

# 10 数学物理方程的 变分原理

- (0) 引言与简介
- (一) 变分问题介绍
- (二) 一维数学物理问题的变分问题
- (三) 高维数学物理问题的变分问题
- (四) 变分问题的近似计算
- (五) 权余量方法以及其他方法

# 引言

有限元算法（Finite Element Method，简称 FEM）是一种非常流行的求解偏微分方程的数值算法。有限元被广泛应用于结构受力分析、复杂边界的麦克斯韦方程求解以及热传导等问题。

加权残差法（Weighted Residual Method）

**举例：求解  $x(t)$ ，使其满足如下微分方程，**

$$\frac{dx}{dt} + x = 0$$

**初始条件  $x(0) = 1$ ，求解区域  $t \in [0, 1]$ 。**

这个方程的解析解是  $e^{-t}$ ，

$$x(t) = e^{-t} = 1 - t + \frac{1}{2!}t^2 - \frac{1}{3!}t^3 + \frac{1}{4!}t^4 + \dots$$

在计算机上展开到无穷阶，就能给出方程的精确解。现实中显然只能考虑有限项，后面忽略的项称作截断误差。假设截断到二阶，

$$x(t) \approx 1 - t + \frac{1}{2!}t^2$$

代入原方程产生残差  $R$ ，

$$\frac{dx}{dt} + x = R, R \neq 0$$

$R$  越接近0，表示近似解越接近精确解。

明知道精确解  $e^{-t}$  不在  $(1, t, t^2)$  这三个函数支撑起的函数空间中，我们仍可做出改进。

使用待定系数展开，

$$x(t) \approx 1 + c_1 t + c_2 t^2$$

此时残差函数

$$R(t) = \frac{dx}{dt} + x(t) = 1 + (1+t)c_1 + (2t+t^2)c_2$$

通过调节  $c_1, c_2$  , 最小化  $R(t)$  的绝对值, 就能得到比  $1 - t + \frac{1}{2!}t^2$  更好的解。

最优情况显然是  $R(t)$  对任意  $t$  都等于零, 但这个做不到。所以存在不同的方法, 减弱约束。下面介绍四种方法,

- 点匹配法: 在有限个点上, 要求  $R(t)$  等于0
- 子区域法: 在有限子域上, 要求  $R(t)$  均值等于0
- 伽辽金法: 对有限检验函数, 要求  $R(t)$  的加权求和等于0
- 最小二乘法: 通过求  $\int_0^1 R^2(t)dt$  的最小值, 解待定系数

## (1) 点匹配法 (Point Matching)

最简单的考虑是，既然有  $c_1, c_2$  两个待定系数，那我们只需要任选两个点，让  $R(t)$  在那两个点上等于 0，就可以得到两个方程，刚好可求解出  $c_1, c_2$ ，

比如，选  $t = 1/3, 2/3$  两个点，得到两个方程。对第一个方程，

$$R\left(\frac{1}{3}\right) = 1 + c_1 \left(1 + \frac{1}{3}\right) + c_2 \left(2 \times \frac{1}{3} + \left(\frac{1}{3}\right)^2\right) = 0$$

化简得到，

$$\frac{4}{3}c_1 + \frac{7}{9}c_2 = -1$$

同理，对  $t = 2/3$ ，得到方程，

$$\frac{5}{3}c_1 + \frac{16}{9}c_2 = -1$$

解出  $c_1 = -0.9310, c_2 = 0.3103$ ，所以近似解为  $x(t) \approx 1 - 0.9310 t + 0.3103 t^2$

## (2) 子域法 (Subdomain method)

根据待定系数的个数  $n$ , 将求解区域分成  $n$  份, 要求  $R(t)$  在每份上的积分等于0。对于这个例子, 分  $t \in [0, 1/2]$  和  $t \in [1/2, 1]$  两个子域, 求解如下两个方程,

$$\int_0^{1/2} R(t) dt = 0, \quad \int_{1/2}^1 R(t) dt = 0$$

即可得到近似解,

$$x(t) \approx 1 - 0.9474 t + 0.3158 t^2$$

### (3) 伽辽金法 (Galerkin method)

如果  $R(t) = 0$  对于任意  $t$  都成立, 显然  $\phi(t)R(t) = 0$  应该对任意检验函数  $\phi(t)$  成立。

根据待定系数个数  $n$ , 选取  $n$  个不同的检验函数  $\phi_i(t)$ ,  $1 \leq i \leq n$ , 要求对每个  $\phi_i$ , 下式满足

$$\int_0^1 \phi_i(t) R(t) dt = 0$$

这样, 就构造了  $n$  个方程, 刚好可解  $n$  个待定系数。

检验函数的一种自然选择是  $x(t)$  展开时使用的基底函数  $1, t^1, t^2$ , 求解

$$\int_0^1 t R(t) dt = 0, \quad \int_0^1 t^2 R(t) dt = 0$$

得到近似解,

$$x(t) \approx 1 - 0.9143 t + 0.2857 t^2$$

#### (4) 最小二乘法 (Least Square Method)

函数极值处一阶导数等于0。求残差平方  $R^2(t)$  积分的最小值, 可以得到待定系数满足的方程,

$$\frac{\partial}{\partial c_1} \int_0^1 R^2(t) dt = 2 \int_0^1 \frac{\partial R}{\partial c_1} R(t) dt = 0$$

对于  $c_1, c_2$  分别求导数, 得到两个方程,

$$\begin{aligned} \int_0^1 \frac{\partial R}{\partial c_1} R dt &= \frac{3}{2} + \frac{7}{3}c_1 + \frac{9}{4}c_2 = 0 \\ \int_0^1 \frac{\partial R}{\partial c_2} R dt &= \frac{4}{3} + \frac{9}{4}c_1 + \frac{38}{15}c_2 = 0 \end{aligned}$$

求解可得近似解,

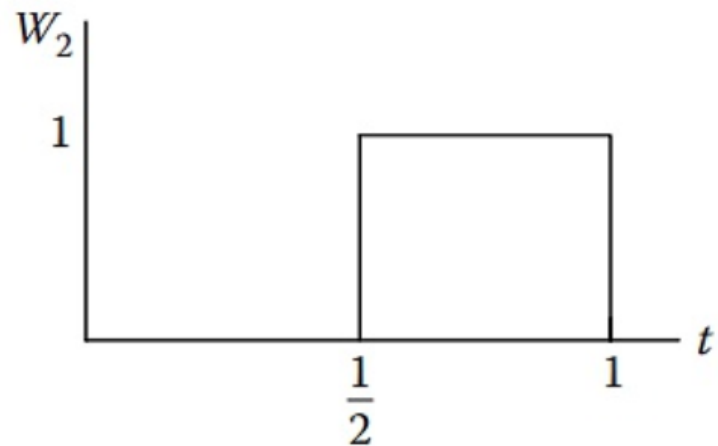
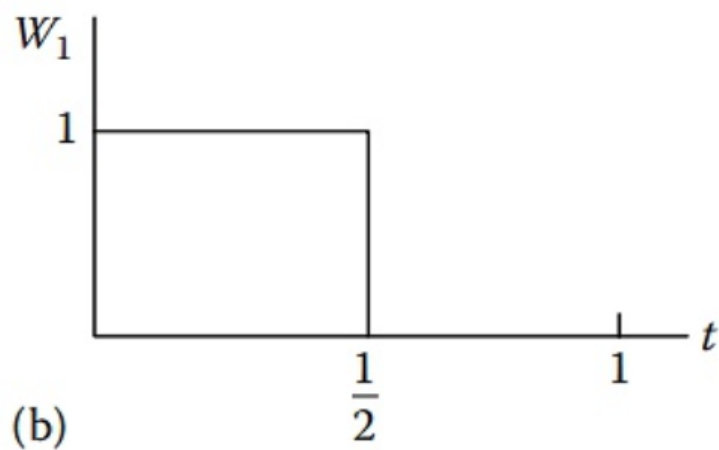
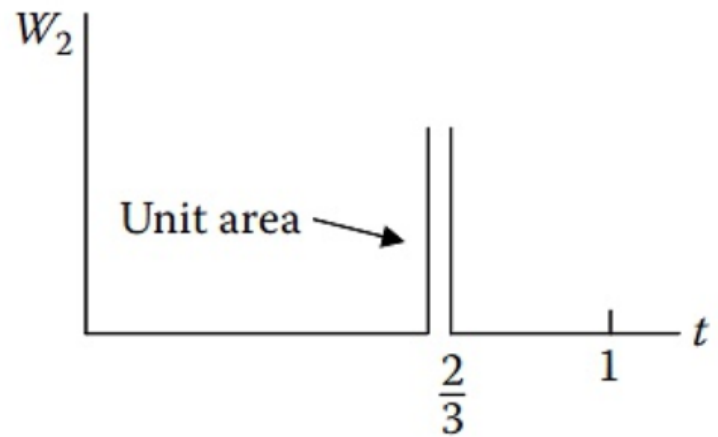
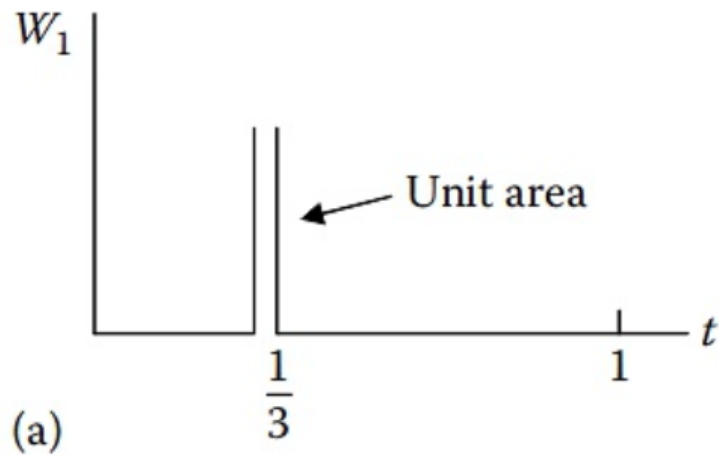
$$x(t) \approx 1 - 0.9427 t + 0.3110 t^2$$

这四种方法可以用一个统一的式子描述, 这个式子跟 Galerkin 给出的式子基本一致,

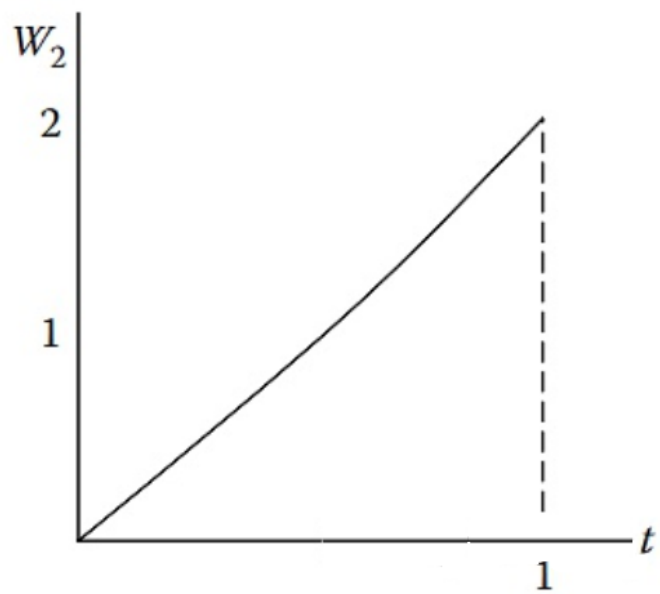
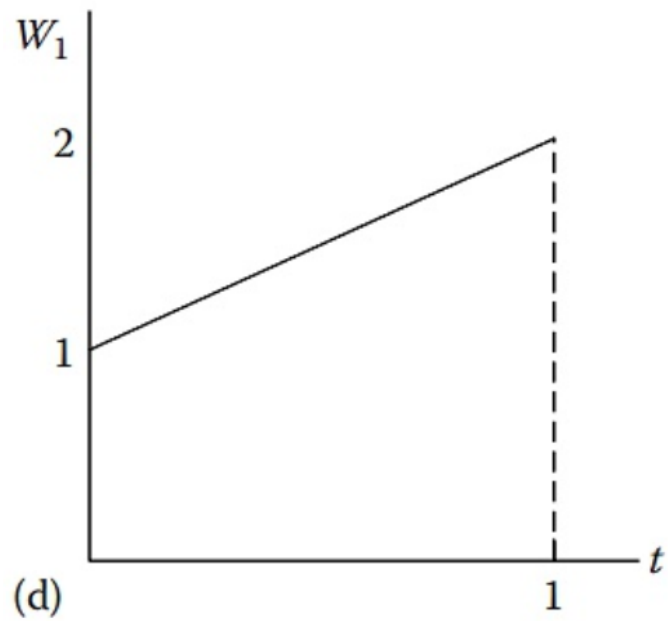
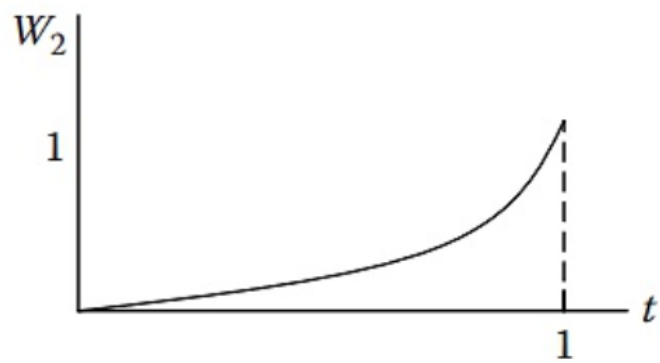
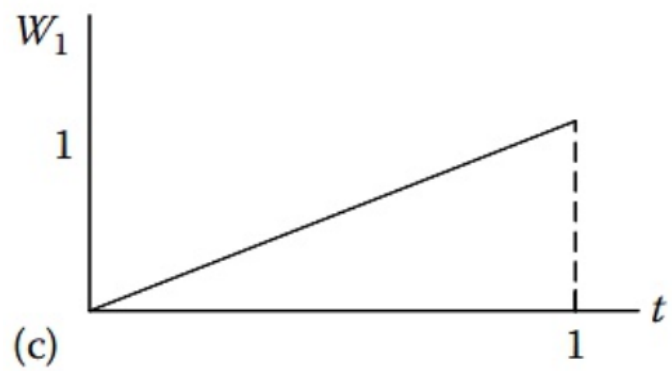
$$\int_0^1 w_j(t) R(t) dt = 0, \quad j = 1, 2$$



参考文献【1】总结了这四种方法对应的检验函数，如下所示，



文献1: Advanced Electromagnetic Computation, Dikshitulu K. Kalluri



## 有限元与有限差分算法的区别

一般的微分方程都可以写成  $Lf(x) = 0$  形式，其中  $L$  称为微分算子 (Differential Operator)。有限差分和有限元的区别在于有限差分对微分算子进行离散化近似，有限元对待求解函数进行基函数展开近似。

$$\underbrace{\text{Differential Operator}}_{\text{finite difference}} \underbrace{f(x)}_{\text{finite element}} = 0$$

有限差分算法(finite difference) 用有限差分来近似微分算子，比如一阶微分，

$$\frac{df}{dx} \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}$$

有限元算法 (finite element) 用子域上的基函数展开来近似想要求解的方程  $f(x)$  ,

$$f(x) = \sum_{k=1}^n a_k \phi_k(x)$$

有了前面加权残差法 (Weighted Residual Method) 的介绍，相信大家对上面这种基函数展开已经有了比较深刻的理解。有限元算法正是广泛使用了上面的 Subdomain 方法与 Galerkin 方法的思想。

上面例子我们知道解析解，所以选择了泰勒展开时前几个最重要的多项式函数作为基底。如果不知道解的性质，如何选择基底函数？**有限元算法应用这么广泛的原因就在于它选择基底函数的方式简单粗暴有用，留在下节详细介绍。**

## 总结

这节介绍了加权残差法，有限元与有限差分的区别。这种使用有限个基底函数展开，再调整待定系数，补偿截断误差的思想，还能在量子力学变分法，量子多体的Ab initio方法，重整化群，密度矩阵重整化群中找到身影，应该说是一种非常有用的计算物理思想。

参考文献：

【1】Advanced Electromagnetic Computation, Dikshitulu K. Kalluri

一阶向前差分  $f(x_i + h) - f(x_i)$

一阶向后差分  $f(x_i) - f(x_i - h)$

一阶中心差分  $f(x_i + h) - f(x_i - h)$

由此可得：

一阶向前差商  $f'(x_i) \approx \frac{f(x_i + h) - f(x_i)}{h}$

一阶向后差商  $f'(x_i) \approx \frac{f(x_i) - f(x_i - h)}{h}$

一阶中心差商  $f'(x_i) \approx \frac{f(x_i + h) - f(x_i - h)}{2h}$

二阶中心差商  $f''(x_i) \approx \frac{f(x_i + h) - 2f(x_i) + f(x_i - h)}{h^2}$

# 偏微分方程差分例子

例如对于方程  $\Delta\phi(x, y) = g$  , 待求解区域为正方形域  $x \in [0, 1], y \in [0, 1]$

将待求解区域划分为矩形网格, 每一个小矩形的长和宽 (步长) 分别为  $h_x$  和  $h_y$  , 使用中心差商, 可得:

$$\left(\frac{\partial^2 \phi}{\partial x^2}\right)_0 \approx \frac{\phi_1 - 2\phi_0 + \phi_3}{h_x^2}$$

$$\left(\frac{\partial^2 \phi}{\partial y^2}\right)_0 \approx \frac{\phi_2 - 2\phi_0 + \phi_4}{h_y^2}$$

对于这里的正方形区域, 自然而然可以选取  $h_x = h_y = h$  , 代入方程, 得:

$$\frac{\phi_1 - 2\phi_0 + \phi_3}{h^2} + \frac{\phi_2 - 2\phi_0 + \phi_4}{h^2} = g_0$$

以此类推, 可以得到对于点  $\phi_{ij}$  , 有:

$$\frac{1}{h^2}(\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}) + \frac{1}{h^2}(\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}) = g_{i,j}$$

这便是点  $\phi_{i,j}$  满足的差分方程，由于该方程包含点  $\phi_{i,j}$  及其上下左右一共5个点，也称为“五点差分格式”。

五点差分格式的截断误差为  $O(h^2)$  阶。如果把左上、左下、右上、右下四个点也用上，为“九点差分格式”，误差可以降低到  $O(h^4)$  阶。

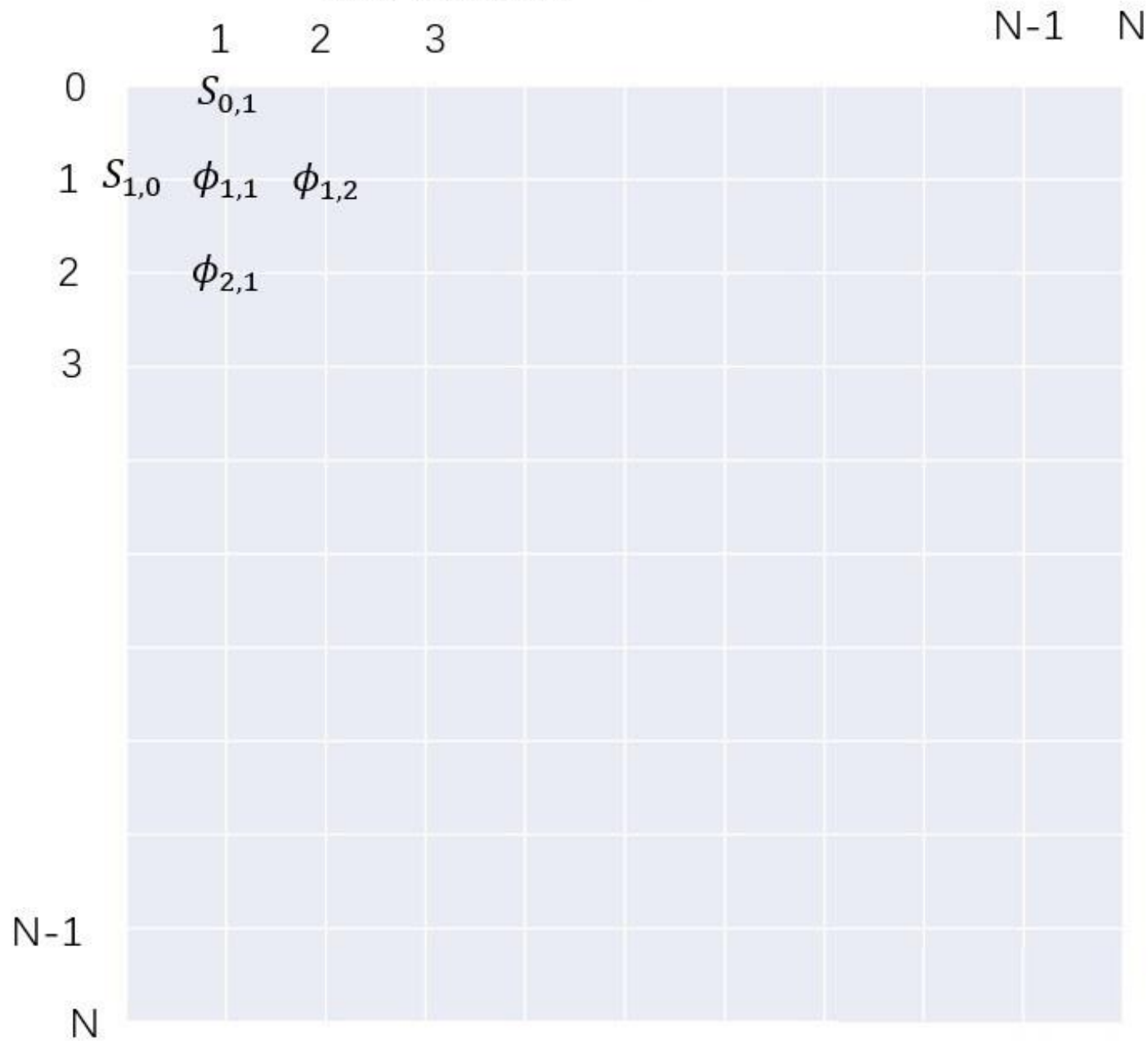
为简便起见，这里以拉普拉斯方程  $\Delta\phi(x,y) = 0$  为例，则五点差分格式简化为：

$$\phi_{i,j} - \frac{1}{4}(\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1}) = 0$$

在边界  $s$  上,  $\varphi(x,0) = \varphi(x,1) = 0, \varphi(0,y) = \varphi(1,y) = 1$

在划分网格时候要注意，如果划分网格为  $N \times N$  个正方形格子，那么待求解的点有  $(N-1) \times (N-1)$  个，因为区域边界上的点（即边界条件）是给定的，不用求。因此，给网格编号时候，从0开始编号再合适不过了，如下：

# 差分格式示意





考虑第一个点  $(1, 1)$ ，根据五点差分格式，它的值由上下左右四个点确定，很显然，上、左分别为边界点  $s_{0,1}$  和  $s_{1,0}$ ，右、下分为为点  $\phi_{1,2}$  和  $\phi_{2,1}$ ，写为差分格式为：

$$\phi_{1,1} - \frac{1}{4}(\phi_{2,1} + s_{0,1} + \phi_{1,2} + s_{1,0}) = 0$$

第一列其他点同理，以此类推，如果将网格第  $j$  列上每个点的解记为：

$$\varphi_j = \begin{pmatrix} \phi_{1,j} \\ \vdots \\ \phi_{2,j} \\ \vdots \\ \phi_{N-1,j} \end{pmatrix}, \text{ 则最终可以写为:}$$

$$G\varphi_1 - \frac{I}{4}\varphi_2 = b_1$$

其中  $I$  为  $(N - 1) \times (N - 1)$  阶单位矩阵，

$$G = \begin{pmatrix} 1 & -1/4 & \cdots & 0 \\ -1/4 & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & -1/4 \\ 0 & \cdots & -1/4 & 1 \end{pmatrix}$$

$$b_1 = -\frac{1}{4} \begin{pmatrix} -s_{0,1} - s_{1,0} \\ \vdots \\ -s_{N-1,0} - s_{N,1} \end{pmatrix}$$

这里出现了一个迭代方程，想解  $\varphi_1$  需要  $\varphi_2$ ，解  $\varphi_2$  需要  $\varphi_3$  .....

然后把这一系列矩阵方程写成一个更大的矩阵方程：

把整个网格的每一列纵向连接为一个列向量：

$$\Phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{N-1} \end{pmatrix}$$

则有：

$K\Phi = B$ ，其中：

$$K = \begin{pmatrix} G & -I/4 & \cdots & 0 \\ -I/4 & G & -I/4 & \vdots \\ \vdots & \vdots & \vdots & -I/4 \\ 0 & \cdots & -I/4 & G \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N-1} \end{pmatrix}$$

终于，我们把要求解的问题化简为了矩阵方程  $A\phi = B$  的形式，可以方便地使用计算机求解，这里A是系数矩阵，列向量  $\Phi$  便是问题的解。

```
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
```

```
N=79 # 矩阵维数
```

```
Imat = -np.eye(N)/4
```

```
def Gmatrix(N): # 生成G矩阵
```

```
    mtr = np.eye(N)
```

```
    for i in range(0, N - 1):
```

```
        mtr[i][i + 1] = -1 / 4
```

```
    for i in range(1, N):
```

```
        mtr[i][i - 1] = -1 / 4
```

```
    return mtr
```

```
def zero_mat(N,m): # 生成m个横向连接N阶零矩阵, m>=1
```

```
    zero_matrix = np.zeros((N,N))
```

```
    for i in range(m-1):
```

```
        zero_matrix = np.block([zero_matrix,np.zeros((N,N))])
```

```
    return zero_matrix
```

```

def zero_arr(N, m): # 生成m个纵向连接N阶列向量, m>=1
    zero_array = np.zeros((N, 1))
    for i in range(m-1):
        zero_array = np.block([[zero_array], [np.zeros((N, 1))]])
    return zero_array

def Kmatrix(N, Gmat, Imat):
    Kmar_list = []
    for i in range(N):
        if i == 0:
            Kmat_line = np.block([np.block([Gmat,
Imat]), zero_mat(N, N-2)])
            Kmar_list.append(Kmat_line)
        elif i == 1:
            Kmat_line = np.block([Imat, Kmar_list[-1][:, :-N]])
            Kmar_list.append(Kmat_line)
        else:
            Kmat_line = np.block([np.zeros((N, N)), Kmar_list[-
1][:, :-N]])
            Kmar_list.append(Kmat_line)
    Kmat = Kmar_list[0]
    for j in range(N-1):
        Kmat = np.block([[Kmat], [Kmar_list[j+1]]])
    return Kmat

```

```
def Barray(N): # 方程右边
    barr = np.ones((N, 1))/4
    barr = np.block([[barr], [zero_arr(N, N-2)]]))
    barr = np.block([[barr], [np.ones((N, 1))/4]])
    return barr
```

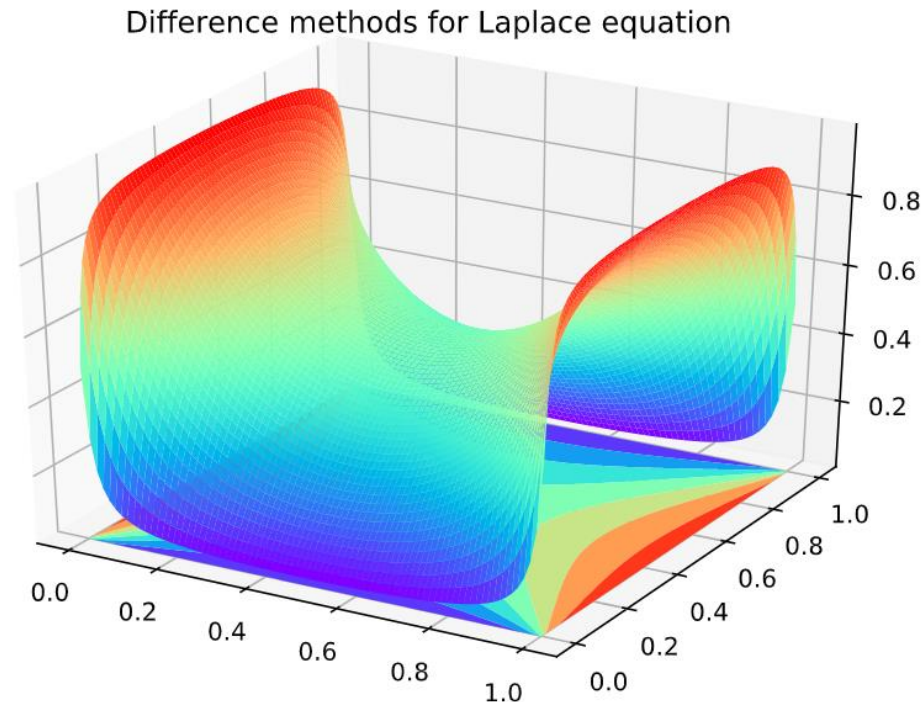
```
Gmat = Gmatrix(N)
Kmat = Kmatrix(N, Gmat, Imat)
Barr = Barray(N)
Phi = np.linalg.solve(Kmat, Barr)
```

```
def trans(N, Phi):
    # 把线性方程组解得的列矢量分割，从左到右排列转化为方矩阵，方便可视化
    phi = Phi[0:N]
    for i in range(N-1):
        j = i+1
        phi = np.block([phi, Phi[(j*N):((j+1)*N)]])
    return phi
```

```

Phi = trans(N,Phi)
x = np.linspace(0,1,N)
y = np.linspace(0,1,N)
X,Y = np.meshgrid(x,y)
fig = plt.figure()
ax = Axes3D(fig)
ax.plot_surface(X,Y,Phi,rstride=1,cstride=1,cmap=plt.get_cmap('rainbow'))
ax.contourf(X,Y,Phi,zdir='z',offset=0,cmap='rainbow')
plt.title('Difference methods for Laplace equation',fontsize='12')
plt.show()

```



# (一) 变分问题介绍



让一个物体从静止开始沿着一个光滑无摩擦的轨道下滑，如果要求下滑过程耗时最短，轨道应该是什么形状？



## 最速降曲线问题 (the brachistochrone problem)

- Johann 利用光学原理类比给出了一种解法，他更牛掰的哥哥 Jacob Bernoulli 想到了另一种解法（Bernoulli 家族盛产数学家，走量的那种）。此外，大名鼎鼎的 Gottfried Leibniz（跟牛顿争微积分发明权的那位），Guillaume de l'Hôpital（洛必达法则就是他的名号）等人都给出了各自的解法
- 1697年1月的某天，牛爵爷此时已在皇家造币厂任职，他开开心心下班回家，发现了来自 Johann Bernoulli 的挑战书。牛爵爷可能有点不太开心，内心也嘀咕着：老子牛逼顿，你们这些外国佬少来跟我在数学问题上嚣张！或者按他自己的原话：*I do not love to be pestered and teased by foreigners about mathematical things*. 于是牛爵爷也忍不住试了试水。以牛爵爷的智商，他大概也就经过了那么一晚上的思考，就运用变分法解决了问题，转手就写了封信寄回给了 Johann
- 不过 Newton 很低调地将他的解答匿名寄了回去。然而 Newton 的解法是如此之风流潇洒，令人拍案叫绝，即便没有署名，Johann 也很快意识到了真正的作者是何方神圣。对此，他做出了著名的评价：*I recognize the lion by his claw mark*
- 其实早在10多年前，牛爵爷在考虑在流体中会受到最小阻力的旋转曲面该是什么形状的问题时，已经构建起了变分原理的基本思想。经过一大批数学家的杰出工作，现在变分法已经成为了数学分析中的求极值问题的一种重要方法。

## 变分法

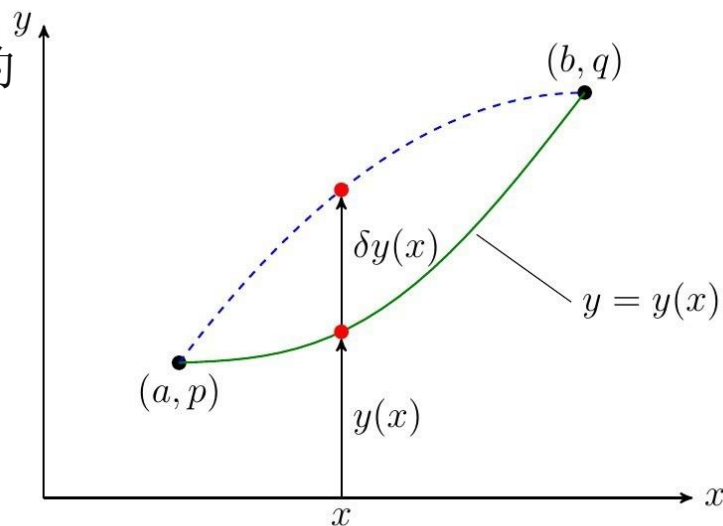
假设我们有两个定点  $(a,p)$  和  $(b,q)$ ，连接这两点的任意曲线的方程  $y=y(x)$  都将满足如下的边界条件：

$$y(a) = p, \quad y(b) = q \quad (1)$$

现在考虑如下形式的定积分：

$$I = \int_a^b f(y, y') dx \quad (2)$$

其中  $f(y, y')$  是关于  $y(x)$  和其一阶导数  $y'(x)$  的函数，我们期望找到一个具体的  $y(x)$ ，使得  $I$  有极值（极大或极小）。



### 注意：

- 在一般的极值问题中，我们考察的是自变量  $x$  的变化,  $x$  取值多少时，函数会有极值。
- 而现在这个新问题的不同之处，我们考察的是函数  $y(x)$  的变化,  $y(x)$  是什么形式时， $I$  会有极值（高大上叫法： $I$  称作函数  $y(x)$  的泛函）。
- 这两类问题依然有共通之处：当  $I$  取极值时，对  $y(x)$  作微小的变化， $I$  在一级近似下应该保持不变。

如果  $y(x)$  有微小改变  $\delta y(x)$  (高大上叫法:  $\delta y(x)$  称作函数  $y(x)$  的**变分**) , 那么  $f(y, y')$  的变化为:

$$\delta f = \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y'} \delta y' \quad (3)$$

$I$  相应的变化为:

$$\delta I = \int_a^b \left[ \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y'} \delta y' \right] dx \quad (4)$$

方括号里的第二项可以改写成  $\frac{\partial f}{\partial y'} \frac{d(\delta y)}{dx}$  , 然后我们可以进行分部积分

$$\begin{aligned} \int_a^b \frac{\partial f}{\partial y'} \delta y' dx &= \int_a^b \frac{\partial f}{\partial y'} d(\delta y) \\ &= \frac{\partial f}{\partial y'} \delta y \Big|_a^b - \int_a^b \delta y \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) dx \end{aligned}$$

由于  $y(x)$  的边界条件固定,  $\delta y(a) = \delta y(b) = 0$ , 所以分部积分出来的第一项为零, 仅第二项有贡献。代回 (4) 式中, 稍作化简可以得到

$$\delta I = \int_a^b \left[ \frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \right] \delta y(x) dx \quad (5)$$

如果  $I$  有极值, 对任意满足边界条件的  $\delta y(x)$  都必须有  $\delta I = 0$ , 这就要求:

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) = 0} \quad (6)$$

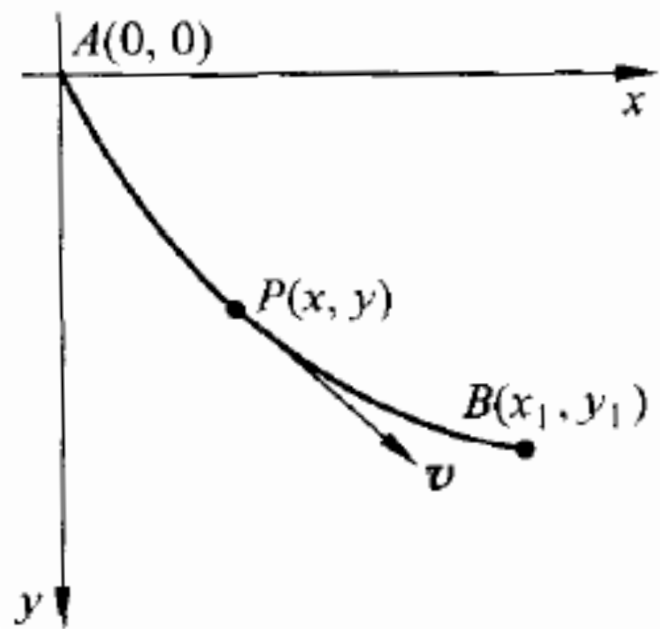
这便是传说中的 **Euler-Lagrange 方程**, 它是变分法的核心定理。有了此等大杀器, 原则上就可以找出所寻求的极值函数  $y(x)$ 。

通常来讲 Euler-Lagrange 方程会是一个二阶的微分方程,  $y(x)$  的通解中含有的两个待定常数刚好可以通过两个边界条件确定。我们下面来举几个例子操练操练。

- 例1 (最速下降曲线问题)确定一条曲线，它连结不在同一垂直线上的两个已知点**A**和**B**，使它具有下列性质：任何质点沿这条线由**A**滑到**B**时所需时间为最少。

$$A(0,0), B(x_1, y_1)$$

$$l : y = y(x), 0 \leq x \leq x_1$$



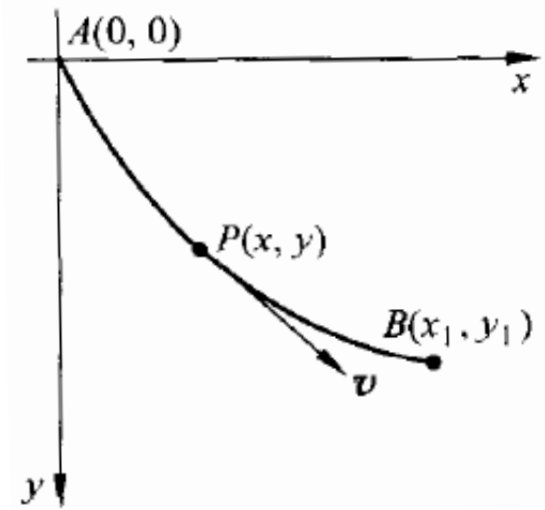
$$mgh = \frac{1}{2}mv^2 \Rightarrow \frac{1}{2}m\left(\frac{ds}{dt}\right)^2 = mgy$$

*because*  $ds = \sqrt{1 + y'^2} dx$


$$\Rightarrow \sqrt{1 + y'^2} \frac{dx}{dt} = \sqrt{2gy}$$

$$\Rightarrow dt = \sqrt{\frac{1 + y'^2}{2gy}} dx$$

$$\Rightarrow T = \int_0^T dt = \int_0^{x_1} \sqrt{\frac{1 + y'^2}{2gy}} dx$$




$$T = \int_0^T dt = \int_0^{x_1} \sqrt{\frac{1 + y'^2}{2gy}} dx$$



$$T(y) = \int_0^T dt = \int_0^{x_1} \sqrt{\frac{1 + y'^2}{2gy}} dx$$

$$K = \{y \mid y \in C^1[0, x_1], y(0) = 0, y(x_1) = y_1\}$$



$$\begin{cases} \text{find } y_0 \in K \text{ s.t.} \\ T(y_0) \leq T(y), \quad \forall y \in K \end{cases}$$

在集合K中求极小。

为了找出让  $T$  取得极小的  $y(x)$ , 我们可以取  $f(y, y') = \sqrt{\frac{1+y'^2}{y}}$ , 再套用 Euler-Lagrange 方程来怒算一波。

$$\frac{\partial f}{\partial y} = -\frac{1}{2} \sqrt{\frac{1+y'^2}{y^3}}, \quad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{y(1+y'^2)}}$$

丢回 (6) 式里面, 我们可以得到这么一个初步的方程:

$$\frac{1}{2} \sqrt{\frac{1+y'^2}{y^3}} + \frac{d}{dx} \left( \frac{y'}{\sqrt{y(1+y'^2)}} \right) = 0 \quad (14)$$

看到这种东西, 要保持平静, 铁了头往下算, 要相信好多恶心的东西会神奇地同归于尽。

$$\begin{aligned} \frac{1}{2} \sqrt{\frac{1+y'^2}{y^3}} + \frac{y''}{\sqrt{y(1+y'^2)}} - \frac{1}{2} \frac{y'^2}{\sqrt{y^3(1+y'^2)}} - \frac{y'^2 y''}{\sqrt{y(1+y'^2)^3}} &= 0 \\ \frac{1}{2} (1+y'^2)^2 + y(1+y'^2)y'' - \frac{1}{2} y'^2 (1+y'^2) - y y'^2 y'' &= 0 \\ \frac{1}{2} (1+y'^2) + y y'' &= 0 \end{aligned} \quad (15)$$



瞧，柳暗花明又一村。不过这还远没完，解这个二阶微分方程还需要一个骚操作。我们对上式乘上一个  $2y'$ ：

$$\begin{aligned}y'(1 + y'^2) + 2yy'y'' &= 0 \\ \frac{d}{dx} [y(1 + y'^2)] &= 0 \quad (16)\end{aligned}$$

感谢 CCAV 这玩意儿居然是个全微分，它要等于零，方括号里那一坨等于常数就完儿事了。且让我们将这个常数写作  $k$

$$\begin{aligned}y(1 + y'^2) &= k \\ y'^2 &= \frac{k}{y} - 1 = \frac{k - y}{y} \\ y' &= \frac{dy}{dx} = \sqrt{\frac{k - y}{y}} \quad (17)\end{aligned}$$

原来的二阶微分方程降次变成了一阶，我们终于可以愉快地分离变量两边积分了

$$x = \int dx = \int \sqrt{\frac{y}{k - y}} dy \quad (18)$$

作三角换元, 设  $y = k \sin^2 \theta$ , 则

$$\begin{aligned}x &= \int \sqrt{\frac{k \sin^2 \theta}{k - k \sin^2 \theta}} d(k \sin^2 \theta) \\x &= \int \frac{\sin \theta}{\cos \theta} 2k \sin \theta \cos \theta d\theta \\x &= \int 2k \sin^2 \theta d\theta = k \int (1 - \cos 2\theta) d\theta \\x &= k\theta - \frac{1}{2}k \sin 2\theta + c \quad (19)\end{aligned}$$

其中  $c$  是积分常数。我们再作逆变换变回到  $y$ , 注意到  $\sin \theta = \sqrt{\frac{y}{k}}$ , 于是

$$\sin 2\theta = 2 \sin \theta \cos \theta = 2\sqrt{\frac{y}{k}} \sqrt{\frac{k-y}{k}}$$

我们可以得到所求的最速降轨道的函数表达式:

$$x(y) = k \sin^{-1} \sqrt{\frac{y}{k}} - \sqrt{y(k-y)} + c \quad (20')$$

轨道起点为  $(0, 0)$ ，很容易得出  $c = 0$ ，于是结果可以进一步简化

$$x(y) = k \sin^{-1} \sqrt{\frac{y}{k}} - \sqrt{y(k-y)} \quad (20)$$

另一方面，轨道终点为  $(a, b)$ ，上面的常数  $k$  还必须满足：

$$k \sin^{-1} \sqrt{\frac{b}{k}} - \sqrt{b(k-b)} = a \quad (21)$$

可以证明，满足边界条件 (21) 的最速降曲线 (20) 将是一条摆线 (cycloid)：它是圆周上的一个定点在圆沿直线滚动时所形成的轨迹。

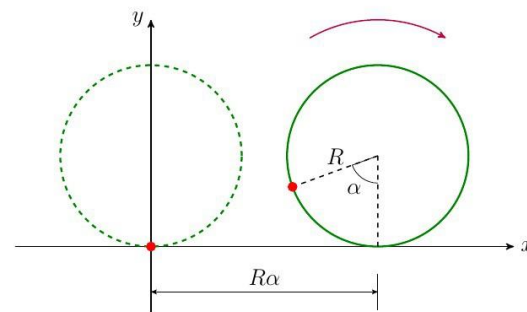
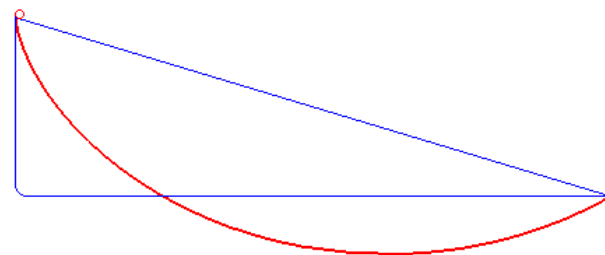
从 (20) 式很难看出这个结论，但其实满足 (17) 式的  $y(x)$  还可以写成如下的参数方程：

$$x(\alpha) = R(\alpha - \sin \alpha), \quad y(\alpha) = R(1 - \cos \alpha) \quad (22)$$

其中  $R = \frac{1}{2}k$  给出了滚动的圆的半径。可以作如下的验算：

$$\begin{aligned} \frac{dy}{dx} &= \frac{\frac{dy}{d\alpha}}{\frac{dx}{d\alpha}} = \frac{\sin \alpha}{1 - \cos \alpha} = \frac{\sqrt{1 - \cos^2 \alpha}}{1 - \cos \alpha} = \sqrt{\frac{1 + \cos \alpha}{1 - \cos \alpha}} \\ \sqrt{\frac{k-y}{y}} &= \sqrt{\frac{k - \frac{1}{2}k(1 - \cos \alpha)}{\frac{1}{2}k(1 - \cos \alpha)}} = \sqrt{\frac{1 + \cos \alpha}{1 - \cos \alpha}} \end{aligned}$$

这验证了 (22) 式的参数方程确实是 (17) 式的解。而参数方程可以更明确地表达出摆线的几何意义 (如图)。



例 2.最小曲面问题,  $x$ - $y$  平面上的有界开区域  $\Omega$  边界为  $\partial\Omega$ , 在边界  $\partial\Omega$  上给定条件  $u|_{\partial\Omega} = \varphi(x, y)$ , 记

$$S(u) = \iint_{\Omega} \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} dx dy$$

$$K = \{u \mid u \in C^1(\bar{\Omega}), u|_{\partial\Omega} = \varphi(x, y)\}$$

最小曲面问题可以写成下面的求泛函极小的问题

$$\begin{cases} \text{find } u_0 \in K \text{ s.t.} \\ S(u_0) \leq S(u), \quad \forall u \in K. \end{cases}$$

所谓**变分问题**，就是在一个函数集合中求泛函的极小或极大的问题。

附注：变分问题也可以求函数  $u \in M$ ，使得

$$J(u) = \max_{v \in M} J(v)$$

但  $\max J(v) = -\min \{-J(v)\}$ ，故求极大值问题可以化为求极小值问题处理。

## 2.变分法基本引理

引理：变分法基本引理，设  $u \in C[a,b]$

且

$$\int_a^b u(x)v(x)dx = 0, \forall v \in C_0^1[a,b]$$

则有  $u(x) \equiv 0, a \leq x \leq b$

### 3. $R^n$ 中的变分问题

(二元函数极值与线性代数方程组的关系)

命题 1.  $x \in R^n$ , satisfy  $Ax - b = 0$



命题 2.  $(Ax - b, y) = 0, \forall y \in R^n$

A 对称



A 对称

正定



命题 3.  $x \in R^n$ , s.t.  $J(x) \leq J(y), \forall y \in R^n$

其中  $J(x) = \frac{1}{2}(Ax, x) - (b, x)$

## (二) 一维数学物理问题的变分问题

$$Lu = -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u,$$

$$p \in C^1[a, b], p(x) \geq p_0 > 0, q \in C[a, b], q(x) \geq 0$$

(1) 考虑如下两点边值问题

$$(\mathbf{P}_1) \begin{cases} -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u = f(x), a < x < b \\ u(a) = 0, \frac{du(b)}{dx} = 0 \end{cases}$$

$$Lu - f = 0$$



$$(P_1) \Rightarrow (P_2)$$

$$(P_1) \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x), a < x < b \\ u(a) = 0, \frac{du(b)}{dx} = 0 \end{cases}$$

设  $u \in C^1[a, b] \cap C^2(a, b)$  为  $(P_1)$  的解

方程两边乘  $v(x)$  再积分，记内积为

$$(u, v) = \int_a^b u(x)v(x)dx$$

我们有  $(Lu - f, v) = 0$

$$\text{即 } -\int_a^b v \frac{d}{dx} \left( p \frac{du}{dx} \right) dx + \int_a^b quv dx - \int_a^b fvd x = 0$$

分部积分得 
$$\int_a^b p \frac{du}{dx} \frac{dv}{dx} dx + \left( p \frac{du}{dx} v \right) \Big|_{x=a} - \left( p \frac{du}{dx} v \right) \Big|_{x=b} + \int_a^b quv dx - \int_a^b f v dx = 0$$

记  $S_0^1 = \{v \mid v \in C^1[a, b], v(a) = 0\}$

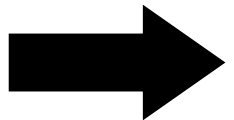
$\frac{du(b)}{dx} = v(a) = 0,$



*find*  $u, v \in S_0^1, \text{ s.t. }$

$$\int_a^b p \frac{du}{dx} \frac{dv}{dx} dx + \int_a^b quv dx - \int_a^b f v dx = 0$$

$$D(u, v) = \int_a^b \left( p \frac{du}{dx} \frac{dv}{dx} + quv \right) dx, \quad F(v) = \int_a^b f v dx$$



$$(\mathbf{P}_1) \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x), a < x < b \\ u(a) = 0, \frac{du(b)}{dx} = 0 \end{cases}$$

$$D(u, v) = \int_a^b \left( p \frac{du}{dx} \frac{dv}{dx} + quv \right) dx, \quad F(v) = \int_a^b f v dx$$

双线性型，对称正定



$$(P_2) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ D(u, v) - F(v) = 0, \quad \forall v \in S_0^1 \end{cases}$$

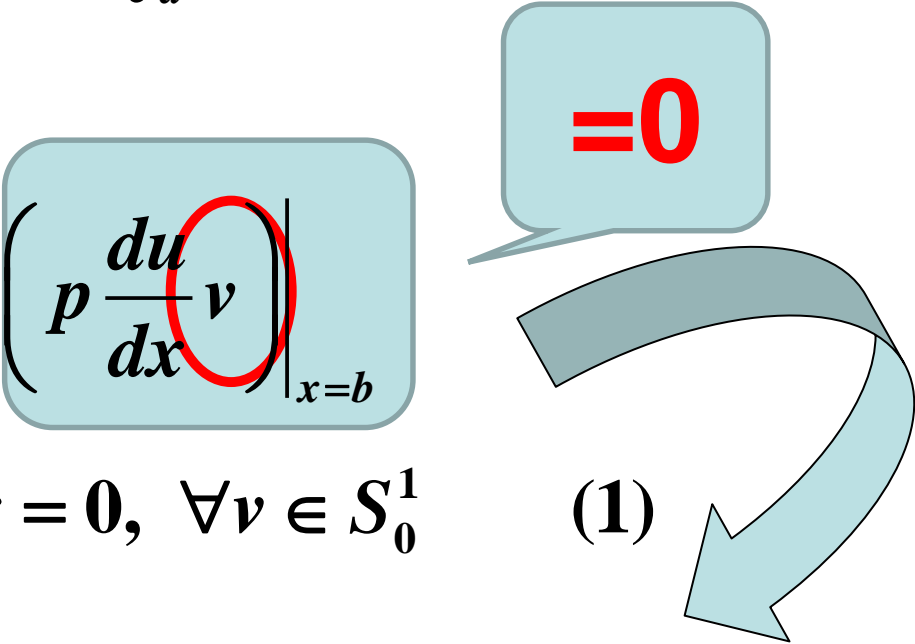
$$S_0^1 = \{v \mid v \in C^1[a, b], v(a) = 0\}$$

$$(P_2) \Rightarrow (P_1)$$

设  $u \in C^1[a,b] \cap C^2(a,b)$ ，则有

$$\int_a^b p \frac{du}{dx} \frac{dv}{dx} dx + \int_a^b quv dx - \int_a^b fvd x = 0, \quad \forall v \in S_0^1$$

分部积分得

$$\begin{aligned}
 & -\int_a^b v \frac{d}{dx} \left( p \frac{du}{dx} \right) dx + \left( p \frac{du}{dx} v \right) \Big|_{x=b} \\
 & + \int_a^b quv dx - \int_a^b fvd x = 0, \quad \forall v \in S_0^1
 \end{aligned}
 \tag{1}$$


取  $v \in V = C_0^1[a,b] = \{v \mid v \in S_0^1, v(b) = 0\} \in S_0^1$

$$S_0^1 = \{v \mid v \in C^1[a,b], v(a) = 0\}$$

$$-\int_a^b v \frac{d}{dx} \left( p \frac{du}{dx} \right) dx + \int_a^b quv dx - \int_a^b f v dx = 0, \quad v \in V = C_0^1[a, b]$$



$$\int_a^b (Lu - f) v dx = 0, \quad \forall v \in C_0^1[a, b]$$


变分学基本引理




$$Lu - f = 0 \quad u \text{ 满足微分方程}$$

将  $Lu - f = 0$  代入(1)

$$-\int_a^b v \frac{d}{dx} \left( p \frac{du}{dx} \right) dx + \left( p \frac{du}{dx} v \right) \Big|_{x=b} + \int_a^b quv dx - \int_a^b f v dx = 0, \quad \forall v \in S_0^1 \quad (1)$$


$$\left( p \frac{du}{dx} v \right) \Big|_{x=b} = 0$$

$p(b) \geq p_0 > 0$ , 由  $v(b)$  任意性


$$\frac{du}{dx} \Big|_{x=b} = 0$$

自然边界条件（包括第二类、第三类边界条件）

(1)  $u(a) = 0$ , 约束边界条件（本质边界条件）

(2)  $\frac{du}{dx} \Big|_{x=b} = 0$ , 自然边界条件

$$\begin{aligned}
 (P_1) & \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x), a \leq x \leq b \\ u(a) = 0, \left. \frac{du}{dx} \right|_{x=b} = 0 \end{cases} \\
 (P_2) & \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ D(u, v) - F(v) = 0, \forall v \in S_0^1 \end{cases}
 \end{aligned}$$

(1) 若  $u$  是  $(P_1)$  的解, 则  $u$  是  $(P_2)$  的解

(2) 若  $u$  是  $(P_2)$  的解, 且

$$u \in C^1[a, b] \cap C^2(a, b)$$

则  $u$  是  $(P_1)$  的解

$$(P_2) \Rightarrow (P_3)$$

$$\left[ \text{find } u \in S^1 \text{ s.t.} \right.$$

$$D(u, v) = \int_a^b \left( p \frac{du}{dx} \frac{dv}{dx} + quv \right) dx, \quad F(v) = \int_a^b f v dx = 0$$

$$\text{令 } J(u) = \frac{1}{2} D(u, u) - F(u)$$

$$= \frac{1}{2} \int_a^b \left[ p \left( \frac{du}{dx} \right)^2 + qu^2 - 2fu \right] dx$$

则它代表在  $(P_1)$  的边界条件下的总势能。

从最小势能原理出发，考虑下面的问题

$$(P_3) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ J(u) \leq J(w) \quad \forall w \in S_0^1 \end{cases}$$

可证  $D(\bullet, \bullet)$  对称正定则  
 $(P_2)(P_3)$  等价



$$(\mathbf{P}_1) \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x), a \leq x \leq b \\ u(a) = 0, \frac{du(b)}{dx} = 0 \end{cases}$$

力学平衡方程定解问题

$$(P_2) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ D(u, v) - F(v) = 0, \quad \forall v \in S_0^1 \end{cases}$$

对应于 (P1) 的  
Galerkin变分问题，来源于  
力学虚功原理

$$(\mathbf{P}_3) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ J(u) \leq J(w) \quad \forall w \in S_0^1 \end{cases}$$

Ritz变分问题，来源于最  
小势能原理

古典解

$$(\mathbf{P}_1) \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x), a \leq x \leq b \\ u(a) = 0, \frac{du(b)}{dx} = 0 \end{cases}$$

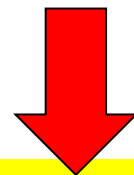
要处理二阶导数要求

$$u \in C^2(a, b) \cap C^1[a, b]$$

$$S_0^1 = \{v \mid v \in C^1[a, b], v(a) = 0\}$$

只要求  $v(a)=0$ （**约束边界条件**，或**本质边界条件**），对第二类第三类边界条件（**自然边界条件**）无强制要求。

$$(\mathbf{P}_2) \begin{cases} \text{find } u \in S_0^1, s.t. \\ D(u, v) - F(v) = 0, \forall v \in S_0^1 \end{cases}$$



$$(\mathbf{P}_3) \begin{cases} \text{find } u \in S_0^1, s.t. \\ J(u) \leq J(w) \quad \forall w \in S_0^1 \end{cases}$$

$$S_0^1 = \left\{ v \mid \int_a^b \left[ v^2 + \left( \frac{dv}{dx} \right)^2 \right] dx < \infty, v(a) = 0 \right\}$$

弱解（广义解）

### (三) 高维数学物理问题的变分问题

#### (1) 第一类边值问题的变分问题

$$(P_1) \Rightarrow (P_2)$$

设  $\Omega$  是一个有界区域,  $\partial\Omega$  是  $\Omega$  的边界, 考虑 Poisson 方程齐次边界条件的第一边值问题,

$$(P_1) \begin{cases} -\Delta u = f, (x, y) \in \Omega, \\ u|_{\partial\Omega} = 0. \end{cases}$$

$$C_0^1(\bar{\Omega}) = \{v \mid v \in C^1(\bar{\Omega}) \mid v|_{\partial\Omega} = 0\}$$

$$\text{取 } S_0^1 = C_0^1(\bar{\Omega}),$$

$$\Rightarrow -\iint_{\Omega} \Delta u \bullet v dx dy = \iint_{\Omega} f \bullet v dx dy, \quad \forall v \in S_0^1$$

推导用到 Green 公式

$$-\iint_{\Omega} \Delta u \bullet v dx dy = \iint_{\Omega} f \bullet v dx dy$$

$$\Rightarrow \iint_{\Omega} \Delta u \bullet v dx dy = - \iint_{\Omega} \nabla u \bullet \nabla v dx dy + \int_{\partial\Omega} \frac{\partial u}{\partial n} v ds$$

$$\text{得到} \int_{\Omega} \nabla u \bullet \nabla v dx dy - \int_{\partial\Omega} \frac{\partial u}{\partial n} v ds = \iint_{\Omega} f \bullet v dx dy, \quad \forall v \in S_0^1$$

$$\text{令 } D(u, v) = \int_{\Omega} \nabla u \nabla v dx dy, \quad F(v) = \int_{\Omega} f v dx dy$$

这样u满足的Garlerkin变分问题是

$$(P_2) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ D(u, v) - F(v) = 0, \quad \forall v \in S_0^1 \end{cases}$$

$$(P_2) \Rightarrow (P_3)$$

问题（P1）可以理解为固定边界膜平衡方程的定解问题，  
对应膜的势能是

$$(P_1) \begin{cases} -\Delta u = f, (x, y) \in \Omega, \\ u|_{\partial\Omega} = 0. \end{cases}$$

$$\begin{aligned} J(u) &= \frac{1}{2} D(u, u) - F(u) \\ &= \frac{1}{2} \iint_D \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 - 2fu \right] dx dy \end{aligned}$$

最小势能原理为

求

$$(P_3) \begin{cases} \text{find } u \in S_0^1 \\ J(u) \leq J(w), \quad \forall w \in S_0^1 \end{cases}$$

$(P_1)$  固定边界模的平衡问题

$$(P_1) \begin{cases} -\Delta u = f, (x, y) \in \Omega, \\ u|_{\partial\Omega} = 0. \end{cases}$$



$(P_2)$  虚功原理：平衡位置 $\mathbf{u}(\mathbf{x})$ 对任意满足齐次边界约束条件的虚位移、惯性力和外力所做功之和为零。

$$(P_2) \begin{cases} \text{find } u \in S_0^1, \text{ s.t.} \\ D(u, v) - F(v) = 0, \quad \forall v \in S_0^1 \end{cases}$$

$(P_3)$  最小势能原理：处于平衡位置的 $\mathbf{u}(\mathbf{x})$ 一定使 $J(\mathbf{u})$ 达到最小。

$$(P_3) \begin{cases} \text{find } u \in S_0^1 \\ J(u) \leq J(w), \quad \forall w \in S_0^1 \end{cases}$$

作业：考虑非齐次及第三类  
第二类边界条件

## 4.重调和方程边值问题的变分问题

$$-\iint_{\Omega} v \Delta u dx dy = \iint_{\Omega} \nabla u \bullet \nabla v dx dy - \oint_{\partial\Omega} v \frac{\partial u}{\partial \vec{n}} ds \quad (1)$$

$$-\iint_{\Omega} u \Delta v dx dy = \iint_{\Omega} \nabla u \bullet \nabla v dx dy - \oint_{\partial\Omega} u \frac{\partial v}{\partial \vec{n}} ds \quad (2)$$

$$(1) - (2) : \iint_{\Omega} (u \Delta v - v \Delta u) dx dy = \oint_{\partial\Omega} \left( u \frac{\partial v}{\partial \vec{n}} - v \frac{\partial u}{\partial \vec{n}} \right) ds$$

$$u \Rightarrow \Delta u : \iint_{\Omega} (\Delta u \Delta v - v \Delta^2 u) dx dy = \oint_{\partial\Omega} \left( \Delta u \frac{\partial v}{\partial \vec{n}} - v \frac{\partial \Delta u}{\partial \vec{n}} \right) ds$$

$$\text{即} \iint_{\Omega} v \Delta^2 u dx dy = \iint_{\Omega} \Delta u \Delta v dx dy + \oint_{\partial\Omega} v \frac{\partial \Delta u}{\partial \vec{n}} ds - \oint_{\partial\Omega} \Delta u \frac{\partial v}{\partial \vec{n}} ds,$$



$$(P_1) \Rightarrow (P_2) \Rightarrow (P_3)$$

$$(P_1) \begin{cases} \Delta^2 u = f, (x, y) \in \Omega \\ u|_{\partial\Omega} = 0, \frac{\partial u}{\partial \vec{n}} \Big|_{\partial\Omega} = 0 \end{cases}$$

$$\iint_{\Omega} v \Delta^2 u dx dy = \iint_{\Omega} \Delta u \Delta v dx dy + \oint_{\partial\Omega} v \frac{\partial \Delta u}{\partial \vec{n}} ds - \oint_{\partial\Omega} \Delta u \frac{\partial v}{\partial \vec{n}} ds, \\ \forall v \in S_0^2$$

$$(P_2) \begin{cases} \text{find } u \in S_0^2 .s.t. \\ D(u, v) - F(v) = 0, \forall v \in S_0^2 \end{cases}$$

$$\text{其中 } S_0^2 = \{v \mid v \in C^2(\bar{\Omega}), v|_{\partial\Omega} = 0, \frac{\partial v}{\partial \vec{n}} \Big|_{\partial\Omega} = 0\}$$

$$D(u, v) = \iint_{\Omega} \Delta u \Delta v dx dy, F(v) = \iint_{\Omega} f v dx dy$$

$$S_0^2 = \{v \mid v \in C^2(\bar{\Omega}), v|_{\partial\Omega} = 0, \frac{\partial u}{\partial \bar{n}} \Big|_{\partial\Omega} = 0\}$$

$$(\mathbf{P}_3) \begin{cases} \text{find } u \in S_0^2, s.t. \\ J(u) \leq J(w) \end{cases}$$

$$\text{其中 } J(w) = \frac{1}{2} \iint_{\Omega} (\Delta w)^2 dx dy - \iint_{\Omega} f w dx dy$$



## (四) 变分问题的近似计算

### 1. Ritz 方法(基于最小势能原理)

考虑 Ritz 形式的变分问题, 求

$$(\mathbf{P}_3) \begin{cases} \text{find } u \in K, s.t. \\ J(u) \leq J(w), \quad \forall w \in K \end{cases}$$

其中

$$J(w) = \frac{1}{2} D(w, w) - F(w)$$

$D(\bullet, \bullet)$  为  $K \times K$  上**对称正定**的双线性泛函,

$F(\bullet)$  为  $K$  上的线性泛函

$\mathbf{K}$  为无限维空间，求解有困难，考虑  $\mathbf{K}$  中的一个有限维的子空间  $S_N$ ，维数为  $N$ ，且

$$S_N = \text{span}\{\varphi_1, \varphi_2, \dots, \varphi_N\} \subset \mathbf{K}$$

$$\mathbf{w}_N = \sum_{i=1}^N c_i \varphi_i, \quad \forall \mathbf{w}_N \in S_N$$

其中  $c_1, c_2, \dots, c_N$  为常数，称  $S_N$  为试探函数空间。

以  $S_N$  代替  $\mathbf{K}$ ，得到  $(P_3) \begin{cases} \text{find } u \in K, s.t. \\ J(u) \leq J(w), \quad \forall w \in K \end{cases}$

的近似解为

$$(P_3') \begin{cases} \text{find } u_N \in S_N, s.t. \\ J(u_N) \leq J(w_N), \quad \forall w_N \in S_N \end{cases}$$

设  $w_N = \sum_{i=1}^N c_i \varphi_i$ , 是一个待定函数,

则  $(P_3)$  为求以下泛函的极小值点  $u_N = \sum_{i=1}^N c_i^0 \varphi_i$  :

$$\begin{aligned} J(w_N) &= \frac{1}{2} D(w_N, w_N) - F(w_N) \\ &= \frac{1}{2} D\left(\sum_{i=1}^N c_i \varphi_i, \sum_{j=1}^N c_j \varphi_j\right) - F\left(\sum_{i=1}^N c_i \varphi_i\right) \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N D(\varphi_i, \varphi_j) c_i c_j - \sum_{i=1}^N F(\varphi_i) c_i \end{aligned}$$

由于

$$\bar{c} = [c_1, c_2, \dots, c_N]^T \neq \bar{0} \Rightarrow w_N = \sum_{i=1}^N c_i \varphi_i \neq 0$$

$$D(\mathbf{w}_N, \mathbf{w}_N) = D\left(\sum_{i=1}^N c_i \varphi_i, \sum_{j=1}^N c_j \varphi_j\right)$$

$$= \bar{\mathbf{c}}^T \left[ D(\varphi_i, \varphi_j) \right] \bar{\mathbf{c}} > 0$$

$$D(\square, \square)$$

$\mathbf{K}$ 上对称正定的双线性泛函

所以  $[D(\varphi_i, \varphi_j)]$  为对称正定矩阵，求解近似变分问题

即为关于  $(c_0, c_1, \dots, c_N)$  求极小值点  $(c_0^0, c_1^0, \dots, c_N^0)$

$$\Rightarrow \left. \frac{\partial J(\mathbf{w}_N)}{\partial c_i} \right|_{(c_1^0, \dots, c_N^0)} = 0, \quad i = 1, \dots, N$$

$$\Rightarrow \sum_{j=1}^N D(\varphi_i, \varphi_j) c_j^0 = F(\varphi_i), \quad i = 1, \dots, N \quad (1)$$

求出  $c_1^0, c_2^0, \dots, c_N^0$  之后，即得近似变分问题的唯一解

$$u_N = \sum_{i=1}^N c_i^0 \varphi_i.$$

把无穷维空间  $V$  中的极值问题，近似为有限维空间  $S_N$  中的极值问题，这就是 **Ritz** 方法的基本思想。

最后通过求解一个线性代数方程组而得出变分问题的近似解。

## 2. Galerkin 方法

考虑 Galerkin 变分问题（基于虚功原理）

设  $(P_2) \begin{cases} \text{find } u \in K, s.t. \\ D(u, v) - F(v) = 0, \forall v \in K \end{cases}$

考虑  $K$  的有限维的子空间  $S_N$  来代替

求  $u_N = \sum_{j=1}^N c_j^0 \varphi_j,$

$$(P_2') \begin{cases} \text{find } u_N \in S_N, s.t. \\ D(u_N, v_N) - F(v_N) = 0, \forall v_N \in S_N \end{cases}$$

其中  $v_N = \sum_{i=1}^N c_i \varphi_i, \quad \forall v_N \in S_N$



$$D(u_N, v_N) - F(v_N) = 0$$

$$u_N = \sum_{j=1}^N c_j^0 \varphi_j, \quad v_N = \sum_{i=1}^N c_i \varphi_i,$$

代入  $D(u_N, v_N) - F(v_N) = 0$ ，得：

$$\sum_{i=1}^N \sum_{j=1}^N D(\varphi_j, \varphi_i) c_j^0 c_i - \sum_{i=1}^N F(\varphi_i) c_i = 0,$$

或者

$$\sum_{i=1}^N \left[ \sum_{j=1}^N D(\varphi_j, \varphi_i) c_j^0 - F(\varphi_i) \right] c_i = 0$$

$$\text{Ritz法: } \sum_{j=1}^N \mathbf{D}(\varphi_i, \varphi_j) c_j^0 = F(\varphi_i), \quad i = 1, \dots, N \quad (1)$$

由  $\mathbf{v}_N \in S_N$  的任意性, 即  $c_1, c_2, \dots, c_N$  的任意性, 可得出线性代数方程组

$$\sum_{j=1}^N \mathbf{D}(\varphi_j, \varphi_i) c_j^0 = F(\varphi_i), \quad (1 \leq i \leq N) \quad (2)$$

它与 **Ritz** 法所得的线性方程组相比, 系数矩阵的形式略有不同, 但目前我们假定双线性泛函  $\mathbf{D}(\mathbf{u}, \mathbf{v})$  对称, 故实际上这两个线性方程组是完全相同的。

$$(\textit{Galerkin}) \begin{cases} \textit{find } u_N \in S_N, s.t. \\ D(u_N, v_N) - F(v_N) = 0, \quad \forall v_N \in S_N \end{cases}$$

$$(\textit{Ritz}) \begin{cases} \textit{find } u_N \in S_N, s.t. \\ J(u_N) \leq J(w_N), \quad \forall w_N \in S_N \end{cases}$$

附注 1 如果不假定  $D(u,v)$  对称, 则变分问题  $(P_2)$  与变分方程  $(P_3)$  是不等价的, 他们对应的近似问题  $(P_2')$  与  $(P_3')$  也是不等价的。此时, 我们应用 **Ritz** 法得不到线性方程组(1), 但用 **Galerkin** 法仍然导出线性方程组 (2)。因此, Galerkin 法比 Ritz 法有更好的适应性, 无论双线性泛函  $D(u,v)$  是否对称都可以应用。

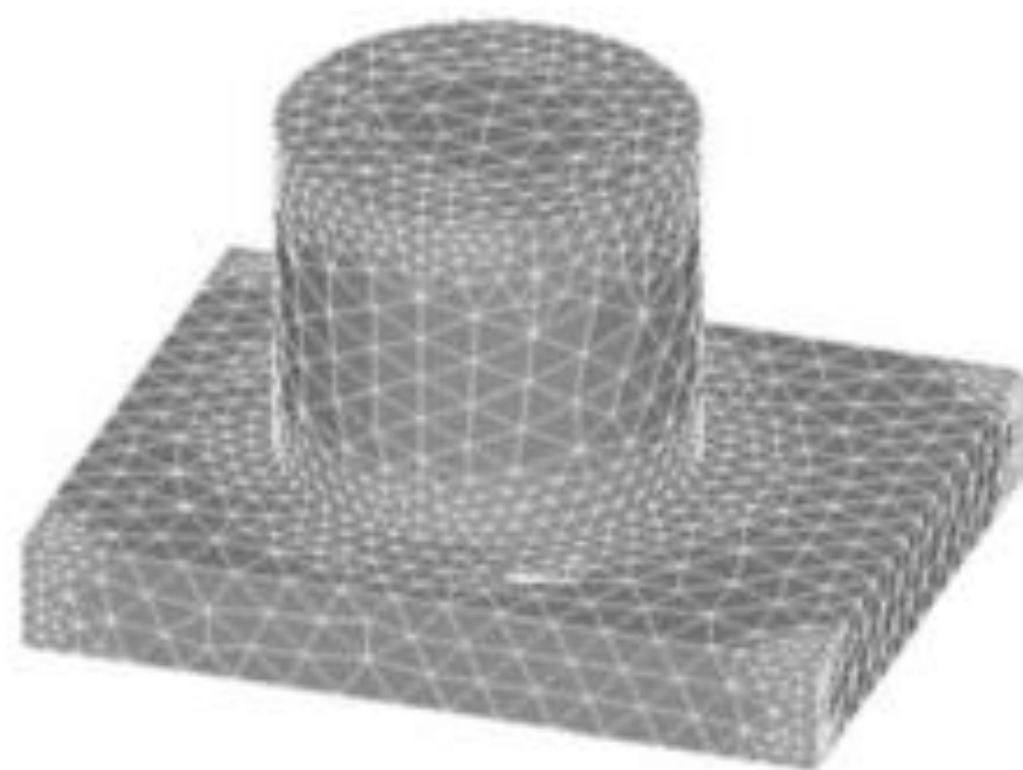
附注 2 无论应用 Ritz 法还是 Galerkin 法，选取子空间  $S_N$  的基  $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$  都不是容易做到的，对二维以上的不规则区域是很难选取的，需要有限元方法来克服这一困难。

**Ritz 法和 Galerkin 法都是在本世纪初提出来的，它们有如下一些优点：**

- 1.如果原问题中的双线性泛函  $D(u,v)$  对称正定，那么离散后得到的线性代数方程组的系数矩阵**仍然保持对称正定**；
- 2.这两种方法都是以变分原理为基础的，因此，椭圆形方程边值问题的第二、第三类边界条件和内边界条件都是自然边界条件，不必作为强制约束在  $V$  中出现，**避免了逼近边界条件中的法向导数所带来的困难。**

虽然 Ritz 法和 Galerkin 法有以上的优点，但在实际应用是还有很多具体**困难**：

1. 对边值问题的第一类边界条件，必须在  $V$  中作为强约束来考虑。如前面所指出的，在二维以上的情况下，当求解区域  $\Omega$  的形状不规则时**难于构造子空间  $S_N$** ；
2. 实际计算时， $D(\varphi_i, \varphi_j)$  和  $F(\varphi_i)$  都是在  $\Omega$  区域上的积分，计算量很大，而且矩阵  $[D(\varphi_i, \varphi_j)]$  **一般是稠密矩阵**，无法节省计算量和存贮量。



(b)

图5 零件自适应网格划分图

# 练习

## 微分方程的定解问题

$$\begin{cases} -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u = f(x), a < x < b \\ u(a) = 0, u(b) = 0 \end{cases}$$

其中  $p \in C^1[a, b]$ ,  $p(x) \geq p_0 > 0$ ,  $q \in C[a, b]$ ,  $q(x) \geq 0$

试推导对应的**Ritz**形式以及**Galerkin**形式的变分问题，  
叙述及证明三个问题之间的关系。



# 补充知识

## Euler-Lagrange 方程：回顾

若期望端点确定的泛函  $I = \int_a^b f(y, y')dx$  取得极值，则目标函数  $y(x)$  符合 **Euler-Lagrange 方程**

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) = 0 \quad (1)$$

如果泛函  $I$  取决于多个变量  $y_i$ ，即

$$I = \int f(y_1, y'_1; y_2, y'_2; \cdots; y_n, y'_n) dx$$

对于这其中的任何一个  $y_i$ ，都会有相应的 Euler-Lagrangian 方程：

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'_i} \right) = 0 \quad (2)$$

## 运动常数与守恒律

在上篇介绍变分原理的文章中，我举了3个求极值问题的例子，写下它们的 Euler-Lagrange 方程后，眼尖的读者很神奇地注意到了这么一件事情：这3个问题对应的二阶微分方程，形式上虽然一个比一个恶心，它们最终都可以降阶化成一个一阶的微分方程，从而使得我们可以顺利地解出目标函数的具体形式。

这其实并不是巧合。在这里，我们就来探讨一下这是为什么。

如果  $f(y, y')$  不显含  $x$ ，那么  $f$  的全微分将有

$$\frac{df}{dx} = \frac{\partial f}{\partial y} \frac{dy}{dx} + \frac{\partial f}{\partial y'} \frac{dy'}{dx} \Rightarrow \frac{df}{dx} = y' \frac{\partial f}{\partial y} + y'' \frac{\partial f}{\partial y'} \quad (3)$$

注意到在之前的例子中，化简出来的微分方程两边同乘以一个  $y'$  就可以化腐朽为神奇。由此启发，现在我们来考虑这么一件事情：如果在 Euler-Lagrange 方程乘上一个  $y'$ ，可以搞出什么名堂？

$$\begin{aligned} y' \left[ \frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \right] &= y' \frac{\partial f}{\partial y} - y' \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \\ &= y' \frac{\partial f}{\partial y} + y'' \frac{\partial f}{\partial y'} - y'' \frac{\partial f}{\partial y'} - y' \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \\ &= \frac{d}{dx} \left( f - y' \frac{\partial f}{\partial y'} \right) \quad (4) \end{aligned}$$

根据 Euler-Lagrange 方程，上式左边为零，这说明右边括号里那一坨必须是一个常数，对应的这个量

$$J = f - y' \frac{\partial f}{\partial y'} \tag{5}$$

被称为**运动常数**（constant of motion），或者叫**第一积分**（first integral）。它具有性质

$$\frac{dJ}{dx} = 0 \tag{6}$$

有了它，二阶的微分方程就可以降为一个更容易解出的一阶方程。

更普遍的，如果泛函  $I$  取决于多个变量  $y$ ，即

$$I = \int f(y_1, y_1'; y_2, y_2'; \cdots; y_n, y_n') dx$$

则相应的运动常数将变成

$$J = f - \sum_i y_i' \frac{\partial f}{\partial y_i'} \tag{7}$$

写到这里，熟悉分析力学的读者可能开始意识到这个运动常数背后所蕴含的意义。一个物理体系的作用量  $S = \int_{t_1}^{t_2} L(q_1, \dot{q}_1; q_2, \dot{q}_2; \cdots; q_n, \dot{q}_n; ) dt$  通常取决于质点们的坐标  $q_i$  以及速度  $\dot{q}_i$ ，而最小作用量原理告诉我们体系的运动方程将使得  $S$  有极值。

用变分原理的这套数学工具：我们将时间  $t$  视作  $x$ ，将质点坐标随时间的变换关系  $q(t)$  视作目标函数  $y(x)$ ，将 Lagrangian  $L(q, \dot{q})$  视作被积函数  $f(y, y')$ ，需要取得极值的作用量  $S$  则是我们的泛函  $I$ 。一通狸猫换太子操作之后，我们很容易得到**分析力学中的 Euler-Lagrange 方程**：

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = 0 \tag{8}$$

在分析力学中，我们可以引入广义动量

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{9}$$

从而定义**体系的 Hamiltonian**：

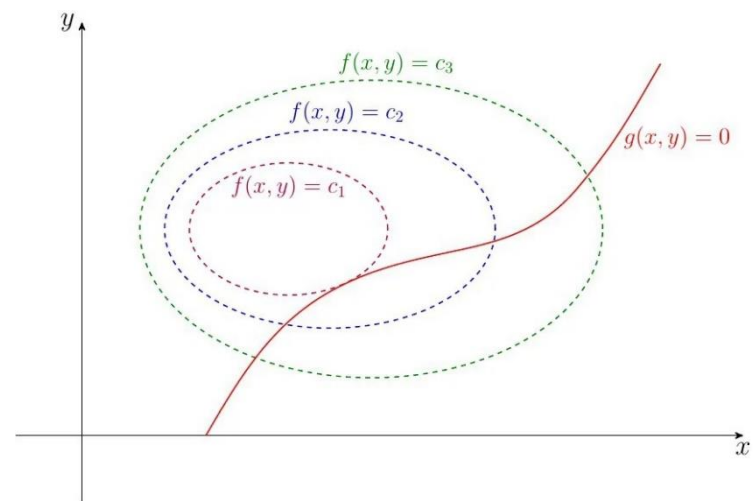
$$H = \sum_i p_i \dot{q}_i - L \tag{10}$$

可以证明（此处略去2000字），如果  $L$  能给出正确的牛顿定律， $L$  的形式大约是一个体系动能项和一个势能项的差值，由此搞出来的 Hamiltonian 会跟体系的能量有关。如果  $L$  不显含时间  $t$ ，也就是  $H$  不显含时间  $t$ ，可以证明  $H$  将是一个守恒量，对应的是一条能量守恒率。类比之下，是不是觉得 Hamiltonian 跟上面的运动常数  $J$  有那么一腿？

让我们索性把  $H$  写写开，有

$$H = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \tag{10'}$$

这不就相当于把  $J$  里边的两项的正负号翻一翻么！所以妙不妙！所谓的运动常数，其实和物理中的能量守恒有着千丝万缕的联系！



## 拉格朗日乘子法

在用变分法求解极值问题时，我们需要寻找的目标函数时常还会受到一些**约束条件** (constraint) 的限制。在上篇文章中，我举了一个简化版的悬链线的例子：我们把链条耷拉在两个固定的支点上，并没有把链条的端点固定悬挂起来，也就是说所解的问题是在两个支点间长度可变的链条的形状。眼尖的读者在评论区很自然地想到，如果就是固定在两个定点间的定长链条，它的形状又该如何求解？这个问题就额外附加了一个绳长恒定的约束条件。而这类受约束的极值问题，可以很漂亮地通过引入拉格朗日乘子来解决。

先来考虑一个二维平面内的问题：如何来寻找函数  $f(x, y)$  在约束条件  $g(x, y) = 0$  限制下的极值？

我们可以想象  $f(x, y) = c$  给出了平面内的一堆**等值线** (contour lines)，而  $g(x, y) = 0$  给出了可取的坐标  $(x, y)$  描绘出的轨迹。如果我们在  $g(x, y) = 0$  描述的曲线上游荡，就会从一条  $f(x, y)$  的等值线，游荡到另一条等值线。当  $f(x, y)$  取得极值  $M$  时，可以看到  $g(x, y) = 0$  的曲线和  $f(x, y) = M$  的曲线刚好相切。

用数学点的语言,  $f(x, y)$  的极值点  $M$  上有:

$$df|_M = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0 \quad (11)$$

但由于约束条件的限制,  $(x, y)$  不能从  $g(x, y) = 0$  这条曲线上下来, 因此还需要有:

$$dg|_M = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy = 0 \quad (12)$$

在极值点处, 现在我们可以把  $d\mathbf{r} = (dx, dy)$  视作可以到处撒丫子乱跑的无穷小位移, 但是必须让 (11) 和 (12) 同时成立, 于是这个方程组的系数必须是线性相关的, 即:

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y} \quad (13)$$

或者写作:

$$\nabla f = \lambda \cdot \nabla g \quad (14)$$

上式的几何意义是: 如果  $f(x, y) = M$  和  $g(x, y) = 0$  相切, 它们各自的法向量一定互相平行 (注意现在还是一个二维平面内的问题), 所以一个家伙的法向量一定可以写成另一个家伙的法向量的倍数形式。

推广到高维问题: 如果还有多于一条的约束条件限制, 那么函数  $f(x; y_1, y_2, \dots, y_m)$  取得极值  $M$  时, 会与所有约束条件  $g_j(x; y_1, y_2, \dots, y_m) = 0$  所描述的曲面相切。换言之,  $f$  在极值点处的法向量, 可以表示成一系列  $g_j = 0$  的法向量的线性组合:

$$\nabla f = \sum_j \lambda_j \nabla g_j \quad (14')$$

这些系数  $\lambda_i$  被称为**拉格朗日乘子** (Lagrangian multipliers)。稍作变换, 我们可以写出

$$\nabla \left( f - \sum_j \lambda_j g_j \right) \cdot d\mathbf{r} = 0 \quad (15)$$

由此引入辅助函数

$$F \equiv f - \sum_j \lambda_j g_j \quad (16)$$

则**最开始要找的约束条件下的极值问题, 就转化为了这个辅助函数的极值问题**。这个辅助函数不仅取决于原来的变量, 还取决于引入的这堆拉格朗日乘子  $\lambda_j$ 。但这么做的好处在于, 原来的约束条件全都自然地隐藏在了这个辅助函数里:

$$\frac{\partial F}{\partial \lambda_j} = 0 \quad \Leftrightarrow \quad g_j = 0 \quad (17)$$

我们接着就可以利用之前变分大法已经得到的结果, 来对付这个  $F$ 。当它取得极值时, 必然有:

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx} \left( \frac{\partial F}{\partial y'_i} \right) = 0 \quad (18)$$

## 举例：悬链线

我们重新回到悬链线（catenary）的问题。这次我们不再投机取巧，而是让链条的两端牢牢地被固定在等高的两点  $(\pm L, h)$  之间。

平衡状态下，体系有势能最小，所以我们要寻找以下函数的极小值：

$$V = \int_{-L}^L \rho g y \sqrt{1 + y'^2} dx \quad (19)$$

若线密度  $\rho$  是常数，因此我们其实只需要找到以下函数的极小值：

$$v = \int_{-L}^L y \sqrt{1 + y'^2} dx \quad (19')$$

它需要满足的约束条件是绳长固定：

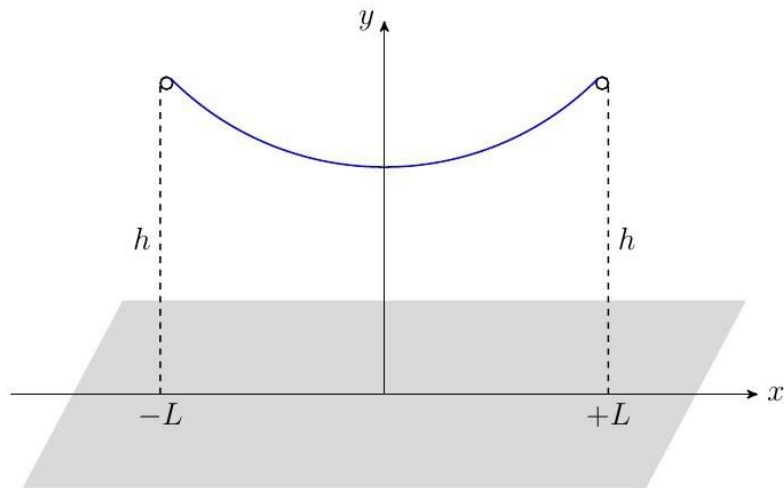
$$s = \int_{-L}^L \sqrt{1 + y'^2} dx = \text{const} \quad (20)$$

引入拉格朗日乘子的辅助函数可以取作

$$F(y, y', \lambda) = v - \lambda s = \int_{-L}^L (y - \lambda) \sqrt{1 + y'^2} dx \quad (21)$$

利用我们本篇第一部分提到的运动常数的结论，将上面的被积函数  $(y - \lambda) \sqrt{1 + y'^2}$  当作  $f$ ，可以试着计算

$$\begin{aligned} J &= f - y' \frac{\partial f}{\partial y'} \\ &= (y - \lambda) \sqrt{1 + y'^2} - y' \cdot \frac{(y - \lambda) y'}{\sqrt{1 + y'^2}} \\ &= \frac{y - \lambda}{\sqrt{1 + y'^2}} \end{aligned} \quad (22)$$



利用 (6) 式  $\frac{dJ}{dx} = 0$ ，马上得到：

$$\frac{d}{dx} \left( \frac{y - \lambda}{\sqrt{1 + y'^2}} \right) = 0 \quad (23)$$

当然，头铁的读者也可以尝试将 (21) 代入到 Euler-Lagrange 方程中进行一通计算（可参考上篇文章的计算过程，非常类似），也可以化简成完全相同的形式。解出这个微分方程，可以证明曲线的形状会满足方程：

$$y(x) = c \cosh \frac{x}{c} + \lambda \quad (24)$$

其中参数  $c$  和  $\lambda$  的选取要匹配边界条件  $y(L) = h$  以及绳子的长度 (20) 式。我们看到，悬链线依然符合双曲函数的形状。