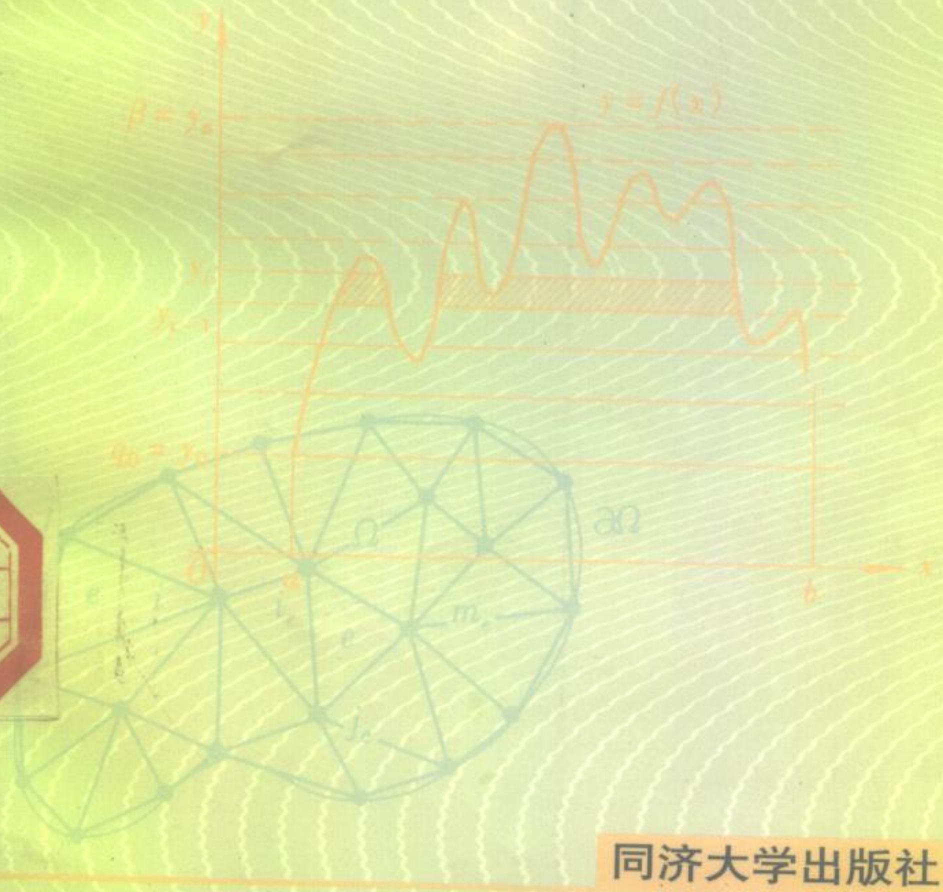
变分方法

徐建平　桂子鹏　编



同济大学出版社

变分方法

徐建平　桂子鹏　编

同济大学出版社

图书在版编目(CIP)数据

变分方法/徐建平,桂子鹏编 .—上海:同济大学出版社,1999.9
ISBN 7-5608-2053-0
Ⅰ.变… Ⅱ.①徐… ②桂… Ⅲ.变分法 Ⅳ.0176
中国版本图书馆 CIP 数据核字(1999)第 30692 号

变 分 方 法
徐建平　桂子鹏　编
同济大学出版社出版
(上海四平路 1239 号　邮编:200092)
浙江新昌印刷厂印刷
开本:850×1168　1/32　印张:6.75　字数:190 千字
1999 年 8 月第 1 版　　1999 年 8 月第 1 次印刷
印数:1500　　定价:10.00 元
ISBN 7-5608-2053-0/O·173

前　言

变分方法是应用数学专业的一门数学基础课,学生通过对本课程的学习,可以对泛函分析的一些基本概念有所了解,并且掌握变分问题古典处理方法和现代的一些处理方法;同时,对变分问题的近似解法也有所了解.变分问题是工程技术中常遇到的一类优化问题,是数学物理方法的重要的一个分支.学生掌握这方面的知识,可为今后学习后继课程、进一步开拓知识面、开展工程和数学方面的研究打好一定的基础.

本书作为教材,可供应用数学专业学生作必修和选修使用,也可作为其他理工科专业研究生的教材或者参考书.其中第一章古典变分方法也可作为工科大学生的学习材料.本教材是作者在多年教学实践中使用的讲义的基础上,通过修改和补充编著而成的.本书的内容包括古典的变分法、泛函分析基础知识、变分原理以及变分问题的近似解等.

本书综合应用经典方法和泛函分析方法来讨论泛函的极值问题,并力求用近代的方法和观点来贯穿全书.在内容的编排上,作者注意到由易到难、循序渐进原则.在文字组织上,条理清楚,通俗易懂,使抽象问题尽量做到形象化,使得读者便于自学.在各章中,通过典型的例子突出其重要的内容,或阐明基本思想和计算技巧.在每一章中都配有适量的习题,并在书后附有全部习题的解答或答案.我们深信,有一定“高等数学”和“微分方程”基础知识的读者,通过一定的努力,是能够掌握书中的内容的.在教学时,本书的内容可以根据教学时数的实际情况作适当的调整.教完本书的全部内容估计需要用 36 个教时.如果仅仅介绍第一章的古典变分问题,只需要 18 个教时左右.

本书由桂子鹏和徐建平编写,徐建平编写第一章,桂子鹏编写第二、三、四章.在本书的编写过程中,作者得到同济大学应用数学系许多教师的支持和帮助,在此表示衷心的感谢.由于编者的水平有限,书中的缺点及不妥的地方在所难免,恳请读者批评指正.

编　者

1998年5月　于同济大学

目　　录

第一章　变分问题

在微积分这一学科形成的初期，人们就已经遇到下面这样的问题.例如，牛顿就曾提出过运动于介质中的旋转体，其体形应具备怎样的条件才能使阻力最小的问题；约翰·伯努利则提出了所谓的捷线问题.这类问题都涉及到更广泛意义下的极值问题，即有关泛函极值问题.我们把这类问题称为变分问题，而处理这类问题的方法则称为变分法.它是微分学中处理有限个变量的函数极值问题的方法的扩展，而随着变分法的发展，则逐步形成并发展成为变分学.另一方面，变分学理论的发展又与力学、物理学等的发展密切相关.从力学的 Hamilton 原理、最小位能原理到目前在自然科学和工程技术中提出的某些“最佳方案”、“最优设计”等，其本质均为变分问题.而当今发展较快的数值计算方法中的有限元方法等则更是以变分学为它的基础.本章从古典的变分法开始讨论，引出有关概念，并为以后各章节的学习作准备.

1.1　变分问题的引入

例 1.1.1　捷线问题

在初速为零且仅受重力作用的情况下，质点沿光滑曲线由点 $A(x_0, y_0)$ 无摩擦地滑行到点 $B(x_1, y_1)$. 为了使滑行时间最短，问沿着滑行的曲线形状应如何？

我们建立坐标系如图 1.1.1 所示. 假设曲线 $\overset{\frown}{AB}$ 的方程为 $y = y(x)$，则质点在 $M(x, y)$ 处的速率为 $v = \sqrt{2gy(x)}$.

若走过路程为　$ds = \sqrt{1 + y'^2(x)}\,dx$,

则所需的时间为 $dT = \frac{ds}{v} = \frac{\sqrt{1+y'^2(x)}}{\sqrt{2gy(x)}}dx.$

故走过整个弧段的总时间为

$$T = J[y] = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2(x)}}{\sqrt{2gy(x)}}dx.$$

根据条件,我们所选用的函数$y(x)$在$[x_0,x_1]$上具有一阶连续导数(记作$y(x)\in C^1[x_0,x_1]$),且$y(x_0)=y_0,y(x_1)=y_1$.

所谓捷线问题,就是要求在满足上述条件的曲线中,找到函数$y=y^*(x)$所表示的曲线,使得$T=T_{\min}=J[y^*]$.

例 1.1.2 最小曲面问题

在空间中,给定一条闭曲线 C.要在这条曲线上张一张曲面,使得它具有最小的面积.

如图 1.1.2 所示,设 Γ 是空间曲线 C 在 xOy 平面上的投影,所张成的曲面的方程为$z=z(x,y)$,于是曲面面积为

$$S = J[z] = \iint_D \sqrt{1+z_x'^2+z_y'^2}dxdy.$$

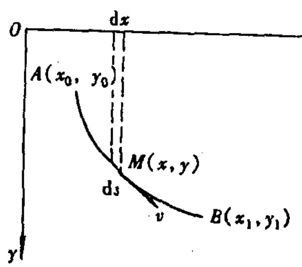


图 1.1.1

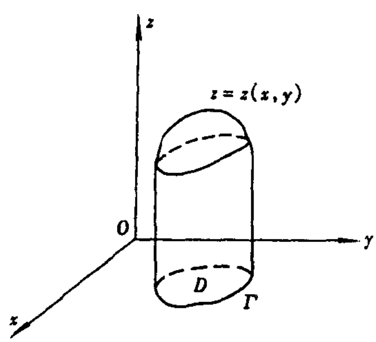


图 1.1.2

根据条件,选用函数$z=z(x,y)\in C^1(D)$,且$z=z(x,y)$所表示的曲面过曲线 C.

本问题就是寻找满足上述条件的函数 $z = z^*(x,y)$，使 $S = S_{\min} = J[z^*]$.

例 1.1.3 悬链线问题

如图 1.1.3 所示，给定过固定点 $A(x_0,y_0)$，$B(x_1,y_1)$ 的柔软且定长为 l 的绳索，在重力作用下，呈下垂状态.

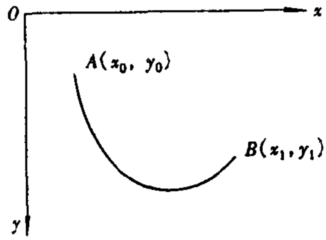


图 1.1.3

假设绳索状态曲线为 $y = y(x)$，则

$$\int_{x_0}^{x} \sqrt{1 + y'^2(x)}\,\mathrm{d}x = l.$$

绳索在平衡状态时，它的重心最低，即

$$\bar{y} = J[y] = \frac{1}{l}\int_{x_0}^{x} y(x)\sqrt{1 + y'^2(x)}\,\mathrm{d}x$$

最大.

本问题就是在 $C^1[x_0,x_1]$ 内寻找一个函数 $y = y^*(x)$，使它在满足 $y_0 = y(x_0)$，$y_1 = y(x_1)$ 及 $\int_{x_0}^{x_1}\sqrt{1 + y'^2(x)}\,\mathrm{d}x = l$ 的条件下，使 $\bar{y} = \bar{y}_{\max} = J[y^*]$.

综观上述问题，可以看到这些问题都是在求依赖于函数的“函数”（这里为积分）的极值. 如果我们把这类依赖于函数的“函数”叫做泛函的话，则变分法的主题就是寻求泛函的极值.

为了讨论变分问题，我们在下节引入泛函及泛函极值的有关概念.

1.2 泛函及泛函极值的概念

1.2.1 泛函的有关概念

1.2.1.1　距离及距离空间

定义 1.2.1　设 X 是一非空集合,若对于 X 中任意元素 x,y,总对应有实数值 $\rho(x,y)$,它满足以下三个条件:

① $\rho(x,y) \geqslant 0$,且 $\rho(x,y)=0$ 的充要条件是 $x=y$;

② $\rho(x,y)=\rho(y,x)\quad \forall x,y \in X$;

③ $\rho(x,z) \leqslant \rho(x,y)+\rho(y,z)\quad \forall x,y,z \in X$,

则称 $\rho(x,y)$ 是 X 中元素 x,y 间的距离.

定义了距离的集合 X 称为距离空间.由于在一个集合中,定义距离的方式可以不同,这样所构成的距离空间也就不同.在集合 X 中引入距离后,我们则称在 X 中引入了拓扑结构.为了区别起见,可把定义了距离 ρ 的距离空间记为 (X,ρ).

例如,在 n 维 Euclid 集 R^n 中,可定义距离

$$\rho_1(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y^2)+\cdots+(x_n-y_n)^2}$$

$$\forall x=(x_1,\cdots,x_n),\ y=(y_1,\cdots,y_n) \in R^n.$$

也可定义距离

$$\rho_2(x,y)=\max_{1\leqslant i\leqslant n}|x_i-y_i|\quad \forall x,y \in R^n.$$

显然 (R^n,ρ_1) 与 (R^n,ρ_2) 具有不同的拓扑结构.

定义 1.2.2　设 X 为距离空间,定义的距离为 $\rho(x,y)$,又 $x_0 \in X$.则称点集

$$B(x_0,\delta)=\{x \in X \mid \rho(x,x_0) \leqslant \delta, \delta>0\}$$

为以 x_0 为中心、半径为 δ 的邻域,简称为 x_0 的 δ 邻域.

定义 1.2.3　设在距离空间 X 中有元素列(点列) $\{x_n\}$ 及元素 x_0,如果有 $\lim\limits_{n\to\infty}\rho(x_n,x_0)=0$,则称 x_0 为元素列(点列) $\{x_n\}$ 的极限元素(极限点)或称 $\{x_n\}$ 收敛到 x_0,记为 $\lim\limits_{n\to\infty}x_n=x_0$ 或 $x_n \to x_0$ $(n\to\infty)$.

距离空间中的收敛点列具有数学分析中讨论过的数列极限的某些基本性质,例如极限唯一性等.

设 X 是配备了距离 ρ 的距离空间 (X,ρ), $\{x_n\}$ 是 X 中的一个

点列.如果对任意给定的$\varepsilon > 0$,总存在正整数$N = N(\varepsilon)$,使当$p,q \geqslant N$时,恒有$\rho(x_p,x_q) < \varepsilon$,则称点列$\{x_n\}$为距离空间$(X,\rho)$的基本序列.

1.2.1.2　完备的距离空间

定义 1.2.4　设(X,ρ)为距离空间.如果(X,ρ)中任何基本序列都是收敛序列且收敛到(X,ρ)中的元素,则称距离空间(X,ρ)是完备的距离空间.

实数空间 R 是完备的,而有理数空间则是不完备的.在不完备的距离空间中添加入一切基本序列所收敛到的极限点后,就可成为一个完备的距离空间.这个空间称为完备化距离空间.

定义 1.2.5　设M是距离空间X的子集.若对任何$x \in X$,总可在M中找到一个点列$\{x_n\}$,使得$\lim\limits_{n\to\infty} x_n = x$,则称$M$为$X$的稠密集.

不完备距离空间在它的完备化距离空间中是稠密的.

1.2.1.3　连续函数空间$C(\Omega)$及$C^m(\Omega)$

设Ω是空间R^n中的有界闭集,$C(\Omega)$表示由定义在Ω上的连续函数全体组成的集合,称为连续函数空间.如果我们在连续空间$C(\Omega)$中引入距离

$$\rho(f,g) = \max_{x\in\Omega} |f(x) - g(x)|, \forall f,g \in C(\Omega).$$

就使$C(\Omega)$成为距离空间.

设$\{f_n(x)\}$为$C(\Omega)$中的基本序列,则$\forall \varepsilon > 0$, $\exists N > 0$,当$n,m \geqslant N$时,恒有

$$\rho(f_n(x),f_m(x)) = \max_{x\in\Omega} |f_n(x) - f_m(x)| < \varepsilon,$$

即对$\forall x \in \Omega$,都有$|f_n(x) - f_m(x)| < \varepsilon$.这是函数列$\{f_n(x)\}$一致收敛的充要条件.故存在$f(x)$,使得当$n \to \infty$时,$f_n(x)$一致收敛于$f(x)$.

由于有“一致收敛的连续函数列必收敛于连续函数”这样一个定理(不证),我们得到结论:连续函数空间$C(\Omega)$是完备的距离

空间.

我们把在 Ω 上具有 m 阶连续(偏)导数的函数全体组成的集合称为 m 阶连续可微函数集合,记作 $C^m(\Omega)$,其中 $m > 0$ 为整数(特别规定 $C(\Omega) \equiv C^0(\Omega)$),称为 m 阶连续可微函数空间.显然,$C^m(\Omega) \subset C^{m-1}(\Omega) \subset C^{m-2}(\Omega) \subset \cdots \subset C^1(\Omega) \subset C^0(\Omega) \equiv C(\Omega)$.在 $C^m(\Omega)$ 中引入距离:

$$\rho(f,g) = \max_{x\in\Omega}\{|f(x) - g(x)|, |\partial^k(f(x) - g(x))|, k = 1,2,\cdots,m\},$$
$$\forall f,g \in C^m(\Omega)$$

其中,$\partial^k f(x)$ 表示 $f(x)$ 的 k 阶(偏)导数.这样,$C^m(\Omega)$ 就成为距离空间,且也是完备的距离空间.

1.2.1.4　线性赋范空间

(1) 线性空间

定义 1.2.6　设 X 是一个集合.在集合 X 上规定称为**加法**及**数乘**的两种运算(合称线性运算)如下:

加法:对于任意 $x,y \in X$,都存在 X 中的唯一元素与之对应,此元素记作 $x + y \in X$.

数乘:对于任意的 $x \in X$ 及任意的数 $\alpha \in K$,都存在 X 中唯一元素与之对应,此元素记作 $\alpha x \in X$,其中 K 为实数域或复数域.

如果这两种运算又满足下列八条性质,则称 X 为线性空间:

① $x + y = y + x, \forall x,y \in X$(加法交换律);

② $x + (y + z) = (x + y) + z, \forall x,y,z \in X$(加法结合律);

③ 存在"零元"$0 \in X$,使 $x + 0 = x, \forall x \in X$;

④ 存在"逆元" $-x \in X$,使 $x + (-x) = 0, \forall x \in X$;

⑤ $\alpha(\beta x) = (\alpha\beta)x, \forall \alpha,\beta \in K, x \in X$(数乘结合律);

⑥ $1 \cdot x = x, \forall x \in X$;

⑦ $(\alpha + \beta)x = \alpha x + \beta x; \forall \alpha,\beta \in K, x \in X$(第一分配律);

⑧ $\alpha(x + y) = \alpha x + \alpha y, \forall \alpha \in K, x,y \in X$(第二分配律).

定义 1.2.7　设集合 X 是集合 Y 的子集.若在 X 和 Y 中定义相同意义的加法和数乘而都成为线性空间,则称线性空间 X 是线

性空间Y的子空间；若X又是Y的真子集，则称X是Y的真子空间.子空间与真子空间分别记为$X \subseteq Y$与$X \subset Y$.

若在连续函数空间$C(\Omega)$中引进线性运算如下：

加法：设$f(x) \in C(\Omega), g(x) \in C(\Omega)$,

则$(f+g)(x) = f(x) + g(x), \forall x \in \Omega$.

数乘：设$f(x) \in C(\Omega), \alpha \in K$,

则$(\alpha f)(x) = \alpha \cdot f(x), \forall x \in \Omega$.

首先，我们来考察$f+g$及αf是否仍为$C(\Omega)$的元素，注意到$C(\Omega)$中元素f的连续性性质，即对任何$f(x) \in C(\Omega)$，都具有性质$\lim\limits_{x \to x_0} f(x) = f(x_0), \forall x_0 \in \Omega$.于是

$$\lim_{x \to x_0}(f+g)(x) = \lim_{x \to x_0}(f(x) + g(x)) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$$
$$= f(x_0) + g(x_0) = (f+g)(x_0), \forall x_0 \in \Omega;$$
$$\lim_{x \to x_0}(\alpha f)(x) = \lim_{x \to x_0} \alpha \cdot f(x) = \alpha \cdot \lim_{x \to x_0} f(x)$$
$$= \alpha \cdot f(x_0) = (\alpha f)(x_0), \forall x_0 \in \Omega.$$

因此　$(f+g)(x) \in C(\Omega), (\alpha f)(x) \in C(\Omega)$.

只要规定“零元”是函数$f(x) \equiv 0$，“逆元”是负函数$(-f)(x) = -f(x)$，就可逐条验证上述八条性质.于是可知$C(\Omega)$是线性空间.

同样，可以证明$C^m(\Omega)$也是线性空间.

(2) 范数和线性赋范空间

定义1.2.8　设X是线性空间.若对于X中的任一元素x都有一个被记为$\|x\|$的实数与它对应，且满足下列三个条件：

① $\|x\| \geqslant 0, \forall x \in X$，且$\|x\| = 0$的充要条件是$x = 0$(零元素)；

② $\|\alpha x\| = |\alpha| \cdot \|x\|, \forall \alpha \in K, \forall x \in X$;

③ $\|x+y\| \leqslant \|x\| + \|y\|, \forall x, y \in X$(三角形不等式)，

则称$\|x\|$为元素x的范数.

定义了范数的线性空间称为线性赋范空间(也称赋范空间).

在 $C(\Omega)$ 中,当 Ω 为有界闭集时,可引入范数

$$\|f(x)\| = \max_{x\in\Omega}|f(x)|, \quad \forall f(x)\in C(\Omega).$$

可以验证范数的三个条件是满足的,即

① $\|f(x)\| \geqslant 0$,且 $\|f(x)\| = 0$ 等价于 $f(x)\equiv 0$,

$$\forall x\in\Omega;$$

② $$\begin{aligned}\|\alpha f(x)\| &= \max_{x\in\Omega}|\alpha\cdot f(x)| \\ &= |\alpha|\cdot\max_{x\in\Omega}|f(x)| = |\alpha|\cdot\|f\|,\end{aligned}$$

$$\forall \alpha\in K, f(x)\in C(\Omega);$$

③ $$\begin{aligned}\|f(x)+g(x)\| &= \max_{x\in\Omega}|f(x)+g(x)| \\ &\leqslant \max_{x\in\Omega}|f(x)| + \max_{x\in\Omega}|g(x)| \\ &= \|f(x)\| + \|g(x)\|,\end{aligned}$$

$$\forall f(x), g(x)\in C(\Omega).$$

因此,$C(\Omega)$ 是线性赋范空间.

在 $C^m(\Omega)$ 中,若 Ω 为有界闭集时,引入范数

$$\|f\|_m = \max_{x\in\Omega}\{|f(x)|, |\partial^k f(x)|, k = 1,2,\cdots,m\}$$

$$\forall f(x)\in C^m(\Omega)$$

显然,它满足范数的三个条件,因此 $C^m(\Omega)$ 也是线性赋范空间.

设 $\|x\|$ 是 X 上定义的范数,则可以派生出自然的距离:

$$\rho(x,y) = \|x-y\|, \quad \forall x,y\in X.$$

显然 ① $\rho(x,y) = \|x-y\| \geqslant 0$

且 $\rho(x,y) = \|x-y\| = 0$ 等价于 $x = y$;

② $\rho(x,y) = \|x-y\| = \|y-x\| = \rho(x,y)$

$$\forall x,y\in X;$$

③ $$\begin{aligned}\rho(x,z) &= \|x-z\| = \|(x-y)+(y-z)\| \\ &\leqslant \|x-y\| + \|y-z\| = \rho(x,y)+\rho(y,z)\end{aligned}$$

$$\forall x,y,z\in X.$$

因此,线性赋范空间必然是距离空间.

定义 1.2.9 设 X 是配备了范数 $\|\cdot\|$ 的线性赋范空间,

$\{x_n\}$ 是 X 中的一个点列.若在 X 中存在 x,使得

$$\rho(x_n,x) = \|x_n - x\| \to 0,\ n \to \infty,$$

则称点列 $\{x_n\}$ 依范数 $\|\cdot\|$ 在 X 中收敛于 x.

若点列 $\{x_n\}$ 在距离 ρ 意义下是基本序列,则称点列 $\{x_n\}$ 是依范数 $\|\cdot\|$ 的基本序列.

定义 1.2.10 设 X 是配备了范数 $\|\cdot\|$ 的线性赋范空间.若 X 中的任何基本序列都是收敛序列,则称 X 是完备的线性赋范空间,也称 Banach 空间.

可知 $C^m(\Omega)$ 在 $\|\cdot\|_m$ 范数下是完备的线性赋范空间——Banach 空间.

定义 1.2.11 设 X 是线性空间,$\|\cdot\|_1$ 和 $\|\cdot\|_2$ 是定义在 X 上的两个范数.若存在正的常数 c,使得 $\|x\|_1 \leqslant c\|x\|_2$, $\forall x \in X$ 成立,则称范数 $\|\cdot\|_2$ 强于范数 $\|\cdot\|_1$.若存在两个正的常数 c_1 和 c_2,使得 $c_1\|x\|_1 \leqslant \|x\|_2 \leqslant c_2\|x\|_1$, $\forall x \in X$ 成立,则称范数 $\|\cdot\|_2$ 和范数 $\|\cdot\|_1$ 等价.

若范数等价,则由此派生的距离、极限、收敛以及由收敛定义的其他概念(如完备性)都是等价的.

1.2.2 泛函极值的概念

设 X 是函数空间.若对于 X 中的子集 D 中每一函数 $y(x)$,按照一定的法则都有确定的数值 J 与它对应,则称 J 是函数 $y(x)$ 在 D 中的泛函,记作 $J = J[y]$,其中,D 称为该泛函的定义域.

定义 1.2.12 设 $J[y]$ 是定义在函数距离空间 X 上的一个泛函.若存在一个 $y^* \in X$ 及某个 $\delta > 0$,使对于一切 $y \in B(y^*,\delta)$ 成立

$$J[y^*] \leqslant J[y] \quad (\text{或 } J[y^*] \geqslant J[y]),$$

其中,$B(y^*,\delta)$ 为以 y^* 为中心的、半径为 δ 的邻域,则称泛函 $J[y]$ 在 y^* 处取得极小值(或极大值).

这里,极小值(或极大值)也称为相对极值,函数 y^* 称为极值函数.

定义 1.2.13 设$J[y]$是定义在函数距离空间 X 上的一个泛函.若存在一个$y^* \in X$,使对一切$y \in X$,恒有

$$J[y^*] \leqslant J[y] \quad (\text{或 } J[y^*] \geqslant J[y]),$$

则称泛函 $J[y]$ 在 y^* 处取得绝对极小值(或绝对极大值).

绝对极大(小)值也称为最大(小)值.在变分问题中所指的极值,一般是指最大(小)值.

我们把 $C(\Omega)$ 中的 y^* 的 δ 邻域称为零阶邻域,记为 $O(y^*,\delta)$,而把 $C^m(\Omega)$ 中的 y^* 的 δ 邻域称为 m 阶邻域,记为 $O_m(y^*,\delta)$.显然

$$O_m(y^*,\delta) \subset \cdots \subset O_k(y^*,\delta) \subset O_{k-1}(y^*,\delta) \subset \cdots \subset O(y^*,\delta).$$

把 $C(\Omega)$ 中的距离

$$\rho(y,y^*) = \max_{x \in \Omega} |y - y^*|$$

叫做 y 与y^* 的零级距离;而把 $C^m(\Omega)$ 中的距离

$$\rho(y,y^*) = \max_{x \in \Omega}\{|y - y^*|, |\partial^k(y - y^*)|, k = 1,2,\cdots,m\}$$

叫做 y 与y^* 的 m 级距离.我们可以看到:如果在阶数低的邻域意义下的极值函数是 y^*,那么,它必然是阶数高的邻域意义下的极值函数,反之不然.因此,我们把低阶邻域定义的极值称为强极值,而把高阶邻域定义的极值称为弱极值.

1.3 泛函极值存在的必要条件

设$J[y]$是定义在距离函数空间 X 上的泛函.我们仿照函数的微分概念来定义泛函的变分概念.

设 $y_0 \in X$,称 $\delta y = y - y_0, \forall y \in X, y \neq y_0$ 为函数 y 在y_0 处的变分.

这里,δy 是 x 的函数,与 Δy 有区别:变分 δy 反映的是整个函