变分问题*

2023-2024 学年秋季学期 数学分析 (III)

作为对多元函数极值问题的拓展, 本讲义简要介绍变分问题 (variational problems). 研究变分问题的理论被称为变分法/变分学 (calculus of variations).

1 变分问题的引入与形式推导

1.1 引言

变分问题是指泛函求极值问题, 这里的泛函 (functional) 是指以一个函数集合为定义域的数量值函数. 我们举一个具体的例子.

例 1.1 (引子问题)**.** 定义

$$V := \{v = v(x) \in C^1([0,1]) : v(0) = v(1) = 0\}.$$

给定 $f = f(x) \in C([0,1])$, 定义泛函

$$F(v) := \int_0^1 \frac{1}{2} v'(x)^2 + f(x)v(x)\,dx.$$

我们的目标是找出 $u \in V$, 使得它是 F 在 V 中的最小值点, 即 $F(u) \le F(v)$ 对任意 $v \in V$ 成立. 同时我们也希望研究 u 的性质.

我们也可以仅要求 u 是 V 中 F 的极小值点, 也就是说对于 "充分靠近" u 的任意的 $v \in V$, 都有 $F(u) \le F(v)$ 成立. 显然, 如果 u 是 F 在 V 中的最小值点, 它也必然是一个极小值点.

在本例中, V 是一个函数的集合, 事实上它还构成了一个无穷维的线性空间. 而泛函 $F = F(v)$ 是一个将 V 中函数映为实数的映射, 我们也常使用记号 $F = F[v]$. 利用 Riemann 积分的可积性理论和 V 的正则性假设, 可以证明 F 在 V 上是良好定义的. F 的值的具体含义依赖于所描述的问题, 通常 $F(v)$ 可以被理解为函数 v 的 "能量" 或者 "分值", 在一些和经济相关的问题中也可以被理解为 "收益" 或者 "损失". 这里我们还没有解释什么叫 v "充分靠近" u, 但这会在后面介绍理论的时候予以严格定义.

上述问题仅仅是变分学中一个非常简单的例子, 像这样的泛函极值问题在数学、科学以及实际生产生活中随处可见:

(i) 在物理中许多可观测到的体系和状态都满足 "能量" 极小原则. 例如光传播的路径是光程取极小的路径, 这被称为费马原理 (Fermat's principle). 又比如在理论力学中, 物体运动的方式由 "最小作用量原理" (the princple of least action) 来确定.

(ii) 数学中有许多由极值问题给出的重要的研究对象, 比如极小曲面 (minimal surfaces): 求空间中的一个曲面使其边界等于一个给定的集合而同时使其曲面面积最小. 想象把一个金属圈放进肥皂水

*最后更新日期：2023 年 10 月 27 日

里, 再拿出来时金属圈上会附上肥皂膜, 由于肥皂膜内部存在的张力总是希望减小膜的面积, 肥皂膜就可以被视为一个极小曲面的模型. 另一个几何中的例子是测地线 (geodesics): 给定曲面/流形上的两点, 在曲面中连接这两点的 "能量极小的" 路径就被称为测地线, 我们可以用它们来定义曲面上的度量和研究曲面的性质.

(iii) 在应用色彩更强烈的问题中也经常能见到泛函求极小的问题, 例如最优输运 (optimal transport) 问题: 想象我们需要将一堆土从一个地方搬运到另一个地方, 搬运过程中需要按照运输方案付出一定的费用, 而我们希望在完成搬运目标的同时使得费用最小. 以一维情形为例, 假设初始土堆在 $\mathbb{R}$ 中的分布由 $\mu(x)$ 给出, 目标土堆的形状由 $\nu(x)$ 给出, 我们要求 $\mu(x)$ 和 $\nu(x)$ 在 $\mathbb{R}$ 上非负且积分都是 1. 假设从 x 点搬运单位量的土到 y 点所需付出的费用为 $c(x,y)$, 比如可以取 $c(x,y)=\frac{1}{2}|x-y|^2$. 最优输运问题就是希望找出最佳的运输方案 $T:\mathbb{R}\to\mathbb{R}$, 将 x 处的土都运到 $T(x)$ 点, 要求这种运输方案能将土堆从以 $\mu(x)$ 为分布变到以 $\nu(x)$ 为分布 (数学记号是 $T_{\#}\mu=\nu$), 同时极小化运输费用

$$\mathcal{C}(T):=\int_{\mathbb{R}} c(x,T(x))\mu(x)\,dx.$$

以上是由 18 世纪法国数学家 Gaspard Monge 提出的最优输运问题的最初提法, 后来由 20 世纪苏联数学家、经济学家 Leonid V. Kantorovich[1] 改进推广后变成人们现在研究的版本. 最优输运问题在数学上和偏微分方程、概率论、几何分析等领域有着密切的关系, 它以及由它衍生出的 Wasserstein 距离等概念也在近年来深度学习和数据科学的繁荣中大放异彩.

关于变分问题, 我们可以提一些基本的问题:

- 泛函极值点是否存在?
- 是否唯一?
- 极值点具有什么样的性质? 这里的 "性质" 可以指极值点所需满足的方程和约束条件, 它的光滑性, 可能存在的奇点的形态, 以及其他在具体问题中人们特别关心的性质.

相较于多元函数极值问题, 变分问题可以视为一个无穷维的优化问题, 因而回答以上这些问题并不容易, 有的甚至需要泛函分析等进阶的知识作为铺垫. 因此, 本讲义将首先介绍在假设极值点的存在性和光滑性的条件下, 如何 (形式地) 推导出极值点所满足的方程. 而这些推导背后的数学理论以及极值点的存在性、唯一性等问题, 我们将尽可能严谨地予以阐述, 但必要时仅做形式化的简要描述. 在阅读的过程中, 读者们可以对照多元函数极值问题的理论, 或许会更容易理解下面的内容.

1.2 推导极值点满足的方程

首先来研究例 1.1 引子问题. 我们假设该问题具有一个极小值点 $u\in V$, 并且假设 u 充分光滑, 比如 $u\in C^2([0,1])$. 我们的目标是推导出 u 所满足的方程.

主要的想法如下. 我们在极小值点 u 的基础上引入扰动: 取一个 "可容许的" 函数 η 作为扰动方向, 考虑 $u+t\eta$, 这里 $t\in\mathbb{R}$. 在选取 η 时需保证 $u+t\eta\in V$. 当 $|t|\ll 1$ 时, $u+t\eta$ 应该会 "十分靠近" u (这里暂时不解释什么叫 "十分靠近", 但在直观上接受这一点并不难). 于是由 u 的极小性, $F(u)\le F(u+t\eta)$ 对任意的 $|t|\ll 1$ 成立, 因此 $t\mapsto F(u+t\eta)$ 这个一元数量值函数在 $t=0$ 处取极小. 所以,

1. 如果函数 $t\mapsto F(u+t\eta)$ 在 $t=0$ 处可导, 那么 $\frac{d}{dt}\big|_{t=0}F(u+t\eta)=0$. 值得注意的是, 这可以被理解为泛函 F 在 u 处沿着方向 η 的 "方向导数" 是 0;

[1] Kantorovich 在 1975 年因 "对稀缺资源的最优配置的研究" 与 Tjalling C. Koopmans 分享了诺贝尔经济学奖. 他也是线性规划的奠基人.

2. 如果函数 $t \mapsto F(u+t\eta)$ 在 $t=0$ 处二阶可导, 那么 $\frac{d^2}{dt^2}\big|_{t=0} F(u+t\eta)=0$.

这样我们就得到了可以计算的等式, 在此基础上就有可能推导出 u 所满足的方程.

具体到例 1.1 中的问题, 由于 V 是线性空间, 我们只需令 $\eta \in V$ 即可保证 $u+t\eta \in V$ 对任意 $t \in \mathbb{R}$ 成立. 这样, 我们计算

$$\begin{aligned}
0 = &\ \frac{d}{dt}\bigg|_{t=0} F(u+t\eta) \\
= &\ \frac{d}{dt}\bigg|_{t=0} \int_0^1 \frac{1}{2}\big(u'(x)+t\eta'(x)\big)^2 + f(x)\big(u(x)+t\eta(x)\big)\,dx \\
= &\ \frac{d}{dt}\bigg|_{t=0} \left[\int_0^1 \frac{1}{2}u'(x)^2 + f(x)u(x)\,dx + t\int_0^1 u'(x)\eta'(x) + f(x)\eta(x)\,dx + \frac{t^2}{2}\int_0^1 \eta'(x)^2\,dx\right],
\end{aligned}$$

所以

$$0 = \int_0^1 u'(x)\eta'(x) + f(x)\eta(x)\,dx.$$

由于以上推导和 η 的具体选取无关, 所以上式对任意 $\eta \in V$ 都成立. 进一步地, 由分部积分得

$$0 = \int_0^1 (-u''(x) + f(x))\eta(x)\,dx. \tag{1}$$

这里没有边界项是因为 $\eta \in V$ 给出了 $\eta(0)=\eta(1)=0$. 此式也对于任意的 $\eta \in V$ 成立. 由于我们假设了 $f \in C([0,1])$ 且 $u \in C^2([0,1])$, 故 $-u''+f \in C([0,1])$, 因此由 η 在 V 中的任意性, 可以证明 (作为练习)

$$-u''(x) + f(x) = 0, \quad \forall\, x \in (0,1). \tag{2}$$

这就是 u 需满足的方程. 再利用 $u \in V$ 给出的边界条件 $u(0)=u(1)=0$, 我们不难推导出 u 的表达式. 具体的计算留给读者完成.

另一个条件 $\frac{d^2}{dt^2}\big|_{t=0} F(u+t\eta)=0$ 对应于 $\int_0^1 \eta'(x)^2\,dx \geq 0$. 这是自然满足的.

以上推导中的 (1) 式表示的是泛函 F 在 u 处沿着方向 η 的 "方向导数" 为 0. 由于多元微积分中方向导数可以由梯度向量和方向向量做内积得到, 而 (1) 式的右边在形式上恰好是 $(-u''+f)$ 和方向 η 的函数 L^2 内积[2], 因此将 $(-u''+f)$ 理解为泛函 F 在 u 处的 "梯度" 看起来十分自然. 事实上, 我们之后会引入记号 $\frac{\delta F}{\delta v}$ 来表示泛函 F 在一个一般点 $v \in V$ 处的 "梯度", 称为 F 的一阶变分 (the first variation). 在本例中, $\frac{\delta F}{\delta v} = -v'' + f$. 因此, (2) 式表达的就是 $\frac{\delta F}{\delta v}\big|_{v=u} = 0$, 这被称为原变分问题的 *Euler-Lagrange* 方程, 它对应于多元函数极值问题中的 "极值点处梯度为零" 的条件.

1.3 更多的例子

上面对于例 1.1 的讨论已经展示了推导泛函极值点所满足的方程 (即 Euler-Lagrange 方程) 的关键想法. 本小节中我们来研究更多的例子.

例 1.2. 定义

$$V := \{v = v(x) \in C^1([a,b]) :\, v(a) = c_1,\, v(b) = c_2\}.$$

[2] 一般地, $[a,b]$ 两个上平方可积的函数 $g=g(x)$ 和 $h=h(x)$ 的 L^2 内积定义为 $\langle g,h\rangle := \int_a^b g(x)h(x)\,dx$. 我们之后会讨论到, 称 (1) 式右端是两个函数的 L^2 内积并不合适, 更准确的说法应该是 $(-u''+f)$ 和 η 的对偶或者配对.

这里 $c_1, c_2 \in \mathbb{R}$ 为给定常数. 另外给定一个 C^2 函数 $h = h(x, v, p) : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. 已知 $u \in V$ 是泛函

$$F(v) := \int_a^b h(x, v(x), v'(x))\, dx$$

在 V 中的极小值点. 我们假设 $u \in C^2([a, b])$. 请推导 u 所满足的 Euler-Lagrange 方程.

解. 仿照上面的方法, 希望在 u 的基础上进行扰动得到 $u + t\eta \in V$. 注意到 V 不是线性空间, 但仍具有线性结构[3], 即对任意的 $v_* \in V$, $\{v - v_* : v \in V\}$ 是一个线性空间. 故定义

$$V_0 := \{v = v(x) \in C^1([a, b]) : v(a) = v(b) = 0\}.$$

这样, 对于任意的 $\eta \in V_0$ 和任意的 $t \in \mathbb{R}$, 都有 $u + t\eta \in V$. 此外, 当 $|t| \ll 1$ 时, $u + t\eta$ "十分靠近" u. 由 u 的极小性, 我们得到

$$\left.\frac{d}{dt}\right|_{t=0} F(u + t\eta) = 0.$$

于是

$$\left.\frac{d}{dt}\right|_{t=0} \int_a^b h(x, u(x) + t\eta(x), u'(x) + t\eta'(x))\, dx = 0.$$

我们希望将求导和积分进行交换, 在 h 的光滑性假设下, 这一论证并不难. 比如由关于 t 的 Taylor 展开,

$$\begin{aligned}
&h(x, u(x) + t\eta(x), u'(x) + t\eta'(x)) - h(x, u(x), u'(x)) \\
={}& t \cdot \left.\frac{d}{ds}\right|_{s=0} \big[h(x, u(x) + s\eta(x), u'(x) + s\eta'(x))\big] + \frac{t^2}{2} R(x, t),
\end{aligned}$$

其中 $R(x, t)$ 关于 x 连续 (由上式即可证), 且存在 $\theta_{x,t} \in (0, 1)$, 使得

$$\begin{aligned}
R(x, t) ={}& \left.\frac{d^2}{ds^2}\right|_{s=t\theta_{x,t}} \big[h(x, u(x) + s\eta(x), u'(x) + s\eta'(x))\big] \\
={}& \eta(x)^2 \cdot \frac{\partial^2 h}{\partial v^2}(x, u(x) + t\theta_{x,t}\eta(x), u'(x) + t\theta_{x,t}\eta'(x)) \\
&+ 2\eta(x)\eta'(x) \cdot \frac{\partial^2 h}{\partial v \partial p}(x, u(x) + t\theta_{x,t}\eta(x), u'(x) + t\theta_{x,t}\eta'(x)) \\
&+ \eta'(x)^2 \cdot \frac{\partial^2 h}{\partial p^2}(x, u(x) + t\theta_{x,t}\eta(x), u'(x) + t\theta_{x,t}\eta'(x)).
\end{aligned}$$

当 $|t| \le 1$ 时, 可以推出估计 $|R| \le C$. 这里 C 是一个仅依赖于函数 h, η, u 的常数[4], 特别地它与 t 无关. 因此, 由导数的定义,

$$\begin{aligned}
&\left.\frac{d}{dt}\right|_{t=0} \int_a^b h(x, u(x) + t\eta(x), u'(x) + t\eta'(x))\, dx \\
={}& \lim_{t\to 0} \frac{1}{t} \int_a^b h(x, u(x) + t\eta(x), u'(x) + t\eta'(x)) - h(x, u(x), u'(x))\, dx \\
={}& \int_a^b \left.\frac{d}{ds}\right|_{s=0} \big[h(x, u(x) + s\eta(x), u'(x) + s\eta'(x))\big]\, dx + \lim_{t\to 0} \int_a^b \frac{t}{2} R(x, t)\, dx
\end{aligned}$$

[3] V 可以被称为一个更大的线性空间 $C^1([a, b])$ 中的一个线性流形, 它可以通过 $C^1([a, b])$ 中的平移成为它的一个线性子空间, 即上面定义的 V_0.

[4] 更具体地说, 常数 C 依赖于 $M := \sup_{x\in[a,b]} |\eta(x)| + \sup_{x\in[a,b]} |\eta'(x)|$, $N := \sup_{x\in[a,b]} |u(x)| + \sup_{x\in[a,b]} |u'(x)|$ 以及 $\sup_{(x,v,p)\in D} |\nabla^2 h(x, v, p)|$, 其中 $D = [a, b] \times [-M - N, M + N] \times [-M - N, M + N]$.

$$= \int_a^b \eta(x) \cdot \frac{\partial h}{\partial v}(x, u(x), u'(x)) + \eta'(x) \cdot \frac{\partial h}{\partial p}(x, u(x), u'(x))\, dx.$$

利用分部积分和边界条件 $\eta(0) = \eta(1) = 0$, 我们得到

$$\int_a^b \eta(x) \left[-\frac{d}{dx} \left(\frac{\partial h}{\partial p}(x, u(x), u'(x)) \right) + \frac{\partial h}{\partial v}(x, u(x), u'(x)) \right] dx = 0$$

对任意 $\eta \in V_0$ 成立. 由于假设了 u 和 h 的 C^2 正则性, 上面括号中的式子为连续函数, 因此由 η 的任意性知

$$-\frac{d}{dx} \left(\frac{\partial h}{\partial p}(x, u(x), u'(x)) \right) + \frac{\partial h}{\partial v}(x, u(x), u'(x)) = 0, \quad \forall\, x \in (a, b). \tag{3}$$

此外 u 需满足边界条件 $u(a) = c_1$, $u(b) = c_2$. □

注记. (3) 为本例中所讨论的具有较为一般形式的泛函 F 所对应的 Euler-Lagrange 方程. (3) 式的左端也可被记为 $\frac{\delta F}{\delta v}\big|_{v=u}$.

有时由极小性不仅能导出极小值点需满足的区域内的方程, 也可以推出它在边界上的约束.

例 1.3. 给定 $c \in \mathbb{R}$. 在

$$V := \{v = v(x) \in C^1([0, 1]) : v(0) = c\}$$

中找一个函数 $u(x)$ 使得它的图像在 $[0, 1]$ 上的总长最小.

解. 任取 $v \in V$, $v = v(x)$ 的函数图像在 $[0, 1]$ 上的总长度为

$$L(v) := \int_0^1 \sqrt{1 + v'(x)^2}\, dx.$$

我们假设上述最小值点 $u \in V$ 存在唯一且 C^2 光滑. 选取扰动空间为

$$V_0 := \{v = v(x) \in C^1([0, 1]) : v(0) = 0\}.$$

这样对任意的 $\eta \in V_0$ 和任意的 $t \in \mathbb{R}$, $u + t\eta \in V$. 与上面的过程类似地, $\frac{d}{dt}\big|_{t=0} L(u + t\eta) = 0$. 利用 L 的表达式, 我们进一步地推导

$$\begin{aligned} 0 = & \left.\frac{d}{dt}\right|_{t=0} \int_0^1 \sqrt{1 + (u'(x) + t\eta'(x))^2}\, dx \\ = & \int_0^1 \left.\frac{d}{dt}\right|_{t=0} \sqrt{1 + (u'(x) + t\eta'(x))^2}\, dx \\ = & \int_0^1 \frac{u'(x)}{\sqrt{1 + u'(x)^2}} \cdot \eta'(x)\, dx. \end{aligned}$$

第二行中交换导数和积分的过程可以类似于上面例子那样进行论证. 由分部积分以及 $\eta(0) = 0$ 得,

$$\begin{aligned} 0 = & \left. \left[\frac{u'(x)}{\sqrt{1 + u'(x)^2}} \cdot \eta(x) \right] \right|_{x=0}^{1} - \int_0^1 \frac{d}{dx} \left(\frac{u'(x)}{\sqrt{1 + u'(x)^2}} \right) \eta(x)\, dx \\ = & \frac{u'(1)}{\sqrt{1 + u'(1)^2}} \cdot \eta(1) - \int_0^1 \frac{d}{dx} \left(\frac{u'(x)}{\sqrt{1 + u'(x)^2}} \right) \eta(x)\, dx. \end{aligned} \tag{4}$$

此式对任意 $\eta \in V_0$ 都成立.

定义 $\tilde{V}_0 := \{v \in C^1([a,b]) : v(0) = v(1) = 0\}$. 由 (4) 式知, 对任意的 $\eta \in \tilde{V}_0$,

$$0 = -\int_0^1 \frac{d}{dx}\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}}\right)\eta(x)\,dx.$$

由于我们假设 $u \in C^2$, 故 $(u'/\sqrt{1+u'^2})'$ 连续, 所以由 η 的任意性得

$$\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}}\right)' = 0, \quad \forall x \in (0,1). \tag{5}$$

将它代回 (4) 知, 对任意的 $\eta \in V_0$,

$$\frac{u'(1)}{\sqrt{1+u'(1)^2}} \cdot \eta(1) = 0.$$

由于 $\eta(1)$ 不一定为 0, 故必须要求 $u'(1) = 0$. 这样, 我们就得到了 u 必须满足 (5), 且 $u(0) = c$, $u'(1) = 0$. 由此可解得 $u(x) \equiv c$. □

下面我们以历史上著名的悬链线问题为例来讨论带约束的变分问题.

例 1.4 (悬链线问题)**.** 考虑一条长度为 $l \in (0,2)$ 的均匀链子, 两端点分别悬于 $(-1,0)$ 和 $(1,0)$. 问在重力作用下链子的稳态形状是什么样的.

讨论. 不妨假设链子的形状可以由 $[-1,1]$ 的一个 (充分光滑的) 函数 $y = f(x)$ 的图像表示, 其中 $f(-1) = f(1) = 0$. 于是由链子总长的条件

$$L(f) := \int_{-1}^1 \sqrt{1+f'(x)^2}\,dx = l. \tag{6}$$

另一方面, 链子在重力作用下呈现出使它自身的重力势能最小的形状. 我们知道链子的重力势能 (如果假设链子的密度为 1) 可以表示为

$$E(f) := \int_{-1}^1 f(x)\sqrt{1+f'(x)^2}\,dx. \tag{7}$$

此式中 $f(x)$ 表示链子的高度, $\sqrt{1+f'(x)^2}\,dx$ 为链子的长度微元因而也是质量微元. 因此, 求链子稳态形状的问题就归结为在约束 (6) 下求泛函 (7) 的最小值点的问题. 我们可以限定在

$$V := \{v = v(x) \in C^1([-1,1]) : v(-1) = v(1) = 0\}$$

中寻找最小值点.

求解这类带约束的变分问题依然可以用 Lagrange 乘子法, 具体做法和求解带约束的多元函数极值问题非常相似. 可定义 Lagrange 增广函数 $F(f,\lambda) := E(f) + \lambda(L(f) - l)$. 在适当的条件下 (我们略去不讨论), 原问题的条件极值点 $f_* \in V$ 需满足存在 $\lambda_* \in \mathbb{R}$, 使得

$$\frac{\delta F}{\delta f}(f_*,\lambda_*) = 0, \quad 且 \quad \frac{\partial F}{\partial \lambda}(f_*,\lambda_*) = 0.$$

不难发现后一式即为约束 $L(f_*) - l = 0$ 本身.

解. 我们假设最小值点 $f_* \in V$ 存在, 唯一, 且 C^2 光滑. 依照上面讨论中的方法, 我们来直接计算 $\frac{\delta F}{\delta f}(f_*,\lambda_*)$. 由于 V 是线性空间, 我们可以取任意 $\eta \in V$ 作为扰动, 就能够保证对任意的 $t \in \mathbb{R}$, 都有 $f_* + t\eta \in V$. 因此由 f_* 的极小性, 我们可以仿照前面的例子计算得到

$$0 = \left.\frac{d}{dt}\right|_{t=0} F(f_* + t\eta, \lambda_*)$$

$$
\begin{aligned}
&= \int_{-1}^{1} \eta(x)\sqrt{1+f_*'(x)^2} + (f_*(x)+\lambda_*)\frac{f_*'(x)\eta'(x)}{\sqrt{1+f_*'(x)^2}}\,dx \\
&= \int_{-1}^{1} \eta(x)\left(\sqrt{1+f_*'(x)^2} - \frac{d}{dx}\left[(f_*(x)+\lambda_*)\frac{f_*'(x)}{\sqrt{1+f_*'(x)^2}}\right]\right)dx.
\end{aligned}
$$

此式对任意 $\eta \in V$ 成立. 由于假设了 $f \in C^2([-1,1])$, 上式括号中的部分连续. 由 η 的任意性,

$$
\sqrt{1+f_*'(x)^2} - \frac{d}{dx}\left[(f_*(x)+\lambda_*)\frac{f_*'(x)}{\sqrt{1+f_*'(x)^2}}\right] = 0, \quad \forall\, x \in (-1,1). \tag{8}
$$

此式即为 $\frac{\delta F}{\delta f}(f_*,\lambda_*) = 0$. 此外, f_* 还需满足

$$
f_*(-1) = f_*(1) = 0, \qquad \int_{-1}^{1} \sqrt{1+f_*'(x)^2}\,dx = l. \tag{9}
$$

下面来求解这个方程. 将 (8) 中的导数展开后得到

$$
\sqrt{1+f_*'(x)^2} = \frac{f_*'(x)^2 + (f_*(x)+\lambda_*)f_*''(x)}{\sqrt{1+f_*'(x)^2}} - (f_*(x)+\lambda_*)f_*'(x)\frac{f_*'(x)f_*''(x)}{(1+f_*'(x)^2)^{3/2}}.
$$

将右边化简后得

$$
\sqrt{1+f_*'(x)^2} = \frac{f_*'(x)^2}{\sqrt{1+f_*'(x)^2}} + \frac{(f_*(x)+\lambda_*)f_*''(x)}{(1+f_*'(x)^2)^{3/2}}.
$$

两边同时乘以 $\sqrt{1+f_*'(x)^2}$ 并化简得

$$
1 = \frac{(f_*(x)+\lambda_*)f_*''(x)}{1+f_*'(x)^2}.
$$

将它进一步写为

$$
\frac{f_*''(x)f_*'(x)}{1+f_*'(x)^2} = \frac{f_*'(x)}{f_*(x)+\lambda},
$$

也就是

$$
\frac{1}{2}\frac{d}{dx}\big[\ln(1+f_*'(x)^2)\big] = \frac{d}{dx}\big[\ln|f_*(x)+\lambda|\big].
$$

因此存在一个常数 $C > 0$ 使得

$$
\sqrt{1+f_*'(x)^2} = C(f_*(x)+\lambda_*).
$$

做变量替换 $g(y) := C(f_*(C^{-1}y)+\lambda_*)$, 那么 $g'(y) = f_*'(C^{-1}y)$, 所以由上式可推出

$$
g'(y)^2 = g(y)^2 - 1, \quad \forall\, y \in (-C, C).
$$

通过求 $s \mapsto 1/\sqrt{s^2-1}$ 的原函数, 我们可以求解这一方程得到

$$
g(y) = \cosh(y + \tilde{C}).
$$

这里 $\tilde{C} \in \mathbb{R}$ 为待定常数. 所以

$$
f_*(x) = C^{-1}\cosh(Cx + \tilde{C}) - \lambda_*.
$$

利用边界条件 $f(-1) = f(1) = 0$ 不难解得

$$
\tilde{C} = 0, \quad \lambda_* = C^{-1}\cosh C.
$$

于是

$$f_*(x) = C^{-1}\big[\cosh(Cx) - \cosh C\big]. \tag{10}$$

最后, 为了确定常数 C, 我们利用 (9) 中关于弧长的条件得

$$l = \int_{-1}^{1} \sqrt{1 + f_*'(x)^2}\, dx = \int_{-1}^{1} \sqrt{1 + \sinh^2(Cx)}\, dx = \int_{-1}^{1} \cosh(Cx)\, dx = \frac{2}{C}\sinh C.$$

这一方程没有显式解, 但我们不难证明存在唯一的 $C > 0$ 满足该方程. 将这一常数 C 代入 (10), 我们就确定了悬链线的形状. □

2 关于变分问题的严格理论

本节中我们将讨论如何将上一节的形式推导严格化, 并简述与变分问题极值点的存在性、唯一性相关的理论.

2.1 泛函分析中的一些基本概念

我们首先介绍一些基本概念. 我们假设读者已经学习过线性代数中线性空间的概念, 不过我们也提醒, 下面讨论的线性空间不一定是有限维的.

定义 2.1 (线性空间上的范数)**.** 假设 X 为一个实数域或者复数域上的线性空间. X 上一个满足如下条件的函数 $\|\cdot\|_X : X \to \mathbb{R}$, $x \mapsto \|x\|_X$ (当不会引起混淆时我们将省略下标, 将它简记为 $\|\cdot\|$) 被称为 X 上的一个范数 (norm, 也称为模):

(1) (正定性) $\|x\| \geq 0$ 对任意 $x \in X$ 成立, 而且 $\|x\| = 0$ 当且仅当 $x = 0$;

(2) (三角不等式) 对任意的 $x, y \in X$, 都有 $\|x + y\| \leq \|x\| + \|y\|$;

(3) (数乘性质) 对任意的 $x \in X$, $\alpha \in \mathbb{F}$, 都有 $\|\alpha x\| = |\alpha|\|x\|$. 这里的 $\mathbb{F}$ 为定义线性空间时的数域, 我们总是假设 $\mathbb{F} = \mathbb{R}$ 或者 $\mathbb{F} = \mathbb{C}$.

例 2.1. (i) $\mathbb{R}^n$ 中通常的向量的欧氏模长 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的一个范数. 这十分容易验证. 事实上, 以上范数的概念就是通常的欧氏向量模长的推广.

(ii) 令 $C([0,1])$ 表示 $[0,1]$ 上全体实值连续函数构成的集合. 不难验证它是一个实线性空间. 对于 $f \in C([0,1])$, 定义

$$\|f\| := \sup_{x\in[0,1]} |f(x)|.$$

则 $\|f\|$ 是 $C([0,1])$ 上的一个范数, 一般记作 $\|f\|_{C([0,1])}$. 证明留作练习.

(iii) 定义平方可和的实数列 (复数列也可以) 的集合

$$l^2 := \left\{ \mathbf{a} = (a_1, \cdots, a_n, \cdots) : \, a_k \in \mathbb{R}, \, \sum_{k=1}^{\infty} |a_k|^2 < \infty \right\}.$$

定义 l^2 上的函数

$$\|\cdot\|_p : \, \mathbf{a} = (a_1, \cdots, a_n, \cdots) \mapsto \left(\sum_{k=1}^{\infty} |a_k|^p \right)^{\frac{1}{p}}.$$

可以证明当 $p \geq 2$ 时, $\|\cdot\|_p$ 是 l^2 上的一个范数. 特别地, $\|\cdot\|_2$ 可以被视为是有限维欧氏空间上通常向量长度概念的推广[5].

然而当 $p \in (0,2)$ 时, $\|\cdot\|_p$ 不是 l^2 上的范数. 这是因为这时候 $\|\cdot\|_p$ 中的无穷求和并不总是收敛的, 因此对某些 l^2 中的元素无法定义这个函数的值. 比如, 考虑数列

$$\mathbf{a} = (1^{-\alpha}, 2^{-\alpha}, \cdots, n^{-\alpha}, \cdots).$$

对于给定的 $p \in (0,2)$, 选取 $\frac{1}{2} < \alpha \leq \frac{1}{p}$. 不难验证这样定义的 $\mathbf{a}$ 属于 l^2, 但是无穷求和 $\sum_{k=1}^{\infty} |a_k|^p$ 不收敛.

(本讲义未完待续)

[5] 之后我们还会指出, 从使得 l^2 成为一个完备度量空间的角度来说, $\|\cdot\|_2$ 也是很特殊的.