变分法的原理及应用

1 绪论

本节主要介绍变分学研究的基本内容与其研究的对象。

1.1 变分学是研究什么的呢？

变分学是一门研究**泛函**极值的学问。许多自然现象、物理学规律、工程问题实际上都能归结为泛函的极值问题。譬如说：光为什么沿直线传播、毛细管液面为什么会弯曲过来、肥皂泡的薄膜为什么会呈现出那样的形状、怎么样盖房子会最省 $\cdots\cdots$

1.2 泛函

泛函：

$$I = M(\text{函数集合}) \longrightarrow R \tag{1}$$

是函数集合M到实数集合R的一个映射。可以把它简单地理解为函数的函数，至于为什么要这样理解，在下文中具体讨论变分法时会直观地阐述一下。

显然，泛函是千变万化的，有各种各样的形式。但一般来说，变分学研究的是：给定一个定义域为$[x_1, x_2] \times \Omega \times R^n$、值域为$R'$的函数

$$f(x, y, y') \in C([x_1, x_2] \times \Omega \times R^n, R') \tag{2}$$

M集合由y的边界条件、约束条件（如哪些点要固定、哪些点要挖掉等）$\cdots\cdots$来确定，从而构成的泛函

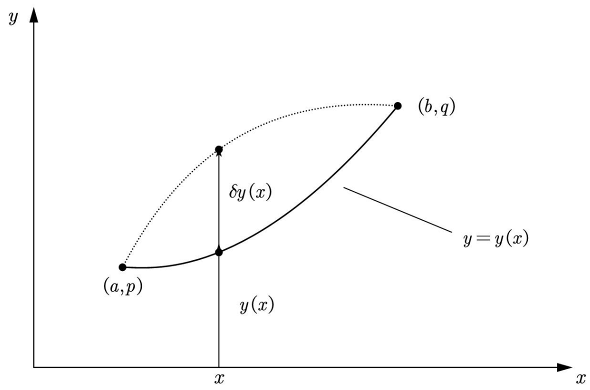
$$I = \int_{x_1}^{x_2} f(x, y, y') dx \tag{3}$$

至于怎么直观地理解这是个什么玩意儿，待下文举出具体例子时你就会明了了。

2 求解泛函极值的工具：欧拉-拉格朗日方程

正如在判断函数是否取到极值时，我们会考虑它的导数值是否为零，判断泛函是否取到极值时我们也需要一种工具：欧拉-拉格朗日方程。本节主要介绍定端变分、无约束条件下的欧拉-拉格朗日方程，并举出一些具体的使用例子。

2.1 从一个例子引入：两点之间为什么线段最短呢？



在直角坐标系中，给定两点$A(x_1,y_1)$、$B(x_2,y_2)$，显然如果要用一条曲线来连接这两个点的话，这个曲线可以是各种各样的，即这个曲线的方程可以是各种各样的。但可以这样来看待这个问题：即便曲线可以是各种各样的，**但每一条曲线在A、B两点间都只对应一个唯一的长度**。从这个角度来看，似乎与一个函数很像：每一个曲线的方程相当于"函数"的自变量，而该曲线对应的一个唯一的长度就相当于"函数"的因变量。这种"函数"就是泛函，它的自变量是函数，因变量是值，所以我们说**泛函是函数的函数**。

那么这个问题不就相当于求一个泛函达到极小值时所对应的"自变量值"（自变量函数的方程）吗？有了这个思路，就可以开始尝试求解了。我们知道曲线的弧微分是：

$$ds=\sqrt{1+y'^2}dx \tag{4}$$

所以连接A、B两点的曲线长度可以表示为：

$$S=\int_{x_1}^{x_2}\sqrt{1+y'^2}dx \tag{5}$$

我们要做的就是找出当S最小时，y的表达式，但从现在看来，我们好像束手无策。至于怎么找，这就是变分学中要学习的东西了。

2.2 回想一下我们是怎样确定函数极值点的？

对于一个一元函数

$$f:\Omega\subset R\longrightarrow R \tag{6}$$

我们是通过它取极值的必要条件来确定极值点的：即在极值点处f需要满足导数为零，从而选出符合该条件的点，极值点就在其中。那么进一步回想，对于一个多元函数，它取极值的必要条件就是在极值点处f需要满足$\nabla f=0$（在极值点处各个方向上的导数为零），但多元函数的这个必要条件是怎么得来的呢？

假设$X_0=(x_0,y_0,z_0\cdots)$是f的极值点，那么对所有$\delta\in R^n$，一定存在关于δ的函数ε，使得$X_0+\varepsilon\delta$在f的定义域中。这相当于做了一个让扰动，它的关键在于自变量经过该扰动后依然在函数的定义域内。那么对于$f(X_0+\varepsilon\delta)$，可以看成是关于ε的一元函数：

$$g(\varepsilon)=f(X_0+\varepsilon\delta) \tag{7}$$

显然，当$\varepsilon=0$时函数取到极值（因为X_0是f的极值点）。所以根据一元函数取极值的必要条件，有$g'(0)=\nabla f(X_0)\cdot\delta=0$，对任意$\delta$都得满足，所以$\nabla f(X_0)=0$. 这样就得到了多元函数取极值的必要条件。

2.3 探讨如何确定泛函的极值点：推导欧拉-拉格朗日方程

借鉴上面的思路，推导出泛函取极值时的必要条件，再通过**必要条件**来确定泛函的极值点。

对于泛函

$$I=\int_{x_1}^{x_2}f(x,y,y')dx \tag{8}$$

假定I在y处取到极值。对y做一个扰动，要想让它经过扰动后依旧在定义域内，则可考虑扰动为一个定义域内的函数。对所有$\delta(x)\subset\Omega$，一定存在ε（一个类似函数的任意变动的数），使得$y+\varepsilon\delta(x)$在y的定义域中。于是做完扰动就得到：

$$Y=y+\varepsilon\delta(x) \tag{9}$$

Y自然可以看成是关于ε的函数。但是要保证

$$\delta(x_1)=\delta(x_2)=0 \tag{10}$$

这是因为我们要解决的问题中两个端点$A:(x_1,y_1)$、$B:(x_2,y_2)$是固定不变的，也就是不能扰动的。这种变分称作端点固定的变分（**定端变分**），那自然也会有端点不固定的情况，这留待之后讨论。

对Y有下列式子：

$$\begin{gathered}Y'=y'+\varepsilon\delta'(x)\\ \frac{dY}{d\epsilon}=\delta(x)\\ \frac{dY'}{d\varepsilon}=\delta'(x)\end{gathered} \tag{11}$$

这些之后会用到的。那么做完这个扰动后，泛函就变成了：

$$I[y+\varepsilon\delta(x)]=\int_{x_1}^{x_2}f(x,Y,Y')dx \tag{12}$$

把它看成关于ε的函数：$F(\varepsilon)=\int_{x_1}^{x_2} f(x,Y,Y')dx$.显然，当$\varepsilon=0$时泛函取到极值（$y$是泛函的极值点）。所以泛函取极值的必要条件就是：

$$\left.\frac{dF(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}=0 \tag{13}$$

并且此时$Y=y$. 接下来，我们就试着求出这个必要条件:

$$\begin{aligned}
\frac{dF(\varepsilon)}{d\varepsilon} &= \int_{x_1}^{x_2}\frac{df(x,Y,Y')}{d\varepsilon}dx \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}\cdot\frac{\partial Y}{\partial \varepsilon}+\frac{\partial f}{\partial Y'}\cdot\frac{\partial Y'}{\partial \varepsilon}]dx \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}\cdot\delta(x)+\frac{\partial f}{\partial Y'}\cdot\delta'(x)]dx \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}\cdot\delta(x)]dx+\int_{x_1}^{x_2}[\frac{\partial f}{\partial Y'}\cdot\delta'(x)]dx \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}\cdot\delta(x)]dx+\int_{x_1}^{x_2}\frac{\partial f}{\partial Y'}d[\delta(x)] \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}\cdot\delta(x)]dx+\left.\frac{\partial f}{\partial Y'}\delta(x)\right|_{x_1}^{x_2}-\int_{x_1}^{x_2}[\delta(x)\cdot\frac{d}{dx}(\frac{\partial f}{\partial Y'})]dx \\
&= \int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}-\frac{d}{dx}(\frac{\partial f}{\partial Y'})]\delta(x)dx
\end{aligned} \tag{14}$$

这里面用到了上面关于Y的三个式子，与$\delta(x_1)=\delta(x_2)=0\rightarrow\left.\frac{\partial f}{\partial Y'}\delta(x)\right|_{x_1}^{x_2}=0$.

所以必要条件就是：$\int_{x_1}^{x_2}[\frac{\partial f}{\partial Y}-\frac{d}{dx}(\frac{\partial f}{\partial Y'})]\delta(x)dx$. 又因为对任意$\delta(x)$都要满足，则只有：

$$\frac{\partial f}{\partial Y}-\frac{d}{dx}(\frac{\partial f}{\partial Y'})=0 \tag{15}$$

又因为取极值时$Y=y$，所以泛函取极值的必要条件是：

$$\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})=0 \tag{16}$$

这，就是**欧拉-拉格朗日方程**，是我们确定泛函极值点的关键工具，相当于”函数在极值点处导数为零“这个工具。

观察一下欧拉-拉格朗日方程，可以知道：想要明白取什么样的函数方程时能让泛函达到极值，关键在于求出具体的泛函表达式$I=\int_{x_1}^{x_2} f(x,y,y')dx$，从而将$f(x,y,y')$代入欧拉-拉格朗日方程中解微分方程来求出函数表达式y。

2.4 例子的解决

有了欧拉-拉格朗日方程，我们就能够找到一个函数方程使得泛函达到极值了。在上文举出的例子问题中，我们已经知道连接A、B两点的曲线长度为：

$$S=\int_{x_1}^{x_2}\sqrt{1+y'^2}dx \tag{17}$$

这里面的$\sqrt{1+y'^2}$就是欧拉-拉格朗日方程中的f，于是有：

$$\begin{gathered}\frac{\partial f}{\partial y}=0\\ \frac{\partial f}{\partial y'}=\frac{y'}{\sqrt{1+y'^2}}\end{gathered} \tag{18}$$

代入欧拉-拉格朗日方程就得到：

$$0-\frac{d}{dx}(\frac{y'}{\sqrt{1+y'^2}})=0\rightarrow\frac{d}{dx}(\frac{y'}{\sqrt{1+y'^2}})=0 \tag{19}$$

所以

$$\frac{y'}{\sqrt{1+y'^2}}=C(const) \tag{20}$$

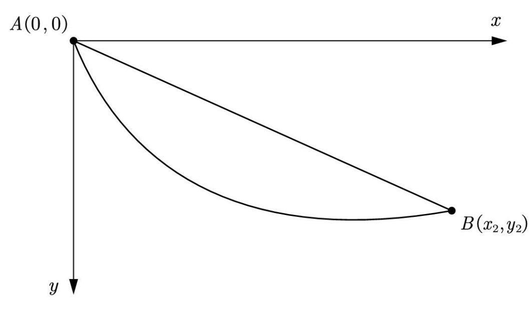
所以$y'=c(const)$，也即：$y=cx+b$，是一条直线。这样就证明了两点之间线段最短。

由此可见欧拉-拉格朗日方程的确是十分强大的吧。

2.5 欧拉-拉格朗日方程的更多应用：最速降线问题、悬链线问题

【最速降线问题】：在竖直平面上两点A、B间要用怎样的一条曲线连接，使得一个不受摩擦的质点在重力的作用下沿着这条曲线从A运动到B所需要的时间最短？

向下建系如图：



假设质点向下的竖直微小位移为y，则由动能定理有：

$$mgy=\frac{1}{2}mv^2 \tag{21}$$

所以得到此时质点的速率为$v=\sqrt{2gy}$，方向沿轨道曲线的切向。此时运动的微小路程（弧微元）为:

$$ds=\sqrt{1+y'^2}dx \tag{22}$$

所以微小用时(微小路程内可视为匀速)为:

$$dt=\frac{ds}{v}=\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}dx \tag{23}$$

所以从A到B的用时为:

$$t=\int_{x_1}^{x_2}\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}dx \tag{24}$$

把常数忽略了，就只看式子$\int_{x_1}^{x_2}\frac{\sqrt{1+y'^2}}{\sqrt{y}}dx$，则:

$$\begin{aligned}\frac{\partial f}{\partial y}&=-\frac{\sqrt{1+y'^2}}{2y\sqrt{y}}\\ \frac{\partial f}{\partial y'}&=\frac{y'}{\sqrt{y(1+y'^2)}}\\ \frac{d}{dx}(\frac{\partial f}{\partial y'})&=\frac{2yy''-y'^2-y'^4}{2y(1+y'^2)\sqrt{y(1+y'^2)}}\end{aligned} \tag{25}$$

代入欧拉-拉格朗日方程就得到:

$$-\frac{\sqrt{1+y'^2}}{2y\sqrt{y}}-\frac{2yy''-y'^2-y'^4}{2y(1+y'^2)\sqrt{y(1+y'^2)}}=0 \tag{26}$$

经过一系列化简后，最终得到:

$$1+y'^2+2yy''=0 \tag{27}$$

解这个微分方程需要一些技巧，注意到:

$$[y(1+y'^2)]'=y'+y'^3+2yy'y'' \tag{28}$$

所以第(27)式可化为:

$$\frac{[y(1+y'^2)]'}{y'}=0 \tag{29}$$

即：$[y(1+y'^2)]'=0$，所以:

$$y(1+y'^2)=C_1\rightarrow y=\frac{C_1}{1+y'^2} \tag{30}$$

令$y'=cot\alpha$，则由(30)式就有:

$$y=\frac{C_1}{1+cot\alpha^2}=\frac{C_1}{2}(1-cos2\alpha) \tag{31}$$

而又有:

$$y' = \frac{dy}{dx} \rightarrow dx = \frac{dy}{y'} = \frac{2C_1 sin\alpha cos\alpha d\alpha}{cot\alpha} = 2C_1 sin^2\alpha d\alpha \tag{32}$$

两边积分就得到:

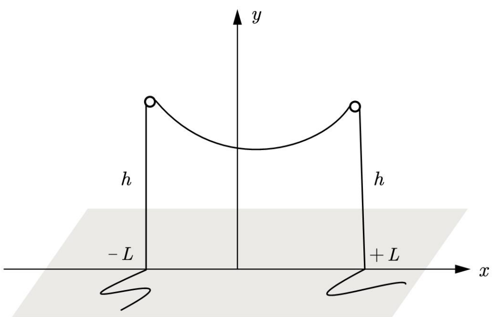
$$x = \frac{C_1}{2}(2\alpha - sin2\alpha) \tag{33}$$

令$k = \dfrac{C_1}{2}$、$t = 2\alpha$，就能得到简洁的参数方程形式:

$$\begin{aligned} x &= k(t - sint) \\ y &= k(1 - cost) \end{aligned} \tag{34}$$

是一条摆线。这就是**最速降线方程**。其中k是未知的常数，需要由具体的约束条件来求得（首末端点的坐标）。

【悬链线问题】：一条匀质软绳悬挂在等高的两点之间，在重力的作用下软绳在两悬挂点间形成的曲线应该是什么形状?



（你可能会觉得这条绳在重力的作用下会直接全部从中间掉下去，但并不一定。比如说两端绳在地面上的部分足够重，中间就掉不下来了。这具体会牵扯到绳长的分布，但已经不是我们讨论的重点，我们只把大体的形状算出来看看。举这个例子是为了和下文中另一个悬链线问题做一个对比。）

整条绳会自发的趋向势能最低的状态，所以我们从势能入手。如图建系，设绳子线密度为ρ，两悬挂点高度为h，落点横坐标为$\pm L$，那么竖直部分绳子的势能就视为不变，不予考虑，只看两悬挂点间的部分。

我们知道弧微分是:

$$ds = \sqrt{1 + y'^2}dx \tag{35}$$

所以一小段绳的重力就是$\rho g\sqrt{1 + y'^2}dx$，所以一小段绳的重力势能就是:

$$dP = \rho g y\sqrt{1 + y'^2}dx \tag{36}$$

所以绳的重力势能就是:

$$P = \int_{-L}^{L} \rho g y \sqrt{1+y'^2} dx \tag{37}$$

同样不看常数，只看式$\int_{-L}^{L} y\sqrt{1+y'^2}dx$，就有：

$$\begin{aligned} \frac{\partial f}{\partial y} &= \sqrt{1+y'^2} \\ \frac{\partial f}{\partial y'} &= \frac{yy'}{\sqrt{1+y'^2}} \\ \frac{d}{dx}(\frac{\partial f}{\partial y'}) &= \frac{y'^2+y'^4+yy''}{(1+y'^2)\sqrt{1+y'^2}} \end{aligned} \tag{38}$$

代入欧拉-拉格朗日方程得到：

$$\sqrt{1+y'^2} - \frac{y'^2+y'^4+yy''}{(1+y'^2)\sqrt{1+y'^2}} = 0 \tag{39}$$

经过一系列化简最终得到：

$$1+y'^2-yy''=0 \tag{40}$$

解这个微分方程也需要一些技巧。注意到$[\frac{y}{\sqrt{1+y'^2}}]' = \frac{y'+y'^3-yy'y''}{(1+y'^2)\sqrt{1+y'^2}}$，所以第(40)式可化为：

$$[\frac{y}{\sqrt{1+y'^2}}]' = 0 \tag{41}$$

即

$$\frac{y}{\sqrt{1+y'^2}} = C_1 \to y' = \sqrt{(\frac{y}{C_1})^2-1} \tag{42}$$

所以：

$$\frac{1}{\sqrt{(\frac{y}{C_1})^2-1}}dy = dx \tag{43}$$

令$y = C_1 cosht$，则$dy = C_1 sinhtdt$，于是第(43)式化为：$dx = C_1 dt$，所以：

$$y = C_1 cosh\frac{x-C_2}{C_1} \tag{44}$$

又因为还需满足约束条件：$y'(0)=0$，所以：$cosh^2(-\frac{C_2}{C_1}) = 1 \to C_2 = 0$，故(44)式化为

$$y = C_1 cosh\frac{x}{C_1} \tag{45}$$

是一个双曲余弦函数。

解决完了这两个例子，不难发现，似乎即便有了欧拉-拉格朗日方程，求解变分问题时还要经过化简、解二阶微分方程⋯⋯依旧并不容易。下面还需要进一步探讨。

3 欧拉-拉格朗日方程的另一种形式：第一积分（运动常数）

在繁杂的实际问题中，欧拉-拉格朗日方程使用起来有时并不简便。本节主要介绍欧拉-拉格朗日方程的一种变形形式，它时常能让问题得到简化。

3.1 推导

在上面解决最速降线问题、悬链线问题的过程中，我们用欧拉-拉格朗日方程分别求得了以下两个微分方程:

最速降线问题:

$$y'^2 + 1 + 2yy'' = 0 \tag{46}$$

悬链线问题:

$$1 + y'^2 - yy'' = 0 \tag{47}$$

接下来回顾一下我们当时的解法。

1 对于(48)式，我们是注意到:

$$[y(1 + y'^2)]' = y' + y'^3 + 2yy'y'' \tag{48}$$

从而将二阶微分方程降价为一阶微分方程求解:

$$y(1 + y'^2) = C_1 \tag{49}$$

2 对于(49)式，我们是注意到:

$$[\frac{y}{\sqrt{1 + y'^2}}]' = \frac{y' + y'^3 - yy'y''}{(1 + y'^2)\sqrt{1 + y'^2}} \tag{50}$$

从而将二阶微分方程降价为一阶微分方程求解:

$$\frac{y}{\sqrt{1 + y'^2}} = C_1 \tag{51}$$

显然用这种注意到的方法来解决问题并不是一件容易的事情。每次使用欧拉-拉格朗日方程后，都要去解一个二阶微分方程，然而一阶微分方程有时都不太好解，更何况是二阶微分方程呢？所以，即便是有了欧拉-拉格朗日方程，解决变分问题似乎也并不容易。

但，在以上两个问题中，无论我们得到的二阶微分方程是什么样子的，最后总能通过一种特殊的方法将其降价为一阶顺利求解。那么这是 一个巧合吗？或者说有没有另一种方法直接让我们得到的微分方程一开始就是一阶的呢？下面就来探讨探讨。

对于一个函数$f(y,y')$不显含x，那么对于f就有：

$$\frac{df}{dx}=\frac{\partial f}{\partial y}y'+\frac{\partial f}{\partial y'}y'' \tag{52}$$

事实上，我们在上面解决最速降线问题、悬链线问题的过程中，“注意到”的内容在某种程度上都是通过在原式两边同时乘上一个y'得到的。那么就不禁去想，如果在欧拉-拉格朗日方程两端同乘一个y'会怎样呢？

在欧拉-拉格朗日方程两端同乘一个y'得到：

$$y'\cdot\frac{\partial f}{\partial y}-y'\cdot\frac{d}{dx}(\frac{\partial f}{\partial y'})=0 \tag{53}$$

利用第(64)式将第(65)式的第一式代换：

$$\frac{df}{dx}-[y''\cdot\frac{\partial f}{\partial y'}+y'\cdot\frac{d}{dx}(\frac{\partial f}{\partial y'})]=0 \tag{54}$$

而中括号里的恰好就是$[y'\cdot\frac{\partial f}{\partial y'}]'$，于是就得到$\frac{d}{dx}(f-y'\cdot\frac{\partial f}{\partial y'})=0$，这意味着左端括号内的式子等于一个常数。

也就是：

$$f-y'\cdot\frac{\partial f}{\partial y'}=\mathcal{J} \tag{55}$$

常数$\mathcal{J}$就称作第一积分或运动常数。

这个式子有什么意义呢？

这个式子其实和欧拉-拉格朗日方程本质上是一样的，可以看成是欧拉-拉格朗日方程的另一种表达形式。但用第一积分这种形式通常能让问题简便一些，这是因为欧拉-拉格朗日方程有$\frac{d}{dx}(\frac{\partial f}{\partial y'})$这样一项，所以得到的总是二阶方程。但第一积分形式就没有这一项，所以直接能够得到一个一阶微分方程，容易解出。

3.2 举例

接下来用第一积分这种形式来解决上面的两个问题，作为对比。

【最速降线问题】：已知泛函为

$$\int_{x_1}^{x_2}\frac{\sqrt{1+y'^2}}{\sqrt{y}}dx \tag{56}$$

所以$\frac{\partial f}{\partial y'}=\frac{y'}{\sqrt{y(1+y'^2)}}$，代入第一积分表达式中得：

$$\frac{\sqrt{1+y'^2}}{\sqrt{y}}-\frac{y'^2}{\sqrt{y(1+y'^2)}}=\mathcal{J} \tag{57}$$

马上就得到了一阶微分方程:

$$\frac{1}{\sqrt{y(1+y'^2)}}=\mathcal{J} \tag{58}$$

剩下的步骤之前已经写过，不再赘述。

【悬链线问题】：已知泛函为

$$\int_{-L}^{L} y\sqrt{1+y'^2}dx \tag{59}$$

所以$\dfrac{\partial f}{\partial y'}=\dfrac{yy'}{\sqrt{1+y'^2}}$，代入第一积分:

$$\frac{y}{\sqrt{1+y'^2}}=\mathcal{J} \tag{60}$$

不再赘述。

4 解决约束条件下定端点变分问题的工具：拉格朗日乘子法

在实际问题中，往往会对许多条件加以约束，为了解决这类约束条件下的问题，需要引入拉格朗日乘子法。本文主要介绍拉格朗日乘子法，并通过具体例子，将其运用到变分问题中。

4.1 从一个例子引入

在上文中的悬链线问题里，题干所给的绳子是一条自然悬挂在两等高点的绳子。但倘若将条件改为：两端固定在等高两点而自然下垂的绳子。即:

【约束条件下的悬链线问题】一条匀质软绳两端点固定在等高的两点之间，在重力的作用下软绳在两悬挂点间形成的曲线应该是什么形状?

又该如何解决呢？这实际上是给原本的变分问题添加了一个绳长恒定的**约束条件**。

4.2 探讨如何解决约束条件下的定端点变分问题：推导拉格朗日乘子法

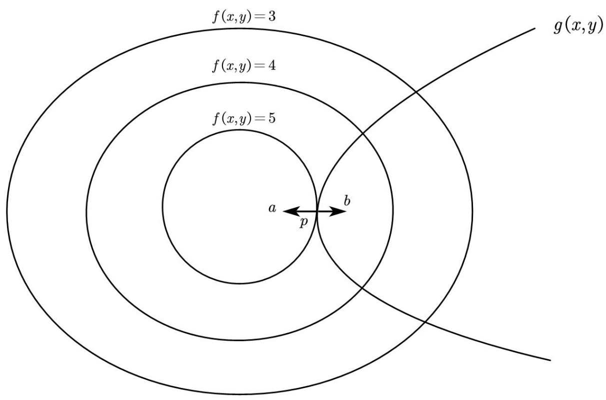
首先思考：如何确定$f(x,y)$在约束条件$g(x,y)=C$下的极值？这里的函数和约束函数都是连续的。

$f(x,y)$在约束条件$g(x,y)=C$下能够取到的点其实就是这两张空间曲面的交点，我们要在这些交点中找到极小值或极大值。

从直观上分析：

1、$f(x,y)$表示一张在$x0y0z$ 坐标系中的曲面。2、 $g(x,y)=C$表示的是一张由一条在$x0y$平面坐标系中的曲线$g(x,y)=0$沿z轴平移而成的柱面。

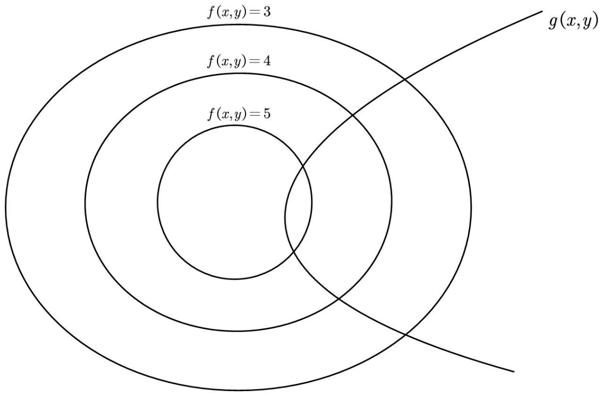
曲面$f(x,y)$在$x0y$平面中对应着一圈又一圈的等值线。将$f(x,y)$与$g(x,y)=C$一起放到$x0y$平面上来看，就如下图：



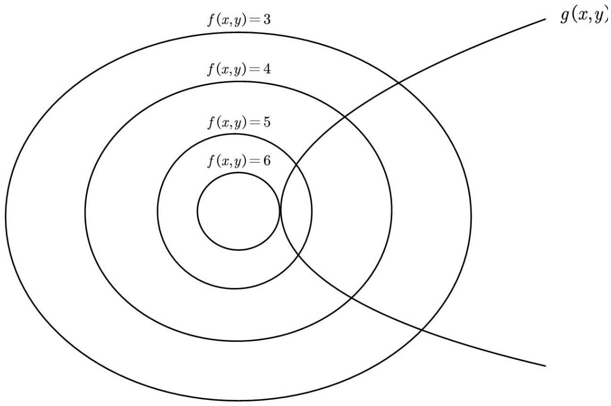
存在一条$f(x,y)$的等值线与曲线$g(x,y)=0$相切，这条等值线对应的值就是$f(x,y)$在约束条件$g(x,y)=C$下的极值。

这时，你也许会问：为什么相切的时候才是极值呢？相交时不能是极值吗？

这里我们考察的曲面都是连续的，也就意味着$f(x,y)$对应的等值线是密集的。考虑相交的情况如下图：



则自然又会存在另一条等值线与曲线$g(x,y)=0$相切，且对应的值更大（或更小）。如下图：



综上所述，在约束条件$g(x,y)=C$下，$f(x,y)$取极值时恰好与曲面$g(x,y)=C$相切，即约束条件下的极值点是函数与约束函数的切点。

用数学语言写出就是：若(x_0,y_0)是$f(x,y)$在约束条件$g(x,y)=C$下的极值点，则以下两式同时成立

$$\begin{aligned} df|_{(x_0,y_0)} &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0 \\ dg|_{(x_0,y_0)} &= \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = 0 \end{aligned} \tag{61}$$

引入向量dr= (dx,dy)，则上式改写为：

$$\begin{aligned} df|_{(x_0,y_0)} &= \nabla f|_{(x_0,y_0)} \cdot dr = 0 \\ dg|_{(x_0,y_0)} &= \nabla g|_{(x_0,y_0)} \cdot dr = 0 \end{aligned} \tag{62}$$

由于对任意dr都必须同时成立，故它们的系数一定是线性相关的（成比例），即：

$$\nabla f|_{(x_0,y_0)} = \lambda \cdot \nabla g|_{(x_0,y_0)} \rightarrow \nabla(f-\lambda\cdot g)|_{(x_0,y_0)} = 0 \tag{63}$$

比例系数λ就叫做**拉格朗日乘子**。

在上式两边同乘向量dr得到:

$$\nabla(f-\lambda\cdot g)|_{(x_0,y_0)}\cdot dr=0 \tag{64}$$

这正是函数取极值的必要条件！由此引入辅助函数:

$$F(x,y,\lambda)=f-\lambda\cdot g \tag{65}$$

则**约束条件下函数的极值问题就转化为了这个辅助函数F在没有约束条件下的极值问题了**。F不仅取决于原来的自变量，还取决于引入的系数λ。这样做就将原函数和约束条件一同放进一个新函数中了，然后就可以对新函数应用欧拉-拉格朗日方程:

$$\frac{\partial F}{\partial y}-\frac{d}{dx}(\frac{\partial F}{\partial y'})=0 \tag{66}$$

这种方法叫做**欧拉-拉格朗日乘子法**。

4.3 例子的解决

回到本节最开始提出的例子，有了拉格朗日乘子法，我们现在可以开始解决它了。

同样设单位绳元的质量为ρ，体系平衡时势能最小，故要寻找以下函数的极小值:

$$I=\int_{-L}^{L}\rho gy\sqrt{1+y'^2}dx \tag{67}$$

忽略常数，单看:

$$i=\int_{-L}^{L}y\sqrt{1+y'^2}dx \tag{68}$$

需满足约束条件—绳长为定值C: $\int_{-L}^{L}\sqrt{1+y'^2}dx=C.$

于是可以用拉格朗日乘子法写出辅助函数:

$$F(y,y',\lambda)=\int_{-L}^{L}(y-\lambda)\sqrt{1+y'^2}dx \tag{69}$$

利用上一节所讲的第一积分: $f-y'\cdot\frac{\partial f}{\partial y'}=\mathcal{J}$. 将$(y-\lambda)\sqrt{1+y'^2}$看作$f$，试着计算:

$$\begin{aligned}\mathcal{J}&=(y-\lambda)\sqrt{1+y'^2}-\frac{\partial(y-\lambda)\sqrt{1+y'^2}}{\partial y'}y'\\&=\frac{y-\lambda}{\sqrt{1+y'^2}}\end{aligned} \tag{70}$$

马上就得出了这个一阶微分方程，由此可以看出第一积分和拉格朗日乘子法的优越性。

所以:

$$\frac{1}{\sqrt{(\frac{y-\lambda}{\mathcal{J}})^2-1}}dy=dx \tag{71}$$

令$y=\lambda+\mathcal{J}cosht$，则最终解得:

$$y=\lambda+\mathcal{J}cosh\frac{x-C}{\mathcal{J}} \tag{72}$$

还要满足$y'(0)=0$，所以$y=\lambda+\mathcal{J}cosh\frac{x}{\mathcal{J}}$. 将$(\pm L,h)$代入可以求得$\mathcal{J}$、$\lambda$.

5 解决变动端点变分问题的工具：横截条件、自然边界条件

在本节之前，我们的讨论都是基于定端变分问题。本节主要介绍变动端点的变分问题，并且介绍解决方法：自然边界条件、横截条件。

5.1 从两个例子引入

在之前的讨论中，无论是有约束条件还是没有约束条件的变分问题，我们分析的都是始末端点(x_0,y_0)、(x_1,y_1)固定的泛函:

$$I(y)=\int_{x_0}^{x_1}f(x,y,y')dx \tag{73}$$

那如果给出的泛函始末端点并不固定，该怎么办呢？考虑下面这两个问题:

1、在$x-y$平面直角坐标系中，给定原点O和一条直线$x=1$，点P是该直线上的任意一点。问从点P到原点O的最速降线方程是什么？
2、在$x-y$平面直角坐标系中，给定原点O和一条直线$y=-x+2$，点P是该直线上的任意一点。问从点P到原点O的最速降线方程是什么？

这两个问题和我们最开始研究的那个最速降线问题是不同的，它不仅要分析小球在两点间怎样降落耗时最短，还要分析小球在哪两点间自然降落耗时最短。可以这样看待这个问题：在直线上任意给定一点，那么它到原点的最速降线都是摆线；但这些摆线各不相同，而我们要找的那条**“直线到原点的最速降线”**正是**“这些摆线”**中耗时最短的那条。具体怎样找，这就是本章要探讨的问题。

本章主要探讨两类变动端点的泛函取极值时应当满足的条件：1、变动端点变分问题的自然边界条件。2、变动端点变分问题的横截条件。它们分别对应着我们开篇提出的问题1和2。

5.2 推导自然边界条件

回头看看问题1，可以发现它是将泛函的右端点限制在$x=1$这条**垂直**于x轴的直线上了。本节讨论的自然边界条件，正是用来解决这类端点在垂直边界（自然边界）上变动的变分问题。

依旧讨论泛函:

$$I(y)=\int_{x_0}^{x!} f(x,y,y')dx \tag{74}$$

左端点$A(x_0,y_0)$、右端点$B(x_1,y_1)$分别沿着$x=x_0$、$x=x_1$运动。回到最初推导欧拉-拉格朗日方程时的步骤,依旧对函数y做一个微小的扰动:$Y=y+\varepsilon\delta(x)$，并将泛函视作ε的函数:

$$F(\varepsilon)=I[y+\varepsilon\delta(x)]=\int_{x_0}^{x_1} f[x,y+\varepsilon\delta(x),y'+\varepsilon\delta'(x)]dx \tag{75}$$

需满足极值条件:

$$\begin{aligned}\frac{dF(\varepsilon)}{d\varepsilon}&=\int_{x_0}^{x_1}[\frac{\partial f(x,y,y')}{\partial y}\delta(x)+\frac{\partial f(x,y,y')}{\partial y'}\delta'(x)]dx\\&=\int_{x_0}^{x_1}\frac{\partial f(x,y,y')}{\partial y}\delta(x)dx+\int_{x_0}^{x_1}\frac{\partial f(x,y,y')}{\partial y'}\delta'(x)dx\\&=\int_{x_0}^{x_1}\frac{\partial f(x,y,y')}{\partial y}\delta(x)dx+\int_{x_0}^{x_1}\frac{\partial f(x,y,y')}{\partial y'}d\delta(x)\\&=\frac{\partial f(x,y,y')}{\partial y'}\delta(x)\Big|_{x_0}^{x_1}+\int_{x_0}^{x_1}\delta(x)[\frac{\partial f(x,y,y')}{\partial y}-\frac{d}{dx}(\frac{\partial f(x,y,y')}{\partial y'})]dx\\&=\frac{\partial f}{\partial y'}\delta(x)\Big|_{x_0}^{x_1}+\int_{x_0}^{x_1}\delta(x)[\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})]dx\\&=0\end{aligned} \tag{76}$$

最开始我们推导欧拉方程时，直接将$\frac{\partial f}{\partial y'}\delta(x)\Big|_{x_0}^{x_1}$视为了0，那是因为端点$A$、$B$固定而导致$\delta(x_0)=\delta(x_1)=0$。可惜的是，现在端点并不固定了，自然而然就不能将这一项直接丢掉了。于是我们得到端点在自然边界上变动的变分问题中，泛函取极值的必要条件:

$$\frac{\partial f}{\partial y'}\delta(x)\Big|_{x_0}^{x_1}+\int_{x_0}^{x_1}\delta(x)[\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})]dx=0 \tag{77}$$

满足这一必要条件的曲线叫做容许曲线吧。注意，**变动端点泛函的容许曲线中包含了固定端点泛函的容许曲线**。现在理清下面这个逻辑：由于变动端点泛函的容许曲线中包含了固定端点泛函的容许曲线，而固定端点泛函的容许曲线是要满足欧拉-拉格朗日方程的，所以**一切变动端点泛函的容许曲线首先需要满足欧拉-拉格朗日方程**，即只要y是泛函的极值曲线，那它一定是满足欧拉-拉格朗日方程的。这是为什么呢？其实也很好理解，比方说假定我们已经找到了变动端点泛函的极值曲线$y*$，那么对于这个已经确定的$y*$来说，它的端点就是固定了的，自然也就满足了欧拉-拉格朗日方程，所以说只要y是泛函的极值曲线，那它一定是满足欧拉-拉格朗日方程的。

所以对于(77)式，将欧拉-拉格朗日方程

$$\frac{\partial f}{\partial y} - \frac{d}{dx}(\frac{\partial f}{\partial y'}) = 0 \tag{78}$$

代入就有:

$$\frac{\partial f}{\partial y'}\delta(x)\Big|_{x_0}^{x_1} + \int_{x_0}^{x_1}\delta(x)[\frac{\partial f}{\partial y} - \frac{d}{dx}(\frac{\partial f}{\partial y'})]dx = \frac{\partial f}{\partial y'}\delta(x_1) - \frac{\partial f}{\partial y'}\delta(x_0) = 0 \tag{79}$$

由于端点该变量$\delta(x_1)$、$\delta(x_2)$是相互独立、互不干扰且不为0的，所以需同时满足:

$$\begin{aligned} \frac{\partial f}{\partial y'}\Big|_{x=x_0} &= 0 \\ \frac{\partial f}{\partial y'}\Big|_{x=x_1} &= 0 \end{aligned} \tag{80}$$

这就是端点在自然边界上变动的变分问题中，泛函的极值曲线应当在两变端点(x_0, y_0)、(x_1, y_1)满足的条件，我们称之为**自然边界条件**。

总结一下，端点在自然边界上变动的泛函取极值的条件就是:

$$\begin{aligned} \frac{\partial f}{\partial y} - \frac{d}{dx}(\frac{\partial f}{\partial y'}) &= 0 \\ \frac{\partial f}{\partial y'}\Big|_{x=x_0} &= 0 \\ \frac{\partial f}{\partial y'}\Big|_{x=x_1} &= 0 \end{aligned} \tag{81}$$

5.3 推导横截边界条件

我们会发现自然边界好像是一种很理想的边界。但现实总是不理想的，无论是在数学里还是在生活中。在许多实际问题中，限制边界并不总是自然边界，比如海岸线是曲折的、肥皂泡是弯曲的……还好无论是怎样的困难，常常也能找到解决的方法。本节讲的横截条件就是用来解决端点在任意形式的曲线上变动的变分问题的。

事实上，这些变端点问题的解决方式归根结底是化变动为不变。我们在论述"**一切变动端点泛函的容许曲线首先需要满足欧拉-拉格朗日方程**"时是采用了这样的思路：先假设找到了极值曲线，这时端点就固定了，从而满足欧拉方程了。现在延续这种思路。

先假设左端点固定在$A(x_0, y_0)$，右端点$B(x_1, y_1)$在曲线$y = b(x)$上自由移动。依然考虑泛函:

$$I = \int_{x_0}^{x_1} f(x, y, y')dx \tag{82}$$

但这个时候上限x_1就不是一个固定的值了。延续最开始的方法，对y做一个微小的扰动:

$$Y = y + \varepsilon\delta(x), \delta(x_0) = 0 \tag{83}$$

（注：对于右端点$B(x_1, y_1)$，扰动之后变为了$B'(X_1, Y_1)$. 因为Y随着ϵ的变化而变化，所以Y_1随着ϵ的变化而变化，而Y_1和X_1又满足边界曲线方程，所以X_1随着ϵ变化而变化。因此右端点x_1其实可以看作关于 ε的函数$x_1(\varepsilon)$。）

依旧将泛函看成关于ε的函数:

$$F(\varepsilon) = I[y + \varepsilon\delta(x)] = \int_{x_0}^{x_1(\varepsilon)} f[x, y + \varepsilon\delta(x), y' + \varepsilon\delta'(x)]dx \tag{84}$$

则:

$$\begin{aligned}
\frac{dF(\varepsilon)}{d\varepsilon} &= x_1'(\varepsilon) \cdot f\Big|_{x=x_1(\varepsilon)} + \int_{x_0}^{x_1(\varepsilon)} [f_y \cdot \delta(x) + f_{y'} \cdot \delta'(x)]dx \\
&= \int_{x_0}^{x_1(\varepsilon)} f_{y'} d[\delta(x)] + \int_{x_0}^{x_1(\varepsilon)} [f_y \cdot \delta(x)]dx + x_1'(\varepsilon) \cdot f\Big|_{x=x_1(\varepsilon)} \\
&= f_{y'}\delta(x)\Big|_{x_0}^{x_1(\varepsilon)} - \int_{x_0}^{x_1(\varepsilon)} [\frac{df_{y'}}{dx}\delta(x)]dx + \int_{x_0}^{x_1(\varepsilon)} [f_y \cdot \delta(x)]dx + x_1'(\varepsilon) \cdot f\Big|_{x=x_1(\varepsilon)} \\
&= f_{y'}\delta(x)\Big|_{x_0}^{x_1(\varepsilon)} + x_1'(\varepsilon) \cdot f\Big|_{x=x_1(\varepsilon)} + \int_{x_0}^{x_1(\varepsilon)} [f_y - \frac{df_{y'}}{dx}]\delta(x)dx
\end{aligned} \tag{85}$$

需要满足极值条件：在$\varepsilon = 0$时$\dfrac{dI}{d\varepsilon} = 0$.即

$$\frac{dF(\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} = f_{y'}\delta(x)\Big|_{x_0}^{x_1(0)} + x_1'(0) \cdot f\Big|_{x=x_1(0)} + \int_{x_0}^{x_1(0)} [f_y - \frac{df_{y'}}{dx}]\delta(x)dx = 0 \tag{86}$$

在之前我们已经说过：**一切变动端点泛函的极值曲线首先需要满足欧拉-拉格朗日方程**。于是有:

$$f_y - \frac{df_{y'}}{dx} = 0 \to \int_{x_0}^{x_1(0)} [f_y - \frac{df_{y'}}{dx}]\delta(x)dx = 0 \tag{87}$$

所以极值条件(86)式化为:

$$f_{y'}\delta(x)\Big|_{x_0}^{x_1(0)} + x_1'(0) \cdot f\Big|_{x=x_1(0)} = 0 \tag{88}$$

由于右端点$B(x_1, y_1)$是在曲线$y = b(x)$上变动的，所以$Y = y(x) + \varepsilon\delta(x)$在右端点$B$:处一定满足:

$$y(x_1) + \varepsilon\delta(x_1) = b(x_1) \tag{89}$$

也即

$$y[x_1(\varepsilon)] + \varepsilon\delta[x_1(\varepsilon)] = b[x_1(\varepsilon)] \tag{90}$$

两边对ε求导:

$$y'\cdot x_1'(\varepsilon)+\delta[x_1(\varepsilon)]+\varepsilon\cdot\delta'[x_1(\varepsilon)]\cdot x_1'(\varepsilon)=b'\cdot x_1'(\varepsilon) \tag{91}$$

在$\varepsilon=0$时上式化为:

$$y'\cdot x_1'(0)+\delta[x_1(0)]=b'\cdot x_1'(0)\xrightarrow{\text{整理得到}}\delta[x_1(0)]=(b'-y')x_1'(0) \tag{92}$$

将(92)式代入(88)式，得:

$$\begin{aligned}f_{y'}\delta(x)\Big|_{x_0}^{x_1(0)}+x_1'(0)\cdot f\Big|_{x=x_1(0)}&=f_{y'}\delta[x_1(0)]-f_{y'}\delta(x_0)+x_1'(0)\cdot f\Big|_{x=x_1(0)}\\&=f_{y'}(b'-y')\cdot x_1'(0)+x_1'(0)\cdot f\Big|_{x=x_1(0)}\\&=[f_{y'}(b'-y')+f\Big|_{x=x_1(0)}]\cdot x_1'(0)\\&=0\end{aligned} \tag{93}$$

这里还用到了条件$\delta(x_0)=0$.

所以变动的端点的横坐标应当满足:

$$f_{y'}\cdot(b'-y')\Big|_{x=x_1(0)}+f\Big|_{x=x_1(0)}=0 \tag{94}$$

这就是这种变端点变分问题泛函的极值曲线应当在变端点(x_1,y_1)处满足的条件，我们称之为**横截条件**。

再推广到左端点$A(x_0,y_0)$也在任意形式的曲线$y=a(x)$上变动的情形，容易得出相应的横截条件为:

$$f_{y'}\cdot(a'-y')\Big|_{x=x_0(0)}+f\Big|_{x=x_0(0)}=0 \tag{95}$$

总结一下，左右两端点再任意形式的曲线上变动的泛函，取极值的条件就是:

$$\begin{aligned}&\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})=0\\&[f_{y'}\cdot(b'(x)-y')+f]\Big|_{x=x_0}=0\\&[f_{y'}\cdot(a'(x)-y')+f]\Big|_{x=x_1}=0\end{aligned} \tag{96}$$

其中$y=a(x)$、$y=b(x)$分别是泛函左右端点所在的边界曲线。

5.4 例子的解决

1、在$x-y$平面直角坐标系中，给定原点O和一条直线$x=1$，点P是该直线上的任意一点。问从点P到原点O的最速降线方程是什么？

我们先前已经求得两定点间的最速降线方程是：

$$\begin{gathered} x(t)=k(t-sint) \\ y(t)=k(1-cost) \\ f(x,y,y')=\sqrt{\frac{1+y'^2}{2gy}} \end{gathered} \tag{97}$$

现在要通过自然边界条件来解出常数k.

将最速降线方程代入自然边界条件：$\frac{\partial f}{\partial y'}\Big|_{x=1}=0$，得到：

$$\frac{\partial f}{\partial y'}\Big|_{x=1}=y'\sqrt{\frac{1}{2gy(1+y'^2)}}\Big|_{x=1}=\sqrt{\frac{(t-sint)sin^2t}{4g}} \tag{98}$$

所以：

$$(t-sint)sin^2t=0\to t=\pi(t\neq 0) \tag{99}$$

所以：

$$x_1=x(\pi)=k\pi=1\to k=\frac{1}{\pi} \tag{100}$$

所以从点P到原点O的最速降线方程是：

$$\begin{gathered} x(t)=\frac{1}{\pi}(t-sint) \\ y(t)=\frac{1}{\pi}(1-cost) \end{gathered} \tag{101}$$

2、在$x-y$平面直角坐标系中，给定原点O和一条直线$y=-x+2$，点P是该直线上的任意一点。问从点P到原点O的最速降线方程是什么？

同样有两定点间的最速降线方程：

$$\begin{gathered} x(t)=k(t-sint) \\ y(t)=k(1-cost) \\ f(x,y,y')=\sqrt{\frac{1+y'^2}{2gy}} \end{gathered} \tag{102}$$

现在要通过横截条件来解出常数k.

将最速降线方程及边界曲线$y=-x+2$代入横截条件

$$[f_{y'}\cdot(b'(x)-y')+f]\Big|_{x=x_1}=0 \tag{103}$$

得:

$$1-\frac{sint}{1-cost}=0 \tag{104}$$

故:

$$\sqrt{2}sin(t+\frac{\pi}{4})=1\rightarrow t=\frac{\pi}{2} \tag{105}$$

所以：$x_1=k(\frac{\pi}{2}-1)$、$y_1=k$，而又要满足：$y_1=-x_1+2$，所以:

$$k=-k(\frac{\pi}{2}-1)+2\rightarrow k=\frac{4}{\pi} \tag{106}$$

所以从点P到原点O的最速降线方程是:

$$\begin{aligned} x(t)&=\frac{4}{\pi}(t-sint)\\ y(t)&=\frac{4}{\pi}(1-cost) \end{aligned} \tag{107}$$

6 $Banach$空间

近代变分学中许多理论都是构建在一个完备的空间：$Banach$空间中的。本节主要简单地介绍赋范线性空间、由范数诱导出来的度量，并基于以上引出$Banach$空间。

6.1 赋范线性空间

定义：（范数）设X为线性空间，范数是定义在X上的实函数:

$$||\cdot||:X\rightarrow R \tag{108}$$

并且满足：1、【非负性】$||x||\geq 0$

2、【非退化性】$||x||=0\iff x=0$

3、【齐次性】$||\alpha x||=||\alpha||\cdot||x||$

4、【三角不等式】$||x+y||\leq||x||+||y||$

定义：（赋范线性空间）设X为线性空间，$||\cdot||$为定义在X上的范数，则$(X,||\cdot||)$为赋范线性空间。

6.2 由范数诱导出来的度量

定义：设$(X,||\cdot||)$是赋范空间，则称

$$d(x,y)=||x-y||,x、y\in X \tag{109}$$

是由范数诱导出来的度量。赋范空间的度量默认使用这种度量。

6.3 $Banach$空间

定义： 设$(X,||\cdot||)$为赋范线性空间，如果在由范数诱导出来的度量下，X是完备的，则称$(X,||\cdot||)$为$Banach$空间。

7 $Ekeland$变分原理、$Palais-Smale$条件与山路定理

本节主要介绍$Ekeland$变分原理、$Palais-Smale$条件，并给出山路定理的描述。

7.1 $Ekeland$变分原理的内涵与证明

定理：在一个完备的度量空间(X,d)中，给定一个下方有界、下半连续的 $f:X\longrightarrow R\cup\{+\infty\}, f\not\equiv+\infty$. 如果$\exists\varepsilon>0$，$\exists x_\varepsilon\in X$ 使得：$f(x_\varepsilon)<\inf\limits_X\{f\}+\varepsilon$，那么就 $\exists y_\varepsilon\in X$使得:

1、$f(y_\varepsilon)\le f(x_\varepsilon)$
2、$d(y_\varepsilon,x_\varepsilon)\le 1$
3、$f(x)>f(y_\varepsilon)-\varepsilon d(y_\varepsilon,x),\forall x\in X\backslash\{y_\varepsilon\}$

证明:

1、2可以就令$y_\varepsilon=x_\varepsilon$，则有$f(y_\varepsilon)=f(x_\varepsilon)$. 故这两条显然成立。$Ekeland$变分原理的核心在于第三条，下面给出第三条的证明思路:

3 : 为了证明第三条，我们现在递推地构造一组序列。令$u_0=x_\varepsilon$，假定已经定义了u_n，那么构造集合S_n:

$$S_n=\{w\in X|f(w)\le f(u_n)-\varepsilon d(u_n,w)\} \tag{110}$$

显然S_n不为空集（因为u_n必然属于S_n）。首先定义$u_{n+1}\in S_n$，即:

$$f(u_{n+1})\le f(u_n)-\varepsilon d(u_n,u_{n+1}) \tag{111}$$

此外还要定义u_{n+1}满足:

$$f(u_{n+1})\le\frac{f(u_n)+\inf\limits_X\{f\}}{2} \tag{112}$$

这是可以做到的。于是我们定义了u_{n+1}是满足(111)、(112)式的元素。由(111)式累加得:

$$\forall m>n:\varepsilon d(u_n,u_m)\le f(u_n)-f(u_m) \tag{113}$$

由于$f(x)$是下方有界、下半连续的，所以:

$$f(u_n)-f(u_m)\to 0(m、n\to\infty) \tag{114}$$

故序列u_n是柯西序列，必有极限。设$\underset{n\to+\infty}{lim}u_n=u^*$. 我们令$y_\varepsilon$就为$u^*$，但这么令的$y_\varepsilon$需要验证一下是否满足$Ekeland$变分原理的第1、2条。

对于第1条，在(113)式中令 $n=0$、$m\to+\infty$就能得到 $f(u^*)\le f(x_\varepsilon)-\varepsilon d(x_\varepsilon,u^*)$，所以是满足的；

对于第2条，同样在(113)式中令$n=0$、$m\to+\infty$，得到$\varepsilon d(x_\varepsilon,u^*)\le f(x_\varepsilon)-f(u^*)$，又$f(x_\varepsilon)-\underset{X}{inf}\{f\}<\varepsilon$，$f(u^*)$不会比下确界更小，故 $\varepsilon d(x_\varepsilon,u^*)\le f(x_\varepsilon)-f(u^*)<\varepsilon$，所以也是满足的。

下面证明第3条。要证明其满足第3条。**用反证法**：假设

$$\exists w\in X\backslash\{u^*\}:f(w)\le f(u^*)-\varepsilon d(u^*,w) \tag{115}$$

在(113)式中令$m\to+\infty$ 得到$f(u^*)\le f(u_n)-\varepsilon d(u_n,u^*)$. 将这个与(113)式合并：

$$f(w)\le f(u_n)-\varepsilon d(u_n,u^*)-\varepsilon d(u^*,w)\le f(u_n)-\varepsilon d(u_n,w) \tag{116}$$

现在就要用到我们定义序列u_n的第二个式子(112)式来找出矛盾了。在(112)式两端取下极限：

$$\varliminf_{n\to+\infty} f(u_{n+1})\le \varliminf_{n\to+\infty}\frac{f(u_n)+\underset{S_n}{inf}\{f\}}{2}\to f(u^*)\le \varliminf_{n\to+\infty}\underset{S_n}{inf}\{f\} \tag{117}$$

由定义可知$w\in S_n$，所以：

$$f(u^*)\le \varliminf_{n\to+\infty}\underset{S_n}{inf}\{f\}\le f(w) \tag{118}$$

但在我们假设的(115)式里 $f(w)\le f(u^*)-\varepsilon d(u^*,w)$，并且$\varepsilon d(u^*,w)>0$，所以我们的假设是：$f(w)<f(u^*)$. 这就推出了矛盾，故假设不成立。于是$y_\varepsilon=u^*$满足$Ekeland$变分原理的第3条，我们完成了第3条的证明。

综上所述，取$y_\varepsilon=u^*$，则这三条定理成立。$Ekeland$变分原理得证。

7.2 $Palais-Smale$条件

定义：设X是一个$Banach$空间，称$f\in C^1(X,R)$在值c满足PS条件（记作PS_c），如果对X中任一序列$\{x_n\}$，只要

$$1、\lim_{n\to+\infty}f(x_n)=c$$
$$2、\lim_{n\to+\infty}||f'(x_n)||=0$$

则$\{x_n\}$必有收敛子列。我们还把序列$\{x_n\}$称为一个PS序列。

7.3 $Ekeland$变分原理的一个应用

用$Ekeland$变分原理证明一个重要推论：设X是一个$Banach$空间，若可微$f : X \to R$下方有界，满足PS条件。则$\underset{X}{inf}\{f\}$是极小值。

证明：f显然下方有界、下半连续，符合$Ekeland$变分原理条件。不妨令$\underset{X}{inf}\{f\} = C$，则根据$Ekeland$变分原理，取$\varepsilon = \dfrac{1}{n}$，可以找到一串序列$\{x_n\}$使：

$$\begin{cases} f(x) > f(x_n) - \frac{1}{n}||x - x_n||, \forall x \neq q x_n (\text{根据第3条}) \\ f(x_n) < C + \frac{1}{n} (\text{根据第1条}) \end{cases} \tag{119}$$

令$x = x_n \pm th$，其中$||h|| = 1$、$t > 0$. 那么变为第一个不等式变为：

$$\frac{f(x_n \pm th) - f(x_n)}{t} \geq -\frac{1}{n} \tag{120}$$

当$t \to 0$时取极限：

$$\lim_{t\to 0} \frac{f(x_n \pm th) - f(x_n)}{t} \geq -\frac{1}{n} \to ||f'(x_n)|| \leq \frac{1}{n} \tag{121}$$

也就是：

$$\lim_{n\to +\infty} ||f'(x_n)|| = 0 \tag{122}$$

又由第二个不等式得：

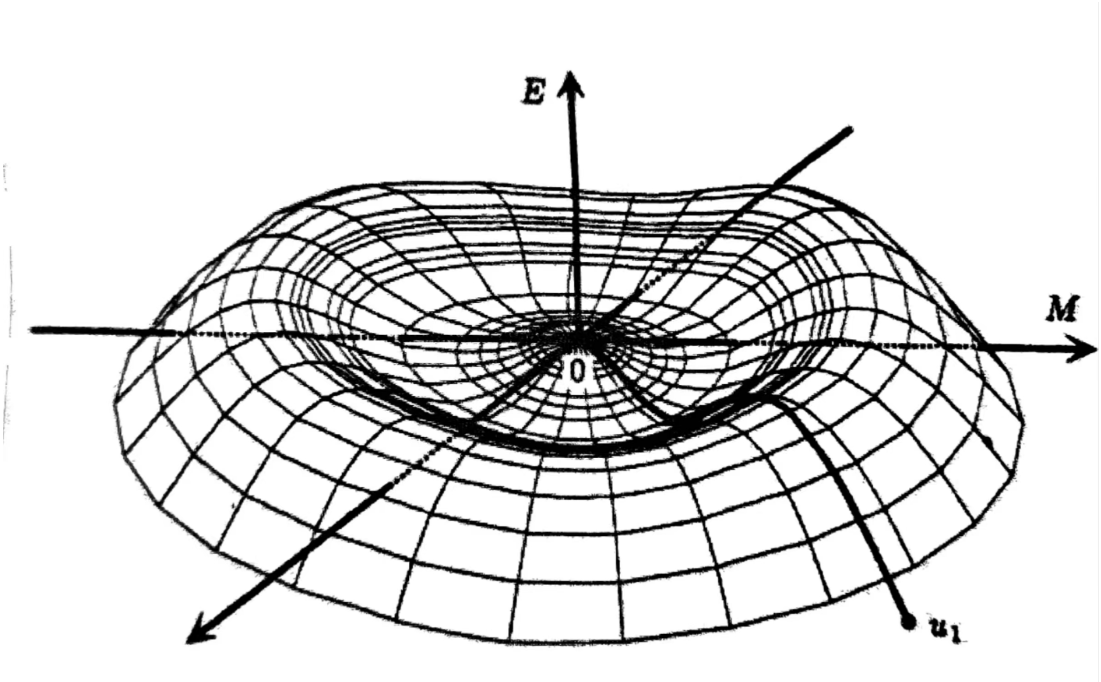
$$\lim_{n\to +\infty} f(x_n) = C \tag{123}$$

所以根据PS条件，x_n有收敛子列。令收敛子列$x_{n_j} \to x^*$，即：

$$\begin{cases} \quad f'(x^*) = 0 \\ \quad f(x^*) = C \end{cases} \tag{124}$$

所以x^*必是临界点，且$f(x^*)$达到下确界C值，所以x^*是极小值点。这就证明了$C = \underset{X}{inf}\{f\}$是极小值。

7.4 山路定理的内涵



这里就只从几何直观上来描述一下这个定理。先来看一下山路定理是干什么的。假设你在爬山，站在图中山谷里的O点处，周围一圈地势都比O点高。你想要爬到山外面的一个低处u_1点，但你又不想爬得太高，于是只能在周围这一圈高的地方中选择一个**“最低的”**来爬。那么要经过这个**“最低的”**地方，该如何选择爬行的山路呢？

在图上看，这个**“最低的”**地方比周围一圈山路都低，但又比O点和u_1点高，所以它并不是极小值或极大值点，它是一个鞍点。**山路定理就是用来判断是否存在鞍点这样不是极值点的临界点**。

容易有一个直观的想法来确定这个**鞍点**的值和寻找经过这个**鞍点**的山路：定义所在点与目的点为P_0、P_1；山谷这个区域叫做Ω，是一个开集，且$P_0 \in \Omega, P_1 \notin \overline{\Omega}$；地势高低用$f \in C^1(X, R)$来描写，函数值的分布需要满足周围一圈的地势都比所在点和目的点高，即：

$$\inf_{x \in \partial\Omega}\{f\} > max\{f(P_0), f(P_1)\} \tag{125}$$

想从P_0到P_1可以走任何的路线，把可能的**“路线”**定义为l:

$$l : [0,1] \xrightarrow{C} X, l(0) = P_0、l(1) = P_1 \tag{126}$$

把所有这些可能的道路进行比较，它们构成了一个集合，定义为：

$$\Gamma = \{l : [0,1] \xrightarrow{C} X | l(0) = P_0、l(1) = P_1\} \tag{127}$$

把每条可能的道路达到的最高点算出来，再取它们的下确界。**鞍点**的值C就是这个下确界的值了：

$$C = \inf_{l \in \Gamma}\{\max_{t \in [0,1]} f[l(t)]\} \tag{128}$$

而此时对应的l就是我们要走的山路了。

这种想法看起来很直观很正确，但其实不确切。如果想让它正确，还需要加上f满足PS条件。下面给出山路定理。

山路定理： 设X是一个$Banach$空间，$f \in C^1(X, R)$，$\Omega \subset X$是一个开集，给定两点$P_0 \in \Omega, P_1 \notin \overline{\Omega}$. 如果$f$满足$PS$条件，又设

$$\underset{x\in\partial\Omega}{inf}\{f\} > max\{f(P_0), f(P_1)\} \tag{129}$$

定义

$$\begin{gathered}\Gamma = \{l : [0,1] \xrightarrow{C} X | l(0) = P_0、l(1) = P_1\} \\ C = \underset{l\in\Gamma}{inf}\{\underset{t\in[0,1]}{max} f[l(t)]\}\end{gathered} \tag{130}$$

则C是f的一个临界值。

8 极小曲面

如果把一根铜丝弯成一条封闭的空间曲线，将它放入肥皂液中，然后轻轻提起，那么肥皂液就会在铜丝框架上张成一个处于平衡状态的薄膜。这个薄膜形成的曲面是什么样的曲面呢？本节主要运用变分法的思想原理，对这个问题进行研究，介绍极小曲面方程。

8.1 问题的表述

对于平面区域Ω，在边界$\partial\Omega$上给定一条封闭的空间曲线L，求一张定义在区域Ω上的曲面S，使得S是所有以L为边界的曲面中面积最小的曲面。

8.2 推导极小曲面方程

依照上文中解决变分问题的思路，假定已经找到了极小曲面。设极小曲面的方程为$f(x, y)$，对它做变分后得到函数：

$$z = F(x, y, \delta) \tag{131}$$

其中$(x, y, \delta) \in \Omega \times (-\varepsilon, \varepsilon)$，$\varepsilon$是一个正数，并且

$$\forall (x, y) \in \Omega:\ F(x, y, 0) = f(x, y) \tag{132}$$

而在扰动之后要求曲面边界不变，故在边界L上的所有点都要满足：

$$F(x, y, \delta) = f(x, y) \tag{133}$$

容易想到去用：

$$g(x,y)=\left.\frac{\partial F(x,y,\delta)}{\partial \delta}\right|_{\delta=0} \tag{134}$$

来刻画扰动后曲面在$\delta=0$时的变形速率，$g(x,y)$是定义在Ω上的连续可微函数。

曲面$f(x,y)$的面积为:

$$S=\int_{\Omega}\sqrt{1+f_x^2+f_y^2}dxdy \tag{135}$$

则做变分并视为δ的函数之后，得到:

$$I(\delta)=\int_{\Omega}\sqrt{1+\left(\frac{\partial F}{\partial x}\right)^2+\left(\frac{\partial F}{\partial y}\right)^2}dxdy \tag{136}$$

则取极值时的必要条件是:

$$\left.\frac{dI(\delta)}{d\delta}\right|_{\delta=0}=0 \tag{137}$$

我们的任务是要将上式左边计算出来。对$F(x,y,\delta)$关于δ做泰勒展开，结合(134)、(136)式:

$$F(x,y,\delta)=f(x,y)+\delta g(x,y)+\delta^2 m(x,y,\delta) \tag{138}$$

这里将δ二次以上的项用$\delta^2 m(x,y,\delta)$表示。于是:

$$\begin{aligned}\frac{\partial F}{\partial x}&=\frac{\partial f}{\partial x}+\delta\frac{\partial g}{\partial x}+\delta^2\frac{\partial m}{\partial x}\\ \frac{\partial F}{\partial y}&=\frac{\partial f}{\partial y}+\delta\frac{\partial g}{\partial y}+\delta^2\frac{\partial m}{\partial y}\end{aligned} \tag{139}$$

所以

$$\begin{gathered}1+\left(\frac{\partial F}{\partial x}\right)^2+\left(\frac{\partial F}{\partial y}\right)^2\\ =1+f_x^2+g_x^2+2\delta\left(f_x g_x+f_y g_y\right)+\delta^2\widetilde{m}(x,y,\delta)\end{gathered} \tag{140}$$

这里又将δ二次以上的项用$\delta^2\widetilde{m}(x,y,\delta)$表示。

将(142)提出$\left(1+f_x^2+g_x^2\right)$，变化为:

$$\left(1+f_x^2+g_x^2\right)\left[1+\delta\frac{2\left(f_x g_x+f_y g_y\right)}{1+f_x^2+g_x^2}+t^2 Q(x,y,\delta)\right] \tag{141}$$

这里已经令$Q(x,y,\delta)=\dfrac{\widetilde{m}(x,y,\delta)}{1+f_x^2+g_x^2}$，并且$Q(x,y,\delta)$是有界函数。又令:

$$P(x,y)=\frac{2\left(f_x g_x+f_y g_y\right)}{1+f_x^2+g_x^2} \tag{142}$$

则:

$$
\begin{aligned}
&\sqrt{1+\left(\frac{\partial F}{\partial x}\right)^2+\left(\frac{\partial F}{\partial y}\right)^2} \\
&=\sqrt{1+f_x^2+g_x^2}\cdot\sqrt{1+\delta P\left(x,y\right)+t^2Q\left(x,y,\delta\right)}
\end{aligned}
\tag{143}
$$

利用$\sqrt{1+a\delta}$在$\delta=0$处的泰勒展开式，得到:

$$
\begin{aligned}
&\sqrt{1+\left(\frac{\partial F}{\partial x}\right)^2+\left(\frac{\partial F}{\partial y}\right)^2} \\
&=\sqrt{1+f_x^2+g_x^2}\cdot\left[1+\frac{\delta}{2}P\left(x,y\right)+\delta^2\widetilde{Q}\left(x,y,\delta\right)\right]
\end{aligned}
\tag{144}
$$

这里又将δ二次以上的项用$\delta^2\widetilde{Q}\left(x,y,\delta\right)$表示，其中$\widetilde{Q}\left(x,y,\delta\right)$是有界函数。所以我们得到了:

$$
\left.\frac{dI\left(\delta\right)}{d\delta}\right|_{\delta=0}=\int_\Omega\frac{\left(f_xg_x+f_yg_y\right)}{\sqrt{1+f_x^2+g_x^2}}dxdy=0
\tag{145}
$$

令$p=\dfrac{\partial f}{\partial x}$、$q=\dfrac{\partial f}{\partial y}$，则(147)式中的被积表达式写成:

$$
\begin{aligned}
&\frac{p\cdot\dfrac{\partial g}{\partial x}+q\cdot\dfrac{\partial g}{\partial y}}{\sqrt{1+p^2+q^2}}=\frac{\partial}{\partial x}\left(\frac{pg}{\sqrt{1+p^2+q^2}}\right)+\frac{\partial}{\partial y}\left(\frac{qg}{\sqrt{1+p^2+q^2}}\right)- \\
&\left[\frac{\partial}{\partial x}\left(\frac{p}{\sqrt{1+p^2+q^2}}\right)+\frac{\partial}{\partial y}\left(\frac{q}{\sqrt{1+p^2+q^2}}\right)\right]\cdot g
\end{aligned}
\tag{146}
$$

对于等式右边的前两项，在Ω上的积分可以用**格林公式**:

$$
\int_L A\left(x,y\right)dx+B\left(x,y\right)dy=\int_\Omega\left(\frac{\partial B}{\partial x}-\frac{\partial A}{\partial y}\right)dxdy
\tag{147}
$$

又由于边界固定，所以$g|_L\equiv 0$. 代入必要条件，得到:

$$
\begin{aligned}
\left.\frac{dI\left(\delta\right)}{d\delta}\right|_{\delta=0}&=-\int_\Omega\left[\frac{\partial}{\partial x}\left(\frac{p}{\sqrt{1+p^2+q^2}}\right)+\frac{\partial}{\partial y}\left(\frac{q}{\sqrt{1+p^2+q^2}}\right)\right]\cdot gdxdy \\
&=0
\end{aligned}
\tag{148}
$$

由于对任意函数$g(x,y)$都要成立，所以需要满足:

$$
\frac{\partial}{\partial x}\left(\frac{p}{\sqrt{1+p^2+q^2}}\right)+\frac{\partial}{\partial y}\left(\frac{q}{\sqrt{1+p^2+q^2}}\right)=0
\tag{149}
$$

将它展开:

$$
\begin{aligned}
&\frac{\partial}{\partial x}\left(\frac{p}{\sqrt{1+p^2+q^2}}\right)+\frac{\partial}{\partial y}\left(\frac{q}{\sqrt{1+p^2+q^2}}\right) \\
&=\frac{1}{\left(1+f_x^2+f_y^2\right)^{\frac{3}{2}}}\left[\left(1+f_y^2\right)f_{xx}-2f_xf_yf_{xy}+\left(1+f_x^2\right)f_{yy}\right] \\
&=0
\end{aligned}
\tag{150}
$$

这就是**极小曲面方程**。

8.3 结论

上述推导过程可以总结为以下的结论:

对于平面区域Ω，在边界$\partial\Omega$上给定一条封闭的空间曲线L，若一张定义在区域Ω上的曲面S是所有以L为边界的曲面中面积最小的曲面，则S的函数表达式必须满足非线性二阶偏微分方程:

$$\left(1+f_y^2\right)f_{xx}-2f_xf_yf_{xy}+\left(1+f_x^2\right)f_{yy}=0 \tag{151}$$

9 参考文献

[1] 张恭庆,《变分学讲义》[M],北京:高等教育出版社,2011,9787040319583

[2] 保继光,朱汝金,唐仲伟,《变分法初步》[M],北京:高等教育出版社,2023,9787040598209

[3] 陈维桓,《极小曲面》[M],大连:大连理工大学出版社,2011,9787561161463

[4] 保继光,李海刚,《偏微分方程基础》[M],北京:高等教育出版社,2018,9787040502220

[5] 陆文端,《微分方程中的变分方法》[M],北京:科学出版社,2003,7030108612