t}=\sqrt{2gy}\,,$$

从而　　$$\mathrm{d}t=\frac{\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}}{\sqrt{2gy}}=\sqrt{\frac{1+y'^2}{2gy}}\mathrm{d}x,$$

故　　$$T=\int_0^T\mathrm{d}t=\int_0^{x_2}\sqrt{\frac{1+y'^2}{2gy}}\mathrm{d}x$$

令　$C_n[a,b]=\{x(t):x(t)$在$[a,b]$有n阶连续导数$\}$，

且　$C_n[a,b]$中的元素的范数为

$$\|x\|=\max_{a\leqslant x\leqslant b}\{|x(t)|,\ |x'(t)|,\ \cdots,\ |x^{(n)}(t)|\},$$

显然，$T=T(y)$可以看成是定义在$C_1[0,x_2]$上的泛函．

捷线问题的数学说法为：

在$C_1[0, x_2]$中选取满足$y(0)=0$，$y(x_2)=y_2$的函数y，使泛函T为最小值。

捷线问题是Johann · Bernoulli 在1696年以会开信形式提出来的，曾引起广泛的注意；历经Leibniz、Newton和Jacob · Bernoulli 等的努力，才得到较完美的解答。

2．短程线的问题

给定曲面$\phi(x, y, z)=0$，求此曲面上A、B两点间长度最缩曲线。这个最短的曲线，叫做短程线（图1-2）．

在曲面$\phi(x, y, z)=0$上的曲线方程可以写成

$$\begin{cases} y=y(x), \\ z=z(x). \end{cases}$$

其中$y(x)$满足$\phi(x, y, z)=0$，因此，$A(x_1, y_1, z_1)$、$B(x_2, y_2, z_2)$两点间曲线长度为

$$L=L[y(x), z(x)]=\int_{x_1}^{x_2}\sqrt{1+y'^2+z'^2}\,dx.$$

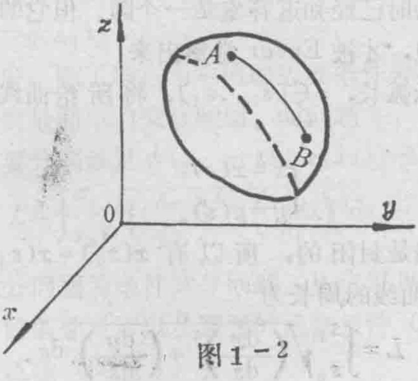

图1－2

设 $C_1^{(2)}[x_1, x_2]=\{(y(x), z(x)): y'(x)$及$z'(x)$在$[x_1, x_2]$上连续$\}$，$C_1^{(2)}$中的元素$u=(y(x), z(x))$是矢量函数，其范数定义为

$$\|u\|=\max_{x_1\leq x\leq x_2}\{|y(x)|, |z(x)|, |y'(x)|, |z'(x)|\},$$

则$L[y(x), z(x)]$可以看作定义在$C_1^{(2)}[x_1, x_2]$上的泛函。于是短程线问题，可以说成：

在函数$y(x)$，$z(x)$满足$\phi(x, y, z)=0$的条件下，从$C_1^{(2)}[x_1, x_2]$选取满足$y(x_1)=y_1$，$z(x)=z_1$，$y(x_2)=y_2$，$z(x_2)=z_2$的矢量函数$(y(x), z(x))$，使泛函$L[y(x), z(x)]$取最小值。

这是一个典型的变分问题。

短程线问题已经在1697年由Johann · Bernoulli所解决。但是这问题的普遍理论，直到后来通过L · Euler，L · Lagrange的努力，才得到解决。

3．等周问题

在长度一定的曲线中，什么曲线所围的面积最大。这个问题在古希腊时已经知道答案是一个圆。但它的变分特性，直到十八世纪，才被Euler觉察出来。

设s表示弧长，$s\in[s_0, s_1]$，将所给曲线用参数形式表达为

$$\begin{cases}x=x(s),\\ y=y(s).\end{cases}$$

因为这条曲线是封闭的，所以有$x(s_0)=x(s_1)$，$y(s_0)=y(s_1)$，这条曲线的周长为

$$L=\int_{s_0}^{s_1}\sqrt{\left(\frac{dx}{ds}\right)^2+\left(\frac{dy}{ds}\right)^2}ds.$$

设 A 为曲线围成的面积，由 Green公式知

$$A=\iint_D \mathrm{d}x\mathrm{d}y=\frac{1}{2}\oint_C x\mathrm{d}y-y\mathrm{d}x$$

$$=\int_{s_0}^{s_1}\left(x\frac{\mathrm{d}y}{\mathrm{d}x}-y\frac{\mathrm{d}x}{\mathrm{d}s}\right)\mathrm{d}s.$$

这样，问题可表述为：

在 $x(s)$，$y(s)$ 满足

$$L=\int_{s_0}^{s_1}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}s}\right)^2}\mathrm{d}s$$

的条件下，在 $C_1^{(2)}[s_0, s_1]$ 中选取满足 $x(s_0)=x(s_1)$，$y(s_0)=y(s_1)$ 的函数 $(x(s), y(s))$，使泛函 $A=A(x(s), y(s))$ 为最大值，这类问题叫等周问题.

在变分问题中，所求函数两端介限已定不变，其端值也不变，这类变分问题，称为边界已定或固定边界的变分问题. 以上三个变分问题，都是边界已定的变分问题. 在问题1中，$y(0)=0$，$y(x_2)=y_2$；在问题2中，$y(x_1)=y_1$，$z(x_1)=z_1$，$y(x_2)=y_2$，$z(x_2)=z_2$；在问题3中，$x(s_0)=x(s_1)$，$y(s_0)=y(s_1)$.

在问题1中，除了端点为定值的边界条件外，没有其它条件，这是一类最简单的变分问题. 而问题2、3中，除边界条件外，还要分别满足 $\phi(x, y(x), z(x))=0$ 和

$$L=\int_{s_0}^{s_1}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}s}\right)^2}\mathrm{d}s.$$

我们称这类变分问题为条件变分问题. 从这里也可看出，条件变分问题中的条件，可以是通常的函数条件（问题2），也可以是某个泛函（问题3）.

这三个有名的变分问题，都是十七世纪末提出的，又都是在十八世纪上半叶解决的．解决的过程中，Euler 和 Lagrange 创立了现在大家都熟悉的变分法，这个变分法后来广泛的用在力学的各个方面，对力学的发展起了很重要的作用．

§1.2 泛函极值

定义 给定泛函$J[y]$，$y\in M$，M是度量空间 $E=(E,\rho)$ 的一个子空间，若在M中某一点y_0的邻域

$$N(y_0,\ \delta)=\{y:y\in M,\ \rho(y_0,\ y)<\delta\}$$

内，下列不等式成立

$$J[y]\geqslant J[y_0]\ (\text{或者}J[y]\leqslant J[y_0])\ ,$$

便称$J[y]$在y_0取相对极小（大）值．若对M中所有的y，下列不等式成立

$$J[y]\geqslant J[y_0]\ (\text{或者}J[y]\leqslant J[y_0])\ ,$$

便称$J[y]$在y_0取得绝对极小（大）值．明显，泛函$J[y]$在y_0取绝对极小（大）值时，也必在y_0取相对极小（大）值，反过来不一定对．

设泛函$J[y]$定义在$C_1[x_1,\ x_2]$的某个子空间M上，y_0是M中的一个点，如果在

$$N_1(y_0,\delta)=\{y:y\in M\text{且}\max_{x_1\leqslant x\leqslant x_2}\{|y-y_0|,|y'-y_0'|\}<\delta\}$$

内，泛函$J[y]$在y_0取相对极小(大)值，便称这种极小（大）值为弱极小（大）值．如果在

$$N(y_0,\delta)=\{y:y\in M\text{且}\max_{x_1\leqslant x\leqslant x_2}\{|y-y_0|\}<\delta\}$$

内，泛函$J[y]$在y_0取相对极小(大)值，便称这种极小（大）值为强极小（大）值。

显然，当$y\in N_1(y_0, \delta)$时，必有$y\in N(y_0, \delta)$，反之不一定对，因此，当泛函$J[y]$在y_0取强极小（大）值时，它也必在y_0取弱极小（大）值，反过来不一定对。

例1 给定泛函

$$J[y]=\int_0^{\pi} y^2(1-y'^2)\mathrm{d}x,\ y\in M,$$

其中，$M=\{y: y\in C_1[0,\pi],\ y(0)=0,\ y(\pi)=0\}$，取$y_0\equiv 0\in M$，则对于$M$中所有满足

$$\max_{0\leqslant x\leqslant \pi}\{|y|,\ |y'|\}<1$$

的y，有$J[y]\geqslant J[y_0]=0$，取$y=\frac{1}{\sqrt{n}}\sin nx$，只要$\frac{1}{\sqrt{n}}<\delta$，即$n>\frac{1}{\delta^2}$，就有$|y|<\delta$，但对此$y$，有

$$\begin{aligned}J[y]&=\int_0^{\pi}\frac{1}{n}\sin^2 nx(1-n\cos^2 nx)\mathrm{d}x\\&=\frac{1}{n}\int_0^{\pi}\sin^2 nx\mathrm{d}x-\frac{1}{4}\int_0^{\pi}\sin^2 2nx\mathrm{d}x\\&=\frac{\pi}{2n}-\frac{\pi}{8}<0\ (n>4).\end{aligned}$$

于是，对任意的$\delta>0$，当$n>\max\left\{4,\ \frac{1}{\delta^2}\right\}$时，虽$\max\limits_{0\leqslant x\leqslant \pi}|y|<\delta$，但$J[y]<J[0]$。因此，$J[y]$在$y\equiv 0$不取强极小值。

例2 $J[y]=\int_{-1}^{1} y^2(1-y'^2)^2\mathrm{d}x,\ y\in M,$

其中　$M=\{y:y\in D_1[-1,\ 1]$且$y(-1)=0,\ y(1)=1\}$，
而　$D_1[-1,\ 1]=\{y:y$在$[-1,\ 1]$连续，y'在$[-1,\ 1]$分段连续$\}$并具有$C_1[-1,\ 1]$的度量。

显然，$J[y]$在M有相对极小值0，并可在M上达到．事实上，取

$$y(x)=\begin{cases}0, & x\in[-1,\ 0];\\ x, & x\in[0,\ 1].\end{cases}$$

那么，$y(x)\in M$，且

$$J[y]=\int_{-1}^{0}0(1-0)^2\mathrm{d}x+\int_{0}^{1}x^2(1-1)^2\mathrm{d}x=0.$$

如果令$M_1=\{y:y\in C_1[-1,\ 1]$，且$y(1)=0 . y(1)=1\}$则$J[y]$在M_1不能达到绝对极小值（注意$y_0\equiv 0\,\overline{\in}\,M_1$，因为$y_0(1)=0\neq 1$）。

§1.3　可微泛函

在古典分析里，中心问题之一是研究可微函数的极值．在变分学中，则主要是研究可微泛函的极值，它具有与可微函数极值问题相类似的特性．其实，变分学的主要目的，在于将古典分析中处理可微函数极值的方法作这样的推广，使我们有可能来解决类似的可微泛函的极值问题．

定义　设E'为线性赋范空间E的一个子空间，且当y及$y+\Delta y$属于E'时，$y+t\Delta y$也属于E'（$-\infty<t<+\infty$），若y从y_0变到$y_0+\Delta y$，$J[y]$的改变量具有形式：

$$\begin{aligned}\Delta J&=J[y_0+\Delta y]-J[y_0]\\&=L[y_0,\ \Delta y]+r[y_0,\ \Delta y],\end{aligned}$$

其中，$L[y_0, \Delta y]$ 是Δy的线性连续泛函，而$r[y_0, \Delta y]$是$\|\Delta y\|\to 0$时较Δy高阶的无穷小，即对任意给定的$\varepsilon>0$，存在 $\delta>0$，使得$\|\Delta y\|<\delta$ 时，恒成立 $|r[y_0, \Delta y]|<\varepsilon\|\Delta y\|$，便称$J[y]$在$y_0$可微，而$L[y_0, \Delta y]$称为$J[y]$在$y_0$处的微分或变分，记为$\delta J$或$\delta J[y_0, \Delta y]$，即

$$\delta J = L[y_0, \Delta y].$$

只可能有一个线性连续泛函满足上述条件，事实上，若

$$\begin{aligned}\Delta J &= L_1[y_0, \Delta y] + r_1[y_0, \Delta y]\\ &= L_2[y_0, \Delta y] + r_2[y_0, \Delta y],\end{aligned}$$

则 $$L_1 - L_2 = r_1 - r_2 = r[y_0, \Delta y]$$

仍是比Δy高阶的无穷小，从而对任何给定的 $\varepsilon>0$，可找到$\delta>0$，使得$\|\Delta y\|<\delta$时，

$$|L_1 - L_2| = |r_1 - r_2| < \varepsilon,$$

于是得$L_1 = L_2$

常记$\Delta y = \delta y$，并称它为宗量y的变分，于是有

$$\delta J[y_0, \Delta y] = \delta J[y_0, \delta y] = L[y_0, \Delta y] = L[y_0, \delta y].$$

记 $\delta y' = (\delta y)'$，有 $\delta y' = (y - y_0)' = y' - y_0' = \delta(y')$。

定理 若$J[y]$在$y_0 \in E'$可微，则对于任何的δy，把$J[y_0 + \eta\delta y]$看作η的函数时，这函数在$\eta = 0$时对η按通常意义是可微的，并且有

$$\left[\frac{\mathrm{d}}{\mathrm{d}\eta} J[y_0 + \eta\delta y]\right]_{\eta=0} = \delta J[y_0, \delta y]$$
$$= L[y_0, \delta y].$$

证 $$\left[\frac{\mathrm{d}}{\mathrm{d}\eta} J[y_0 + \eta\delta y]\right]_{\eta=0}$$

$$= \lim_{\eta \to 0} \frac{J[y_0 + \eta\delta y] - J[y_0]}{\eta}$$

$$= \lim_{\eta \to 0} \frac{L[y_0, \eta\delta y] - r[y_0, \eta\delta y]}{\eta}$$

$$= \lim_{\eta \to 0} \left[L[y_0, \delta y] + \frac{r[y_0, \eta\delta y]}{\eta} \right]$$

$$= \lim_{\eta \to 0} \frac{r[y_0, \eta\delta y]}{\eta} + L[y_0, \delta y].$$

因$r[y_0, \eta\delta y]$是$\eta \to 0$时较$\eta\delta y$高阶的无穷小。故对任意给定的$\varepsilon > 0$，存在$\delta > 0$，当$\| \eta\delta y \| < \delta$时，恒有

$$|r[y_0, \eta\delta y]| < \varepsilon \| \eta\delta y \| = |\eta| \| \delta y \| \varepsilon.$$

$$\left| \frac{r[y_0, \eta\delta y]}{\eta} \right| < \varepsilon \| \delta y \|.$$

故 $$\lim_{\eta \to 0} \frac{r[y_0, \eta\delta y]}{\eta} = 0.$$

于是

$$\left[\frac{d}{d\eta} J[y_0 + \eta\delta y] \right]_{\eta = 0} = L[y_0, \delta y].$$

要注意，在求$\delta J[y_0, \delta y]$时，$J[y_0 + \delta y]$ 中的 δy 是变量，而在求$\frac{d}{d\eta} J[y_0 + \eta\delta y]$的过程中$\delta y$则是常量。

如果$J[y]$在E'处处可微，则当固定 y 及δy时，$J[y, \eta\delta y]$对每个η可微。事实上，命$\eta = \eta_0 + \tau$，有

$$J[y + \eta\delta y = J[y + \eta_0 \delta y + \tau\delta y]$$

$$= J[y_0 + \tau\delta y] (y_0 = y + \eta_0 \delta y).$$

据以上定理，$J[y_0 + \tau\delta y]$ 在 $\tau = 0$ 可微，于是 $J[y + \eta\delta y]$在