第二章 变分法

（Variational Approach）

- 最优控制所要解决的问题：在一定的约束条件下，求使性能指标达到极大（或极小）值的控制函数。

  约束条件——一般是由向量微分方程描述的控制对象特性

  性能指标——一般是用泛函来描述

- 也就是说，最优控制问题实际上是在微分方程约束下求泛函的条件极值问题，而从数学上看这就是一个**变分问题**，需要用**变分法**求解。
- 变分法是近代数学中的一个完整分支，是研究最优控制问题的重要工具。

2.1 赋范线性空间

- 此节为有关的数学基础知识。 供参考，不展开。

2.2 线性算子及泛函

1. 线性算子
2. 线性算子的微商
3. 泛函
4. 泛函宗量及其变分
5. 线性泛函
6. 连续泛函
7. 泛函变分

1. 线性算子

- 定义2-10： n维线性空间R^n到m维线性空间R^m的线性算子（映射）是指一确定对应规律 $y = f(x)$，使每一个$x \in R^n$有一个对应的$y \in R^m$，且满足线性条件

  1）$f(x_1+x_2) = f(x_1) + f(x_2)$， $x_1, x_2 \in R^n$

  2）$f(\alpha x) = \alpha f(x)$， $\forall x \in R^n$，$\alpha \in R$

微积分中的线性函数关系 $y = f(x)$ 就是一种线性算子。

2. 线性算子的微商

- 定义2-11：设$y = f(x)$ 是n维线性空间R^n中的子集D到m维线性空间R^m的算子，且在D中当由点x转到$x + \Delta x$时，y变为$y + \Delta y = f(x + \Delta x)$。若存在一个由$D$到$R^m$的线性算子$k$，使

$$f(x + \Delta x) - f(x) = k\Delta x + \theta\|\Delta x\|$$

  - 其中 $\theta\|\Delta x\|$ 是 $\|\Delta x\|$ 的高阶无穷小量，即

$$\lim_{\|\Delta x\| \to 0} \frac{\theta\|\Delta x\|}{\|\Delta x\|} = 0$$

- 则称线性算子k为$f(x)$在x处的微商，记为 $k = f'(x)$，并称$f(x)$在x处可微。

可同样定义二阶至n阶微商，并解释$f(x)$在x处n次可微含义

二阶微商　$(f'(x))'=f''(x)$

……

……

n阶微商　$(f^{(n-1)}(x))'=f^{(n)}(x)$

回顾微积分中函数 $y=f(x)$ 的增量、高阶无穷小、微分等相应概念。

3 .泛函

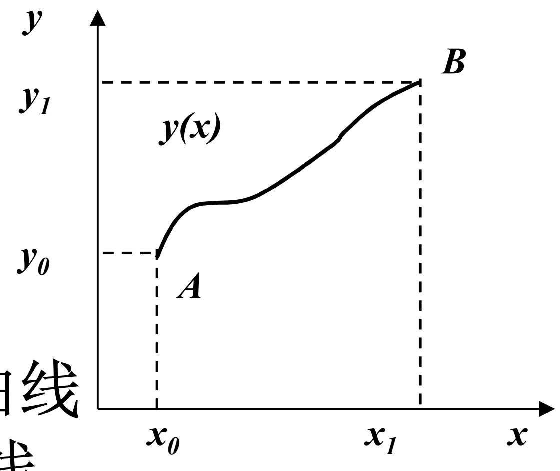
- 定义2-12：由赋范线性空间R^n到数域R的算子称为R^n上的泛函。
- 例：求曲线弧长的公式即为泛函

xy平面上两点$A(x_0,y_0)$, $B(x_1,y_1)$ 间的弧长公式为

$$\overset{\frown}{AB} = \int_{x_0}^{x_1} \sqrt{1+y'^2}\,dx$$

曲线（函数）$y = y(x)$通过A、B两点，曲线（函数）不同则弧长不同，即弧长是曲线（函数）$y = y(x)$的函数，记为$J[y(x)]$，则

$$\overset{\frown}{AB} = \int_{x_0}^{x_1} \sqrt{1+y'^2}\,dx = J[y(x)]$$



曲线弧长示意图

- 由二维空间（xy平面）映射到一维空间（长度）

4.泛函宗量及其变分

- 泛函$J[y(x)]$的宗量是函数 $y(x)$;

- 泛函宗量的变分: $\delta y(x) = y(x) - y_0(x)$

（回顾函数关系$y = f(x)$ 中的自变量x）

5. 线性泛函

- 定义2-13：线性空间R^n上的泛函J为R^n上的线性泛函，*iff*

$$J(\alpha x+\beta y)=\alpha J(x)+\beta J(y)\ ,\forall x,y\in R^n,\forall \alpha,\beta\in R$$

例如：积分

$$\int_a^b[\alpha\Phi(x)+\beta\Psi(x)]dx=\alpha\int_a^b\Phi(x)dx+\beta\int_a^b\Psi(x)dx$$

为线性泛函。

6.连续泛函

- 定义2-14：线性空间R^n上的泛函J为R^n上的连续泛函，*iff*
  对任意$x_0, x\in R^n, \varepsilon\in R$, $x= x_0+\varepsilon\Delta x$，当
  $\|x-x_0\|\to 0(\varepsilon\to 0)$
  时，有 $J(x)\to J(x_0)$
- 连续泛函的重要性在于，任意一点的泛函值可以用该点附近的泛函值任意逼近。
- 在有穷维线性空间上，任何线性泛函都是连续的。

7.泛函变分

- 泛函变分可以从两种不同描述角度加以定义。
- 定义2-15（1）：若赋范线性空间R^n上的泛函$J(x)$作为算子在x_0处是可微的，则其微分$J'(x)\Delta x$称为泛函$J(x)$在$x = x_0$处的变分，记为$\delta J(x_0,\Delta x)$。
- 该定义表明泛函$J(x)$的变分就是泛函增量 $\Delta J=J(x+\Delta x) - J(x)$的线性主部。

- 相应可定义：

$$\delta^2 J(x_0,\Delta x) = J''(x_0)(\Delta x)^2$$为泛函的二阶变分，

………

$$\delta^n J(x_0,\Delta x) = J^{(n)}(x_0)(\Delta x)^n$$为泛函的n阶变分。

泛函变分

- 定义2-15（2）：泛函$J(x_0+\varepsilon\Delta x)$($|\varepsilon|\in[0，1]$，$\varepsilon=0$时泛函为$J(x_0)$，$\varepsilon=1$时泛函为$J(x_0+\Delta x)$) 对$\varepsilon$的导函数在$\varepsilon=0$时的值，即

$$\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\Delta x)\bigg|_{\varepsilon=0}$$

称为$J(x)$ 在$x=x_0$处的变分。

上述两种定义本质相同，即有：

$$\delta J(x_0,\Delta x)=\frac{\partial}{\partial\varepsilon}J(x_0+\varepsilon\Delta x)\bigg|_{\varepsilon=0}\ ,\quad 0\leq|\varepsilon|\leq 1 \qquad \textbf{(2-2-1)}$$

- 定理**2-1**：设$J(x)$ 是赋范线性空间R^n上的泛函，若在$x = x_0$处可微，则其变分为（**2-2-1**）式。

证明：

$\because J(x)$在$x = x_0$处可微

$$\therefore \frac{\partial}{\partial \varepsilon} J(x_0 + \varepsilon \Delta x)\Big|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{J(x_0 + \varepsilon \Delta x) - J(x_0)}{\varepsilon}$$

$$= \lim_{\varepsilon \to 0} \frac{\delta J(x_0, \varepsilon \Delta x) + \theta(x_0, \varepsilon \Delta x)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{J'(x_0)\varepsilon \Delta x + \theta(x_0, \varepsilon \Delta x)}{\varepsilon}$$

$$= \lim_{\varepsilon \to 0} \frac{\varepsilon J'(x_0)\Delta x}{\varepsilon} + \lim_{\varepsilon \to 0} \frac{\theta(x_0, \varepsilon \Delta x)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{\varepsilon\, \delta J(x_0, \Delta x)}{\varepsilon} + \lim_{\varepsilon \to 0} \frac{\theta(x_0, \varepsilon \Delta x)}{\varepsilon}$$

$$= \delta J(x_0, \Delta x)$$

同样可证，若在$x = x_0$处$J(x)$ n次可微，则其n阶变分为

$$\delta^n J(x_0 + \Delta x) = \frac{\partial^n}{\partial \varepsilon^n} J(x + \varepsilon \Delta x)\Big|_{\varepsilon=0}, \quad 0 \le |\varepsilon| \le 1$$

2.3 变分原理

- 变分原理讨论泛函的极值及其条件问题，是变分法的理论基础。

- 三个概念：
    - 泛函的极值
    - 泛函极值的必要条件
    - 泛函极值的充分条件

泛函的极值

- 定义2-16：设$\boldsymbol{J}(\boldsymbol{x})$ 为$\boldsymbol{R}^n$上某子集$\boldsymbol{D}$中的泛函，对于$\boldsymbol{D}$中某一点 $\hat{\boldsymbol{x}} \in \boldsymbol{D}$ ，称泛函$\boldsymbol{J}(\boldsymbol{x})$在 $\boldsymbol{x} = \hat{\boldsymbol{x}}$ 处达到极小（或极大）值，是指$\boldsymbol{J}(\boldsymbol{x}) \geq \boldsymbol{J}(\hat{\boldsymbol{x}})$ (或 $\boldsymbol{J}(\boldsymbol{x}) \leq \boldsymbol{J}(\hat{\boldsymbol{x}})$)。其中，

$$\boldsymbol{x} \in \boldsymbol{U}(\hat{\boldsymbol{x}}, \sigma) \subset \boldsymbol{D}, \sigma > 0$$

$$\boldsymbol{U}(\hat{\boldsymbol{x}}, \sigma) = \left\{ \boldsymbol{x} \,\middle|\, \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| < \sigma ,\ \boldsymbol{x} \in \boldsymbol{R}^n \right\}$$

$\hat{\boldsymbol{x}}$被称为泛函$\boldsymbol{J}(\boldsymbol{x})$的极小（或极大）点。

泛函极值的必要条件

- 定理2-2：设$J(x)$ 是在R^n的某个开子集 $G \subset R^n$上定义的泛函，且在$x=\hat{x}$ 处有一阶变分，如果泛函$J(x)$在 $\hat{x}$ 处达到极值，则其一阶变分为0，即

$$\delta J(\hat{x},\Delta x)=0$$

- 证明：设$J(x)$在 $x=\hat{x}$ 处达到极小值，则存在一正数$\sigma>0$，使当 $\hat{x}+\Delta x\in U(\hat{x},\sigma)$ 时，有 $J(\hat{x}+\Delta x)\geq J(\hat{x})$。对于确定的 $\hat{x}$ 和 $\hat{x}+\Delta x$，作为ε的函数，$J(\hat{x}+\varepsilon\Delta x)=\varphi(\varepsilon),(0\leq\varepsilon\leq 1)$在$\varepsilon=0$处达到极小值，则有$\varphi'(\varepsilon)\big|_{\varepsilon=0}=0$，即

$$\delta J(\hat{x},\Delta x)=\frac{\partial}{\partial\varepsilon}J(\hat{x}+\varepsilon\Delta x)\Big|_{\varepsilon=0}=0$$

。证毕。

- 该必要条件非常重要，是后续讨论的基础。

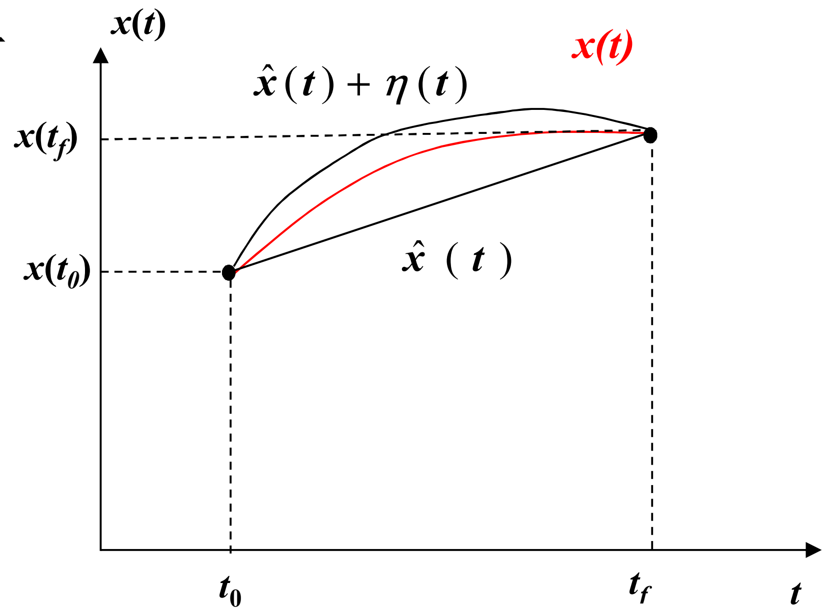
泛函极值的充分条件

- 定理2-3：设$J(x)$ 是在R^n的某个开子集 $G \subset R^n$上定义的泛函，且在 $x = \hat{x}$处有二阶变分，如果泛函$J(x)$在$\hat{x}$ 处一阶变分为0，且二阶变分大于0，$\delta^2 J(\hat{x}, \Delta x) > 0$，则$J(x)$在 $\hat{x}$处达到极小值；反之，如果 $\delta^2 J(\hat{x}, \Delta x) < 0$，泛函$J(x)$在 $\hat{x}$处达到极大值。

2.4 无约束条件的泛函极值问题
—Euler方程

- 考虑轨线 $x(t)$，设其始端 $x(t_0)=x_0$ 和终端 $x(t_f)=x_f$ 均属已知，试寻求连续可微的极值轨线 $\hat{x}(t)$，使性能泛函
- $$J(x)=\int_{t_0}^{t_f} L[x(t),\dot{x}(t),t]\,dt \qquad \text{(2-4-1)}$$
  达到极值，其中被积函数 $L[x(t),\dot{x}(t),t]$是连续可微函数。
- 轨线$x(t)$除始端和终端固定及连续可微要求外，不受其他条件约束。

- 假定极值轨线为$\hat{x}(t)$，其附近一容许轨线为 $\hat{x}(t)+\eta(t)$，其中$\eta(t)$连续可微。



极值轨线与容许轨线

- $\hat{x}(t)$和$\hat{x}(t)+\eta(t)$两轨线间所有容许轨线可表示为

$$x(t)=\hat{x}(t)+\varepsilon\eta(t),\ 0\le|\varepsilon|\le 1 \qquad \text{（2-4-2）}$$

- 当$\varepsilon=0$时，即为极值曲线 $\hat{x}(t)$，有 $\hat{x}(t)=x(t)\big|_{\varepsilon=0}$
- 将（2-4-2）代入（2-4-1）得

$$J(x)=\int_{t_0}^{t_f} L[\hat{x}(t)+\varepsilon\eta(t),\dot{\hat{x}}(t)+\varepsilon\dot{\eta}(t),t]\,dt \qquad \text{（2-4-3）}$$

泛函极值必要条件——一阶变分=0

- （2-4-3）式表明$J(x)$是ε的函数，极值轨线上满足

$$\delta J(x)=\frac{\partial J(x)}{\partial \varepsilon}\Big|_{\varepsilon=0}=0 \qquad \text{（2-4-4）}$$

- 由（2-4-1）、（2-4-2）和（2-4-3）式定义显然有

$$\lim_{\varepsilon\to 0} x(t)=\hat{x}(t) \quad 和 \quad \lim_{\varepsilon\to 0} J(x)=J(\hat{x})$$

- 由（2-4-3）和（2-4-4）可得

$$\frac{\partial J(x)}{\partial \varepsilon}\Big|_{\varepsilon=0}=\int_{t_0}^{t_f}\left\{\eta(t)\frac{\partial L[x(t),\dot{x}(t),t]}{\partial x(t)}+\dot{\eta}(t)\frac{\partial L[x(t),\dot{x}(t),t]}{\partial \dot{x}(t)}\right\}dt=0$$

即

$$\boxed{\int_{t_0}^{t_f}\eta(t)\frac{\partial L[x(t),\dot{x}(t),t]}{\partial x(t)}dt+\int_{t_0}^{t_f}\dot{\eta}(t)\frac{\partial L[x(t),\dot{x}(t),t]}{\partial \dot{x}(t)}dt=0} \qquad \text{（2-4-5）}$$

泛函极值必要条件的具体描述，推导*Euler*方程的起点。

需变换为$\eta(t)$同类项

对（**2-4-5**）式左边第二项分部积分，有

$$\int_{t_0}^{t_f} \dot{\eta}(t) \frac{\partial L}{\partial \dot{x}} dt = \frac{\partial L}{\partial \dot{x}} \eta(t) \Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \eta(t) \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} dt \qquad \text{(2-4-6)}$$

将（**2-4-6**）式代入（**2-4-5**）式，有

$$\int_{t_0}^{t_f} \eta(t) \left\{ \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right\} dt + \frac{\partial L}{\partial \dot{x}} \eta(t) \Big|_{t_0}^{t_f} = 0 \qquad \text{(2-4-7)}$$

因两端点固定，$\eta(t_0) = \eta(t_f) = 0$，故（**2-4-7**）式化为

$$\int_{t_0}^{t_f} \eta(t) \left\{ \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right\} dt = 0 \qquad \text{(2-4-8)}$$

至此，需要引入变分预备定理。

- 变分预备定理：

  设$\boldsymbol{M(t)}$是区间$[t_0, t_1]$上的$\boldsymbol{n}$维连续向量函数，如果对于任意连续向量函数$\boldsymbol{\eta(t)}$，$\eta(t_0)=\eta(t_1)=0$，皆有

  $\int_{t_0}^{t_1} M(t)\eta(t)dt = 0$ ，则在区间$[t_0, t_1]$上$M(t)\equiv 0$。

- 根据变分预备定理，由于在极值轨线$\hat{x}(t)$处（2-4-8）式对任意$\eta(t)$都应成立，所以有

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \equiv 0 \qquad (2\text{-}4\text{-}9)$$

至此，可以得到以下定理。

- 定理2-4：设已知轨线 $x(t)$及其始端 $x(t_0) = x_0$和终端 $x(t_f) = x_f$，则其使性能泛函

$$J(x) = \int_{t_0}^{t_f} L[x(t), \dot{x}(t), t]\, dt \tag{2-4-10}$$

取极值的必要条件是轨线 $x(t)$为微分方程

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0 \tag{2-4-11}$$

的解。

微分方程（2-4-11）的展开形式可写成

$$L_x - L_{\dot{x}\,t} - L_{\dot{x}\,x}\dot{x} - L_{\dot{x}\,\dot{x}}\ddot{x} = 0 \tag{2-4-12}$$

- 方程（2-4-11）即为欧拉方程(**Euler Equation**)，也称为欧拉－拉格朗日方程(**Euler-Lagrange Equation**)。
- 一般情况下，欧拉方程是二阶非线性微分方程，属于两点边值问题。

- 例**2-1**：设轨线$x(t)$的始端 $x(0)=0$ ，终端 $x(\frac{\pi}{2})=1$ ，求使性能泛函 $J(x)=\int_0^{\frac{\pi}{2}}[\dot{x}^2(t)-x^2(t)]\,dt$ 达到极值的极值轨线。
- 解：这里 $L[x(t),\dot{x}(t),t]=\dot{x}^2(t)-x^2(t)$

  由欧拉方程 $\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}=0$ ， 有 $-2x(t)-\frac{d}{dt}[2\dot{x}(t)]=0$

  整理得 $\ddot{x}(t)+x(t)=0$

  解此微分方程得 $x(t)=c_1\cos t+c_2\sin t$

  考虑边界条件$x(0)=0$ 和 $x(\frac{\pi}{2})=1$，得 $c_1=0$ ，$c_2=1$

  $\therefore x(t)=\sin t$ 即为所求的极值曲线。 #
- 以上欧拉方程的推导是将$J(x)$看作是ε的函数，按一般微积分运算中求极值方法处理。

- 另一种方法可以直接应用变分的定义表达式求性能泛函极值。
- 对性能泛函

$$J(x)=\int_{t_0}^{t_f} L[\hat{x}(t)+\varepsilon\eta(t),\dot{\hat{x}}(t)+\varepsilon\dot{\eta}(t),t]\,dt \qquad (2\text{-}4\text{-}13)$$

将L在ε=0的邻域展开为Taylor级数

$$L[\hat{x}(t)+\varepsilon\eta(t),\dot{\hat{x}}(t)+\varepsilon\dot{\eta}(t),t]$$

$$=L[\hat{x}(t),\dot{\hat{x}}(t),t]+\frac{\partial L}{\partial x}\varepsilon\eta(t)+\frac{\partial L}{\partial \dot{x}}\varepsilon\dot{\eta}(t)+H.O.T. \qquad (2\text{-}4\text{-}14)$$

其中，*H.O.T.*为关于$\eta(t)$和$\dot{\eta}(t)$的高阶无穷小项。

- 泛函的增量为 $\Delta J=J(\hat{x}+\varepsilon\eta)-J(\hat{x})$

即 $$\Delta J=\int_{t_0}^{t_f}\{L[\hat{x}(t)+\varepsilon\eta(t),\dot{\hat{x}}(t)+\varepsilon\dot{\eta}(t),t]-L[\hat{x}(t),\dot{\hat{x}}(t),t]\}dt$$

$$=\int_{t_0}^{t_f}\{\frac{\partial L}{\partial x}\varepsilon\eta+\frac{\partial L}{\partial \dot{x}}\varepsilon\dot{\eta}+H.O.T.\}dt \qquad (2\text{-}4\text{-}15)$$

- 可定义 $x(t)$ 和 $\dot{x}(t)$ 的一阶变分为 $\delta x = \varepsilon\eta(t)$ 和 $\delta\dot{x} = \varepsilon\dot{\eta}(t)$，则有

$$\Delta J = \int_{t_0}^{t_f} \{\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial \dot{x}}\delta\dot{x} + H.O.T.\}dt \qquad (2\text{-}4\text{-}16)$$

- 取泛函增量 ΔJ的线性主部即为其一阶变分

$$\delta J = \int_{t_0}^{t_f} \{\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial \dot{x}}\delta\dot{x}\}dt \qquad (2\text{-}4\text{-}17)$$

- 分部积分后有

$$\delta J = \int_{t_0}^{t_f} \left\{\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right\}\delta x dt + \frac{\partial L}{\partial \dot{x}}\delta x\Big|_{t_0}^{t_f} \qquad (2\text{-}4\text{-}18)$$

- 当端点固定时有 $\delta x(t_0) = \delta x(t_f) = 0$，所以有 $\frac{\partial L}{\partial \dot{x}}\delta x\Big|_{t_0}^{t_f} = 0$ ，即有

$$\delta J = \int_{t_0}^{t_f} \left\{\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right\}\delta x dt \qquad (2\text{-}4\text{-}19)$$

- 由泛函极值必要条件、为任意取值以及变分预备定理，即可求得欧拉方程为

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0 \qquad (2\text{-}4\text{-}20)$$

与（2-4-11）式结果相同。

- 欧拉方程可以推广到***n***维向量微分方程，即***n***维状态空间。
- 定理**2-5**：

在***n***维状态空间中，已知状态向量 $\mathbf{x}(t)=[x_1(t),x_2(t),\cdots\cdots,x_n(t)]^{\mathrm{T}}$ 的起点 $\mathbf{x_0}(t)=[x_{10}(t),x_{20}(t),\cdots\cdots,x_{n0}(t)]^{\mathrm{T}}$和终点 $\mathbf{x_f}(t)=[x_{1f}(t),x_{2f}(t),\cdots\cdots,x_{nf}(t)]^{\mathrm{T}}$ ，则性能泛函

$$J[\mathbf{x}(t)]=\int_{t_0}^{t_f} L[\mathbf{x}(t),\dot{\mathbf{x}}(t),t]\,dt \qquad \text{（2-4-21）}$$

取极值的必要条件，是轨线 $\mathbf{x}(t)$ 为向量微分方程

$$\frac{\partial L}{\partial \mathbf{x}}-\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}}=0 \qquad \text{（2-4-22）}$$

的解，其中***x***应有连续的二阶导数，而***L***则至少两次连续可微。

★理解（**2-4-22**）式的展开形式！