第五讲：变分法基础

最优控制的数学理论之一

张杰

人工智能学院
中国科学院大学

复杂系统管理与控制国家重点实验室
中国科学院自动化研究所

2017 年 9 月 26 日

Table of Contents

Table of Contents

最简变分问题

问题 1 (最简变分问题)

求函数 $x(t) : [t_0, t_f] \to \mathbb{R}^n$，在给定的初始和终端时刻 t_0, t_f 满足，$x(t_0) = x_0$, $x(t_f) = x_f$，且最小化性能指标

$$J(x) = \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t)\,\mathrm{d}t. \tag{1}$$

其中 g 取值于 $\mathbb{R}$，二阶连续可微。

函数 $x(t) : [t_0, t_f] \to \mathbb{R}^n$ 及其导数 $\dot{x}(t) : [t_0, t_f] \to \mathbb{R}^n$ 记为

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \dot{x}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix},$$

最优控制问题

问题 2 (最优控制问题)

1. 被控对象的状态方程为

$$\dot{x}(t) = f(x(t), u(t), t), \quad x(t_0) = x_0.$$

2. 容许控制, $u \in \mathcal{U}, \quad x \in \mathcal{X}$.
3. 目标集, $x(t_f) \in \mathcal{S}$

$$\mathcal{S} = [t_0, \infty) \times \{x(t_f) \in \mathbb{R}^n : m(x(t_f), t_f) = 0\}$$

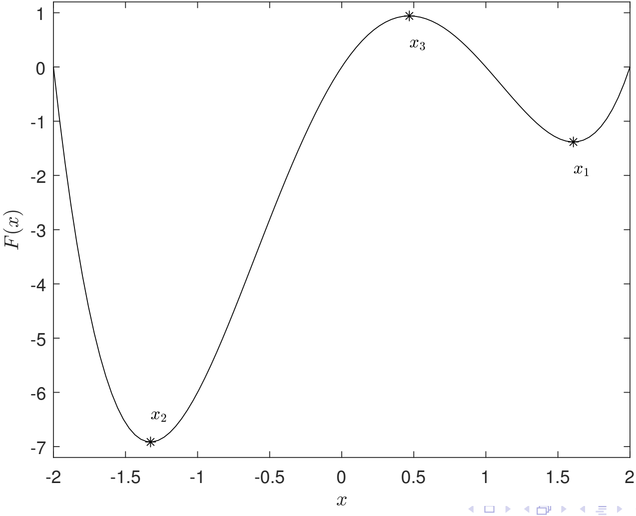
4. 求分段连续的 u, 以最小化性能指标

$$J(u) = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)\, \mathrm{d}t.$$

Table of Contents

局部极值和全局极值

局部极小：x 的 δ-邻域内的 x' 都有 $F(x) \leq F(x')$。全局：$\delta = \infty$



泰勒展开

用泰勒公式对函数多项式近似

$$F(x+\Delta x)=F(x)+\Delta x^{\mathrm{T}}\nabla F(x)+\frac{1}{2}\Delta x^{\mathrm{T}}\nabla^2 F(x)\Delta x+\dots$$

$$\nabla F(x)=\begin{bmatrix}\frac{\partial F}{\partial x_1}(x)\\ \vdots\\ \frac{\partial F}{\partial x_n}(x)\end{bmatrix},\quad \nabla^2 F(x)=\begin{bmatrix}\frac{\partial^2 F}{\partial x_1^2}(x) & \dots & \frac{\partial^2 F}{\partial x_1\partial x_n}(x)\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 F}{\partial x_n\partial x_1}(x) & \dots & \frac{\partial^2 F}{\partial x_n^2}(x)\end{bmatrix}.$$

最优的必要条件（一阶条件）：$\nabla F(x)=0$

最简变分问题

问题 3 (最简变分问题)

求函数 $x(t):[t_0,t_f]\to\mathbb{R}^n$，在给定的初始和终端时刻 t_0,t_f 满足，$x(t_0)=x_0$, $x(t_f)=x_f$，且最小化性能指标，关于 x 的泛函

$$J(x)=\int_{t_0}^{t_f} g(x(t),\dot{x}(t),t)\,\mathrm{d}t.$$

其中 g 取值于 $\mathbb{R}$，二阶连续可微。

函数 $x(t):[t_0,t_f]\to\mathbb{R}^n$ 及其导数 $\dot{x}(t):[t_0,t_f]\to\mathbb{R}^n$ 记为

$$x(t)=\begin{bmatrix} x_1(t)\\ \vdots \\ x_n(t)\end{bmatrix},\quad \dot{x}(t)=\begin{bmatrix} \dot{x}_1(t)\\ \vdots \\ \dot{x}_n(t)\end{bmatrix},$$

泛函的定义

定义 1 (泛函, functional)

从任意集合 M 到实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$ 的映射称为"泛函";

在变分学和最优控制中，泛函定义域为函数集合，只取实值

例 1 (泛函)

$t_1 \in \mathbb{R}$, 集合 M 为 $\mathbb{R} \to \mathbb{R}$ 的一阶连续可微函数全体, 对于 $u \in M$

$$I_1(u) = \max_{t \in \mathbb{R}} |u(t)|, \quad I_2(u) = u(t_1), \quad I_3(u) = \int_0^{+\infty} u^2(t)dt.$$

都是关于函数 u 的泛函。

定义：函数的范数与距离

定义 2 (范数)

范数 $\|\cdot\|$ 可用于空间中两点的距离 $d(x,x') = \|x - x'\|$

- 正定: $\|x\| \geq 0$, 且 $\|x\| = 0$ iff $x = 0$
- 齐次: $\|ax\| = |a|\|x\|$, $a \in \mathbb{R}$
- 次可加: $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$

Remark 1

函数空间中的范数是欧氏空间范数的抽象

$$\|x\| = (\sum_{i=1}^{n} |x_i|^2)^{1/2}.$$

定义：一些函数空间和范数

定义 3 (连续函数的范数)

设 Ω 为 $[a,b]$ 到 $\mathbb{R}$ 的连续函数全体。对 $x \in \Omega$ 可定义范数 $\|\cdot\|_0$：

$$\|x\|_0 = \max_{t\in[a,b]} |x(t)|. \tag{2}$$

定义 4 (连续可微函数的范数)

设 Ω 为 $[a,b]$ 到 $\mathbb{R}$ 连续可微函数全体。对 $x \in \Omega$ 可定义范数 $\|\cdot\|_1$：

$$\|x\|_1 = \max_{t\in[a,b]} |x(t)| + \max_{t\in[a,b]} |\dot{x}(t)|. \tag{3}$$

从函数极值到泛函极值

定义 5 (泛函极小值)

定义域 M 是一类函数的集合。泛函 $J(x): M \to \mathbb{R}$ 在 x 达到局部极小值，若存在 $\epsilon > 0$ 使得:

$$J(x) \le J(x'), \text{ if } \|x' - x\| < \epsilon, x \in M.$$

Remark 2 (函数极小值)

定义域 $\Omega \subset \mathbb{R}^N$ 是开集。函数 $F(x): \Omega \to \mathbb{R}$ 在 x 达到局部极小值，若存在 $\epsilon > 0$ 使得:

$$F(x) \le F(x'), \text{ if } \|x' - x\| < \epsilon, x \in \Omega.$$

从函数增量到泛函增量

定义 6 (泛函增量)

若函数 x 和 $x+\delta x$ 都在泛函 J 的定义域中，定义泛函增量 ΔJ 为:

$$\Delta J := J(x+\delta x) - J(x). \tag{4}$$

或记为 $\Delta J(x, \delta x)$ 以强调泛函增量与函数 x 和 δx 有关

Remark 3 (函数增量)

若实数 x 和 $x+\Delta x$ 都在函数 F 的定义域中，$F(x)$ 取得函数小极值，函数的增量 $\Delta F(x, \Delta x) \geq 0$

$$\Delta F(x, \Delta x) = F(x+\Delta x) - F(x) \approx \dot{F}(x)\Delta x + o(|\Delta x|) \tag{5}$$

例子：求泛函增量

例 2 (求泛函增量)

求泛函 $J : C[t_0, t_f] \to \mathbb{R}$ 的增量

$$J(x) = \int_{t_0}^{t_f} x^2(t) dt.$$

解：求泛函增量

$$
\begin{aligned}
J(x) &= \int_{t_0}^{t_f} x^2(t)dt, \\
\Delta J(x, \delta x) &= J(x+\delta x) - J(x) \\
&= \int_{t_0}^{t_f} [x(t) + \delta x(t)]^2 dt - \int_{t_0}^{t_f} x^2(t)dt \\
&= \int_{t_0}^{t_f} 2x(t)\delta x(t)dt + \int_{t_0}^{t_f} [\delta x(t)]^2 dt.
\end{aligned}
$$

增量可表示成两个部分，前者与 δx 成“线性”关系；后者则是 δx 的高阶无穷小项

定义：线性泛函

定义 7 (线性泛函)

称 J 是 x 的线性泛函，若其满足“齐次性, homogeneity”条件:

$$J(ax) = aJ(x), \forall a \in \mathbb{R}, x \in \Omega, ax \in \Omega$$

以及“可加性, additivity”条件

$$J(x_1 + x_2) = J(x_1) + J(x_2), \forall x_1, x_2, x_1 + x_2 \in \Omega.$$

例 3 (线性泛函)

可以验证，下列关于 δx 的泛函是线性的

$$\int_{t_0}^{t_f} 2x(t)\delta x(t)dt$$

函数变分与泛函的变分

定义 8 (泛函的变分)

若泛函增量可写为函数变分的线性泛函和高阶无穷小两部分:

$$\Delta J(x, \delta x) = \delta J(x, \delta x) + g(x, \delta x)\|\delta x\|, \tag{6}$$

则称 δJ 是 J 对于 x 的变分，称 J 对 x 可微

$$\lim_{\|\delta x\| \to 0} \{g(x, \delta x)\} = 0$$

Remark 4

泛函变分可与函数导数作为对比，函数的增量与之十分类似

$$\Delta F(x, \Delta x) = \dot{F}(x)\Delta x + o(|\Delta x|)$$

例子：求泛函变分

例 4 (求泛函变分)
求泛函 $J: C[t_0, t_f] \to \mathbb{R}$ 的变分

$$J(x) = \int_{t_0}^{t_f} x^2(t)dt.$$

$$\Delta J(x, \delta x) = \int_{t_0}^{t_f} 2x(t)\delta x(t)dt + \int_{t_0}^{t_f} [\delta x(t)]^2 dt.$$

只需证明，后项为高阶无穷小项

泛函变分—高阶无穷小 1/2

取函数范数

$$\|\delta x\| = \|\delta x\|_0 = \max_{t\in[t_0,t_f]} \{|\delta x(t)|\},$$

$$\begin{aligned}
\int_{t_0}^{t_f} [\delta x(t)]^2 \,\mathrm{d}t &= \|\delta x\|_0 \cdot \int_{t_0}^{t_f} |\delta x(t)| \cdot \frac{|\delta x(t)|}{\|\delta x\|_0} \,\mathrm{d}t \\
&= \|\delta x\|_0 \cdot \int_{t_0}^{t_f} |\delta x(t)| \cdot \frac{|\delta x(t)|}{\max_{t\in[t_0,t_f]}\{|\delta x(t)|\}} \,\mathrm{d}t \\
&\le \|\delta x\|_0 \cdot \int_{t_0}^{t_f} |\delta x(t)| \,\mathrm{d}t \\
&\le \|\delta x\|_0 \cdot \int_{t_0}^{t_f} \max_{t\in[t_0,t_f]}\{|\delta x(t)|\} \,\mathrm{d}t \\
&= \|\delta x\|_0 \cdot \int_{t_0}^{t_f} \|\delta x\|_0 \,\mathrm{d}t = \|\delta x\|_0 \cdot [t_f - t_0] \|\delta x\|_0.
\end{aligned}$$

泛函变分—高阶无穷小 2/2

若 $\|\delta x\|_0 \to 0$，则：

$$\lim_{\|\delta x\|_0 \to 0} [t_f - t_0] \|\delta x\|_0 = 0.$$

即后项为高阶无穷小得证。于是求得，泛函 J 对 x 的变分为：

$$\delta J(x, \delta x) = \int_{t_0}^{t_f} 2x(t)\delta x(t)\,\mathrm{d}t. \tag{7}$$

利用微积分方法计算变分

引理 1 (利用微积分方法计算变分)

若泛函 J 对函数 x 可微，则可计算泛函变分如下：

$$\delta J(x, \delta x) = \frac{\mathrm{d}}{\mathrm{d}\alpha} J(x + \alpha\delta x)\Big|_{\alpha=0}, \quad \alpha \in \mathbb{R}. \tag{8}$$

Remark 5

作为对比

$$J(x + \alpha\delta x) = J(x) + \frac{\mathrm{d}}{\mathrm{d}\alpha} J(x + \alpha\delta x) + o(\alpha)$$
$$F(x + \Delta x) = F(x) + \dot{F}(x)\Delta x + o(|\Delta x|)$$

使用本引理可很容易计算泛函的变分

证明：通过求导计算变分

求泛函增量即可证明.

$$\Delta J(x,\delta x) = J(x+\delta x) - J(x) = \delta J(x,\delta x) + g(x,\delta x)\cdot \|\delta x\|$$

$$\begin{aligned} J(x+\alpha\delta x) - J(x) &= \delta J(x,\alpha\delta x) + g(x,\alpha\delta x)\cdot \|\alpha\delta x\| \\ &= \alpha\delta J(x,\delta x) + g(x,\alpha\delta x)\cdot \|\alpha\delta x\| \end{aligned}$$

$$\begin{aligned} \frac{d}{d\alpha} J(x+\alpha\delta x)\Big|_{\alpha=0} &= \lim_{\alpha\to 0} \frac{J(x+\alpha\delta x) - J(x)}{\alpha} \\ &= \delta J(x,\delta x) + \lim_{\alpha\to 0} \frac{g(x,\alpha\delta x)\cdot \|\alpha\delta x\|}{\alpha} \\ &= \delta J(x,\delta x) + \lim_{\alpha\to 0} g(x,\alpha\delta x)\cdot \frac{|\alpha|\|\delta x\|}{\alpha} \\ &= \delta J(x,\delta x). \end{aligned}$$

例子：泛函变分的简便求解

例 (求泛函变分)

求泛函 $J : C[t_0, t_f] \to \mathbb{R}$ 的变分

$$J(x) = \int_{t_0}^{t_f} x^2(t)dt.$$

$$\begin{aligned}
\delta J(x, \delta x) &= \frac{d}{d\alpha} J(x + \alpha\delta x)\Big|_{\alpha=0} \\
&= \frac{d}{d\alpha} \int_{t_0}^{t_f} (x + \alpha\delta x)^2 dt\Big|_{\alpha=0} \\
&= \int_{t_0}^{t_f} [2(x + \alpha\delta x)\delta x] dt\Big|_{\alpha=0} = \int_{t_0}^{t_f} 2x(t)\delta x(t)dt.
\end{aligned}$$

定理：泛函极值一阶条件

定理 1 (泛函极值一阶条件)

$x \in M$，M 是一类函数的开集合，泛函 J 对 x 可微。若 x 使 J 取极值，则对任意容许的 δx 有

$$\delta J(x, \delta x) = 0. \tag{9}$$

【其中容许的 δx 指，若 $x \in \Omega$ 则 $x + \delta x \in \Omega$】

Remark 6

泛函变分可对比函数导数，上述泛函极值的一阶条件则可对比函数极值的一阶条件：$\forall \Delta x$，需满足 $\dot{F}(x)\Delta x = 0$

简要证明

若 x 是最优解，而 $\delta J(x,\delta x)\neq 0$，不妨设其大于零。由线性性，

$$\delta J(x,+\alpha\delta x)=+\alpha\delta J(x,\delta x)>0. \tag{10}$$
$$\delta J(x,-\alpha\delta x)=-\alpha\delta J(x,\delta x)<0. \tag{11}$$

$$\Delta J(x,\alpha\delta x)=\delta J(x,\alpha\delta x)+g(x,\alpha\delta x)\|\alpha\delta x\|.$$

$$\Delta J(x,+\alpha\delta x)>0. \tag{12}$$
$$\Delta J(x,-\alpha\delta x)<0. \tag{13}$$

不是极小也不是极大，矛盾！

例子：根据泛函极值一阶条件求泛函极值

例 5 (求泛函极值)

求泛函 $J : C[t_0, t_f] \to \mathbb{R}$ 极小值点

$$J(x) = \int_{t_0}^{t_f} x^2(t)dt.$$

若 x 使 J 取极值，则对任意容许的 δx，

$$0 = \delta J(x, \delta x) = \int_{t_0}^{t_f} 2x(t)\delta x(t)dt. \tag{14}$$

容易证明，$2x(t) = 0$

使用经典变分法求泛函极值的基本过程

- 求泛函变分（根据定义求，或引理 1）
- 求解泛函极值条件（根据定理 1，“导数为零”）

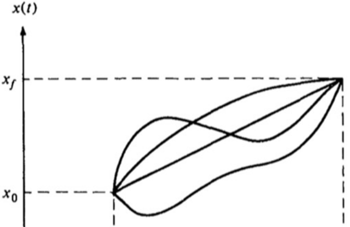
最简变分问题

问题 (最简变分问题)

求函数 $x(t):[t_0,t_f]\to\mathbb{R}^n$，在给定的初始和终端时刻 t_0,t_f 满足，$x(t_0)=x_0$, $x(t_f)=x_f$，且最小化性能指标，关于 x 的泛函

$$J(x)=\int_{t_0}^{t_f} g(x(t),\dot{x}(t),t)\,\mathrm{d}t.$$

其中 g 取值于 $\mathbb{R}$，二阶连续可微。



1/5 计算泛函增量

$$\begin{aligned}\Delta J(x,\delta x) &= J(x+\delta x) - J(x)\\ &= \int_{t_0}^{t_f} g(x(t)+\delta x(t), \dot{x}(t)+\delta\dot{x}(t), t)dt\\ &\quad - \int_{t_0}^{t_f} g(x(t),\dot{x}(t),t)dt.\end{aligned}$$

$$\frac{\partial g}{\partial x}(x,\dot{x},t) = \begin{bmatrix} \frac{\partial g}{\partial x_1}(x,\dot{x},t) \\ \vdots \\ \frac{\partial g}{\partial x_n}(x,\dot{x},t) \end{bmatrix}, \quad \frac{\partial g}{\partial \dot{x}}(x,\dot{x},t) = \begin{bmatrix} \frac{\partial g}{\partial \dot{x}_1}(x,\dot{x},t) \\ \vdots \\ \frac{\partial g}{\partial \dot{x}_m}(x,\dot{x},t) \end{bmatrix}.$$

2/5 化简泛函增量，得到泛函变分

将该积分在 $x(t), \dot{x}(t)$ 泰勒展开

$$\Delta J = \int_{t_0}^{t_f} \{g(x(t), \dot{x}(t), t) + \frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) \cdot \delta x(t) + \frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) \cdot \delta \dot{x}(t) + o(\| \cdot \|) - g(x(t), \dot{x}(t), t)\} dt.$$

得到泛函的变分

$$\delta J(x, \delta x, \dot{x}, \delta \dot{x}) = \int_{t_0}^{t_f} \{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) \cdot \delta x(t) + \frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) \cdot \delta \dot{x}(t)\} dt$$

3/5 也可通过微积分计算泛函变分

$$
\begin{aligned}
J(x) &= \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t) dt. \\
\delta J &= \frac{d}{d\alpha} J(x + \alpha \delta x) \Big|_{\alpha=0} \\
&= \frac{\partial}{\partial \alpha} \int_{t_0}^{t_f} g(x(t) + \alpha \delta x(t), \dot{x}(t) + \alpha \delta \dot{x}(t), t) dt \Big|_{\alpha=0} \\
&= \int_{t_0}^{t_f} \{ \frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) \cdot \delta x(t) + \frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) \cdot \delta \dot{x}(t) \} dt
\end{aligned}
$$

4/5 分部积分去掉变分和导数之间的依赖

$\delta x(t)$ 和 $\delta \dot{x}(t)$ 有导数关系

$$\delta \dot{x}(t) = \frac{d}{dt}\delta x(t)$$

使用分部积分公式

$$\begin{aligned}
&\delta J(x, \delta x, \dot{x}, \delta \dot{x}) \\
&= \int_{t_0}^{t_f} \{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) \cdot \delta x(t) + \frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) \cdot \delta \dot{x}(t)\} dt \\
&= \int_{t_0}^{t_f} \{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{d}{dt}[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)]\} \cdot \delta x(t) dt \\
&\quad + [\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) \cdot \delta x(t)]\Big|_{t_0}^{t_f}.
\end{aligned}$$

5/5 泛函极值一阶条件

由 $x(t_0), x(t_f)$ 已经给定，有 $\delta x(t_0) = \delta x(t_f) = 0$，于是

$$\delta J(x, \delta x) = \int_{t_0}^{t_f} \left\{ \frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}, t)\right] \right\} \cdot \delta x(t)\, \mathrm{d}t.$$

泛函极值必要条件是对于任意容许的 δx，

$$\delta J(x, \delta x) = 0$$

可得 J 取极值的必要条件——欧拉-拉格朗日方程

$$\boxed{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\right] = 0. \qquad (15)}$$

欧拉-拉格朗日方程的特殊情况求解

$$\frac{\partial g}{\partial x}(x(t),\dot{x}(t),t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial g}{\partial \dot{x}}(x(t),\dot{x}(t),t)\right] = 0.$$

$$\frac{\partial g}{\partial x} - \frac{\partial^2 g}{\partial \dot{x}\partial x}\dot{x} - \frac{\partial^2 g}{\partial \dot{x}\partial \dot{x}}\ddot{x} - \frac{\partial^2 g}{\partial \dot{x}\partial t} = 0$$

二阶方程化简为一阶?

- Case 1. No $\dot{x}$, i.e. $g = g(x,t)$.
- Case 2. No x, i.e. $g = g(\dot{x},t)$.
- Case 3. No t, i.e. $g = g(x,\dot{x})$.

欧拉-拉格朗日方程的特殊情况之一

问题 4 (Case 1. No $\dot{x}$: $g = g(x,t)$)

$$\frac{\partial g}{\partial x}(x(t),\dot{x}(t),t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial g}{\partial \dot{x}}(x(t),\dot{x}(t),t)\right] = 0$$
$$g = g(x,t).$$

可得

$$\frac{\partial g}{\partial x}(x(t),\dot{x}(t),t) = 0 \tag{16}$$

例子：根据欧拉-拉格朗日方程求泛函极值

例 (求泛函极值)

求泛函 $J : C[t_0, t_f] \to \mathbb{R}$ 的极值

$$J(x) = \int_{t_0}^{t_f} x^2(t)dt.$$

$$\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) = 0 \Rightarrow 2x(t) = 0 \Rightarrow x(t) = 0.$$

欧拉-拉格朗日方程的特殊情况之二

问题 5 (Case 2. No x: $g = g(\dot{x}, t)$)

$$\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\Big] = 0$$
$$g = g(\dot{x}, t).$$

得到

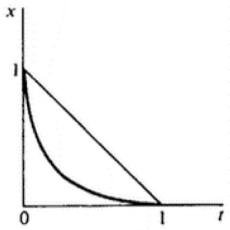
$$\begin{aligned} \frac{d}{dt}\Big[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\Big] &= 0 \\ \frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t) &= c \end{aligned} \tag{17}$$

例子：两点之间线段最短

例 6 (两点之间线段最短)

如图，起止两点，$t_0 = 0, t_f = 1, x(t_0) = 1, x(t_f) = 0$。求两点最短的连线。即，求函数 x 满足上述边界条件且最小化性能指标

$$J(x(t)) = \int_{t_0}^{t_f} \sqrt{1 + \dot{x}^2(t)}dt. \tag{18}$$



例子：两点之间线段最短

$g(x,\dot{x},t)=\sqrt{1+\dot{x}^2}$ 是 $\dot{x}$ 的函数，No x

$$\frac{\dot{x}}{(1+\dot{x}^2)^{1/2}}=c_1 \tag{19}$$

$$\dot{x}=\pm\sqrt{c_1^2/(1-c_1^2)} \tag{20}$$

$$x=\pm\sqrt{c_1^2/(1-c_1^2)}t+c_2. \tag{21}$$

带入初值终值 $x(0)=1,x(1)=0$，$x=-t+1$，为连接两点的线段

欧拉-拉格朗日方程的特殊情况之三

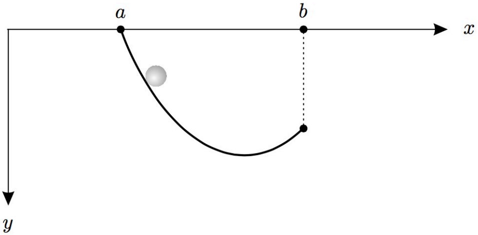
问题 6 (Case 3. No t: $g = g(x, \dot{x})$)

$$
\begin{aligned}
&\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\right] = 0 \\
&\frac{\partial g}{\partial x} - \frac{\partial^2 g}{\partial \dot{x} \partial x}\dot{x} - \frac{\partial^2 g}{\partial \dot{x} \partial \dot{x}}\ddot{x} - \frac{\partial^2 g}{\partial \dot{x} \partial t} = 0 \\
&g = g(x, \dot{x}).
\end{aligned}
$$

$$
\begin{aligned}
0 &= \frac{\partial g}{\partial x} - \frac{\partial^2 g}{\partial \dot{x} \partial x}\dot{x} - \frac{\partial^2 g}{\partial \dot{x} \partial \dot{x}}\ddot{x} \\
0 &= \dot{x}\left[\frac{\partial g}{\partial x} - \frac{\partial^2 g}{\partial \dot{x} \partial x}\dot{x} - \frac{\partial^2 g}{\partial \dot{x} \partial \dot{x}}\ddot{x}\right] = \frac{d}{dt}\left[\frac{\partial g}{\partial \dot{x}}\dot{x} - g\right] \\
&\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\dot{x} - g(x(t), \dot{x}(t), t) = c \qquad (22)
\end{aligned}
$$

例子：最速降线问题

例 7 (最速降线问题, Brachistochrome Curve Problem)



1/3 例子：最速降线问题

$$T(y) = \int_a^b \frac{\sqrt{1 + [dy/dx]^2}}{\sqrt{2gy}} dx$$

被积函数没有 t 的情况

$$\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\dot{x} - g(x(t), \dot{x}(t), t) = c$$

$$\frac{[dy/dx]^2}{\sqrt{2gy(1 + [dy/dx]^2)}} - \sqrt{\frac{1 + [dy/dx]^2}{2gy}} = c$$

$$y = \frac{1}{2gc^2(1 + [dy/dx]^2)}$$

2/3 例子：最速降线问题

$$y = \frac{1}{2gc^2(1 + [dy/dx]^2)}$$

使用参数法，令 $dy/dx = \mathrm{ctg}(\theta)$

$$y = \frac{1}{2gc^2(1 + \mathrm{ctg}^2(\theta))} = \frac{\sin^2(\theta)}{2gc^2}$$

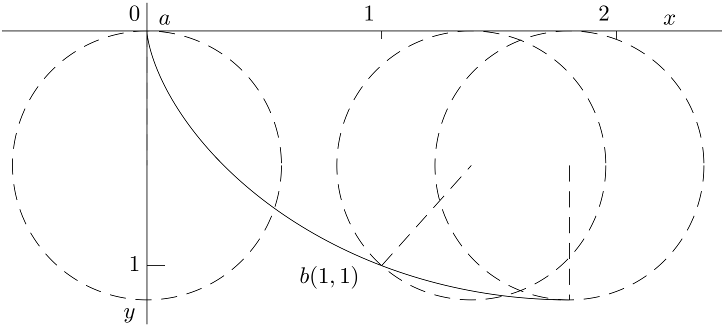
$$dx = \frac{1}{\mathrm{ctg}(\theta)} \frac{2\sin(\theta)\cos(\theta)}{2gc^2} d\theta = \frac{\sin^2(\theta)}{gc^2} d\theta = \frac{1 - \cos(2\theta)}{2gc^2} d\theta$$

$$x = \frac{2\theta - \sin(2\theta)}{4gc^2} + c_1$$

引入 $\alpha = 2\theta$，

$$x(\alpha) = r[\alpha - \sin\alpha], \quad y(\alpha) = r[1 - \cos\alpha].$$

3/3 最速降线问题的解：旋轮线



$a = (0, 0)$, $b = (1, 1)$，可得待定系数 $r \approx 0.573$，

$$x(\alpha) = r[\alpha - \sin\alpha],\ y(\alpha) = r[1 - \cos\alpha].$$

Table of Contents

欧拉-拉格朗日方程与哈密尔顿函数

$$J(x) = \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t) dt.$$

以状态变量的“方向”为“控制变量” $u(t)$

$$\begin{cases} \dfrac{\partial g}{\partial x}(x(t), u(t), t) - \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\dfrac{\partial g}{\partial u}(x(t), u(t), t)\right] = 0, \\ u(t) = \dot{x}(t). \end{cases} \tag{23}$$

引入哈密尔顿函数

定义哈密尔顿函数:

$$\mathcal{H}(x,u,p,t) := g(x,u,t) + p \cdot f(x,u,t). \tag{24}$$

通过对哈密尔顿函数简单计算，立即可得:

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial p} &= u \\ \frac{\partial \mathcal{H}}{\partial x} &= \frac{\partial g}{\partial x} \\ \frac{\partial \mathcal{H}}{\partial u} &= p + \frac{\partial g}{\partial u}. \end{aligned}$$

取 $x(t), u(t)$ 为上述一阶常微分方程组-(23) 的解，再定义该最优控制问题的协态（costate）为:

$$p(t) := -\frac{\partial g}{\partial u}(x(t),u(t),t).$$

计算规范方程

可得状态和协态的导数分别为：

$$\dot{x}(t) = u(t) = \frac{\partial \mathcal{H}}{\partial p}(x(t), \dot{x}(t), p(t), t),$$
$$\dot{p}(t) = \frac{\mathrm{d}}{\mathrm{d}t}[-\frac{\partial g}{\partial u}(x(t), u(t), t)].$$

由欧拉-拉格朗日方程-(23)，协态变量的导数可进一步化简为使用哈密尔顿函数表示的：

$$\begin{aligned}\dot{p}(t) &= -\frac{\partial g}{\partial x}(x(t), u(t), t)\\ &= -\frac{\partial \mathcal{H}}{\partial x}(x(t), u(t), t).\end{aligned}$$

以及：

$$\frac{\partial \mathcal{H}}{\partial u}(x(t), u(t), t) = p(t) + \frac{\partial g}{\partial u}(x(t), u(t), t) = 0.$$

变分问题的“极小值原理”

$$0 = \frac{\partial \mathcal{H}}{\partial u} \tag{25}$$

$$\dot{x} = +\frac{\partial \mathcal{H}}{\partial p} \tag{26}$$

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x}. \tag{27}$$

变分问题与最优控制问题：区别

	变分问题	停车例子	导弹例子
状态方程	无	$\dot{x} = f(x, u, t)$	$\dot{x} = f(x, u, t)$
目标	x_f,t_f fix	x_f,t_f fix	x_f free, t_f fix
性能指标	$J(x)$	$J(u)$	$J(u)$

Remark 7 (经典变分求最优控制所需)

- 需处理约束条件
- 需要处理不同的控制目标（边界条件）

Table of Contents

微分方程约束的泛函极值

问题 7 (微分方程约束的泛函极值)

$x(t)$ 初值终值 $x(t_0) = x_0, x(t_f) = x_f$，$t_f$ *fixed*，需要满足约束

$$F(x(t), \dot{x}(t), t) = 0 \tag{28}$$

最小化性能指标

$$J(x) = \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t) dt \tag{29}$$

F 是 l 维约束

1/4 拉格朗日乘子法

最小化 Lagrangian

$$\bar{J}(x,p)=\int_{t_0}^{t_f}\left[g(x(t),\dot{x}(t),t)+p(t)\cdot F(x(t),\dot{x}(t),t)\right]\mathrm{d}t. \tag{30}$$

记

$$\bar{g}(x(t),\dot{x}(t),p(t),t)\overset{\text{def}}{=}g(x(t),\dot{x}(t),t)+p(t)\cdot F(x(t),\dot{x}(t),t),$$

2/4 E-L 方程

$$
\begin{aligned}
0 =& \frac{\partial \bar{g}}{\partial x_1}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{x}_1}(x(t), \dot{x}(t), p(t), t)\Big], \\
& \dots \\
0 =& \frac{\partial \bar{g}}{\partial x_n}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{x}_n}(x(t), \dot{x}(t), p(t), t)\Big], \\
0 =& \frac{\partial \bar{g}}{\partial p_1}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{p}_1}(x(t), \dot{x}(t), p(t), t)\Big], \\
& \dots \\
0 =& \frac{\partial \bar{g}}{\partial p_l}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{p}_l}(x(t), \dot{x}(t), p(t), t)\Big].
\end{aligned}
$$

3/4 简单整理

前 n 个方程可写成

$$0 = \frac{\partial \bar{g}}{\partial x}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{x}}(x(t), \dot{x}(t), p(t), t)\Big] \tag{31}$$

后 l 个方程中，由于 $\bar{g} = g(x, \dot{x}, t) + p \cdot F(x, \dot{x}, t)$ 中并没有 $\dot{p}$ 项，因而后项都为零，前项为

$$\frac{\partial \bar{g}}{\partial p_j}(x(t), \dot{x}(t), p(t), t) = F_j(x(t), \dot{x}(t), t), \quad j = 1, 2, \ldots, l.$$

于是，

$$0 = F(x(t), \dot{x}(t), t). \tag{32}$$

4/4 有微分约束情况下的欧拉方程

得到了有微分约束情况下的泛函极值必要条件，Euler-Lagrange 方程：

$$0 = \frac{\partial \bar{g}}{\partial x}(x(t), \dot{x}(t), p(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial \bar{g}}{\partial \dot{x}}(x(t), \dot{x}(t), p(t), t)\right] \tag{33}$$

$$0 = F(x(t), \dot{x}(t), t). \tag{34}$$

以及题设中给定的初始时刻状态与终端时刻状态

$$x(t_0) = x_0, \quad x(t_f) = x_f.$$

对比 Euler-Lagrange 方程，除问题引入的等式约束外，二者在形式上的区别仅为拉格朗日乘子

积分约束的泛函极值

问题 8 (积分约束的泛函极值)

初值终值 $x(t_0) = x_0, x(t_f) = x_f$，$t_f$ *fixed*，$x(t)$ *需要满足约束*

$$\int_{t_0}^{t_f} b(x(t), \dot{x}(t), t)dt = B \tag{35}$$

最小化性能指标

$$J(x) = \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t)dt \tag{36}$$

1/3 引入新的状态变量

引入新的状态变量

$$z(t) := \int_{t_0}^{t} b(x(\tau), \dot{x}(\tau), \tau) d\tau. \tag{37}$$

求导可得

$$\dot{z}(t) = b(x(t), \dot{x}(t), t) \tag{38}$$

则原问题转化为增加一维状态，且满足约束条件

$$z(t_0) = 0,\ z(t_f) = B. \tag{39}$$

2/3 化为微分约束泛函极值

引入拉格朗日乘子

$$\bar{g}(x(t),\dot{x}(t),\dot{z}(t),p(t),t) := g(x(t),\dot{x}(t),t) + p(t)[b(x(t),\dot{x}(t),t) - \dot{z}(t)] \quad (40)$$

得到欧拉-拉格朗日方程

$$0 = \frac{\partial \bar{g}}{\partial x}(x(t),\dot{x}(t),\dot{z}(t),p(t),t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{x}}(x(t),\dot{x}(t),\dot{z}(t),p(t),t)\Big], \quad (41)$$

$$0 = \frac{\partial \bar{g}}{\partial z}(x(t),\dot{x}(t),\dot{z}(t),p(t),t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{z}}(x(t),\dot{x}(t),\dot{z}(t),p(t),t)\Big], \quad (42)$$

$$0 = b(x(t),\dot{x}(t),t) - \dot{z}(t). \quad (43)$$

方程-(42) 前半为 0，后半得到 $p(t) = \lambda$ 为常数

积分约束的泛函极值 3/3

解：（得到一阶条件）

$$0 = \frac{\partial \bar{g}}{\partial x}(x(t), \dot{x}(t), \dot{z}(t), \lambda, t) - \frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial \bar{g}}{\partial \dot{x}}(x(t), \dot{x}(t), \dot{z}(t), \lambda, t)\Big] \tag{44}$$

$$0 = b(x(t), \dot{x}(t), t) - \dot{z}(t), \tag{45}$$

以及边界条件 $z(t_0) = 0, z(t_f) = B$ 其中

$$\bar{g}(x(t), \dot{x}(t), \dot{z}(t), \lambda, t) = g(x(t), \dot{x}(t), t) + \lambda[b(x(t), \dot{x}(t), t) - \dot{z}(t)]$$

例子：积分约束的泛函极值

例 8

$$\min J(x) = \int_0^1 \dot{x}^2(t)dt$$
$$\text{s.t.} \int_0^1 x(t)dt = B$$
$$x(0) = 0, x(1) = 2.$$

解：积分约束的泛函极值

解：

引入变量 $z(t) = \int_0^t x(\tau)d\tau$ 有 Lagrangian

$$\bar{g} = \dot{x}(t)^2 + \lambda(x(t) - \dot{z}(t))$$

$$0 = \bar{g}_x - \frac{d}{dt}\bar{g}_{\dot{x}} = \lambda - 2\ddot{x} \Rightarrow \dot{x} = \frac{1}{2}\lambda t + c_1 \Rightarrow x = \lambda t^2 + c_1 t + c_2 \tag{46}$$

$$\dot{z}(t) = x(t) \Rightarrow z = \frac{\lambda}{3}t^3 + \frac{c_1}{2}t^2 + c_2 t + c_3 \tag{47}$$

边界条件 $x(0) = 0, x(1) = 2, z(0) = 0, z(1) = B$, 得
$c_1 = 6B - 4, c_2 = c_3 = 0, \lambda = 6 - 6B$