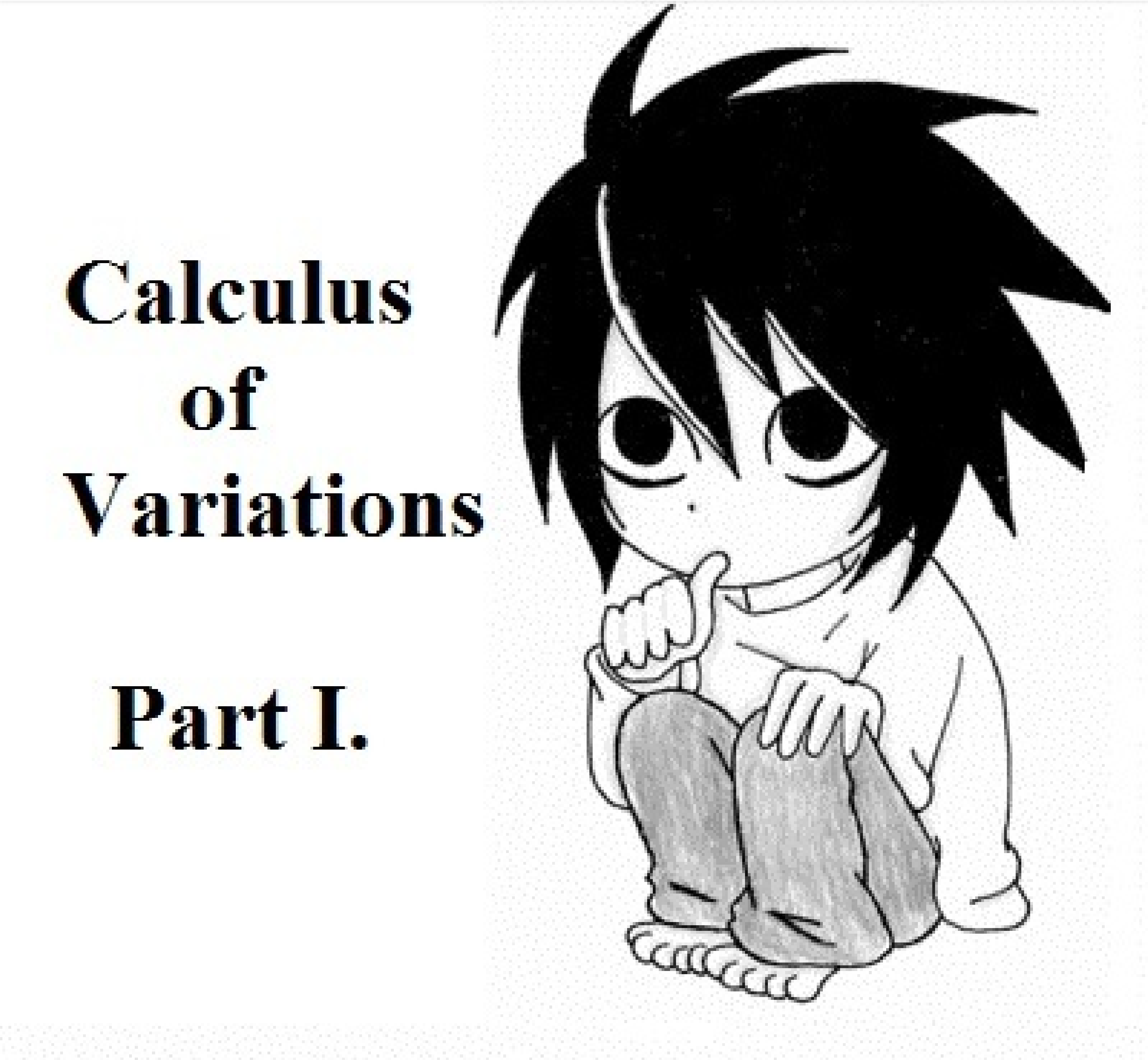


变分法简介Part 1.（Calculus of Variations）

Dr.Stein
计算力学

1. 泛函数 (Functionals)

简而言之，泛函数是函数的函数，即它的输入是函数，输出是实数。而这个输出值取决于一个或多个函数（输入）在一整个路径上的积分而非像一般函数一样取决于离散的变量。这样说可能还是比较抽象，不过坚持看到下文的*Example 1*就可以更好理解了。

通常在变分法中，泛函数是一个积分，记做 I。

$$I(y) = \int_{x1}^{x2} F dx$$

为了说明方便，我们先姑且设 F 是 y 和 y' 的函数，所以我们可以进一步将泛函数 I 写成：

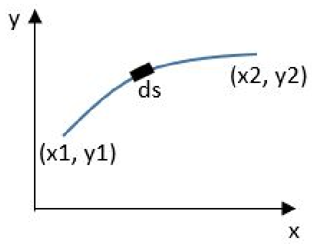
$$I(y)=\int_{x1}^{x2}F(y,y';x)dx$$

积分里面我用分号;将 x 和前面的 y 隔开代表 y 和 y' 是 x 的函数。一般 F 和 y 的函数关系是已知的，所以想要最大或最小化泛函数，实际上是通过选择适当的函数 $y(x)$ 。

为了透彻理解这个概念，我们可以来看一个简单的例子。

Example 1.

一个最简单直观的例子是求两个固定点之间的最短路径。当然大家都知道两点之间直线最短，这里可以用变分法做出解释。



如上图所示路径是一任意路径，我们取区中一小段微元 ds ，可以容易计算微元断的长度为：

$ds^2=dx^2+dy^2=[1+(y')^2]dx^2$,即：

$$ds=\sqrt{1+(y')^2}dx$$

积分得到总的路径长度为：

$$L=\int_{x1}^{x2}ds=\int_{x1}^{x2}\sqrt{1+(y')^2}dx$$

这个例子中， L 是泛函数， $\sqrt{1+(y')^2}$ 是拉格朗日函数 F ，我们想要找一个函数 $y(x)$ 使得泛函数 L 的值最小。这次Part 1.的任务就是为解决这个问题做准备。Part 2.中我们会用变分法证明这个 $y(x)$ 确实是直线的方程。

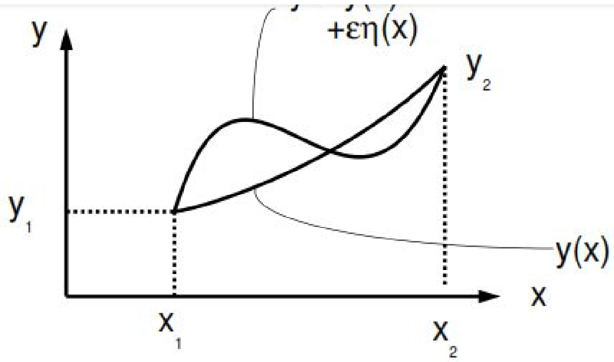
2. 泛函数的极值

这里重申下，泛函数 I 在区间 $[x_1,x_2]$ 上的值取决于积分路径的选择，即取决于函数 $y(x)$ 的选择。我们有理由假设存在一个这样的 $y(x)$ ，可以使得泛函数 I 取到极值。而在这个 $y(x)$ 附近的任意路径我们记做 $\tilde{y}(x)$ 。另外，我们假设 $y(x)$ 两阶可微。通过引入一个微小量 $\epsilon\ll 1$ 和一个任意可微函数 $\eta(x)$ ，我们可以用 $y(x)$ 表示 $\tilde{y}(x)$:

$$\tilde{y}(x)=y(x)+\epsilon\eta(x)$$

这样做的好处是对于一个给定的 $\eta(x)$ ，我们可以通过改变 ϵ 的值来得到无穷多的路径，同时对于任何 $\eta(x)$ ，当 $\epsilon=0$ 的时候， $\tilde{y}(x)$ 和 $y(x)$ 重合。

图像直观表示如下图：



由于在边界条件的限制，$\eta(x_1)=\eta(x_2)=0$。这样就能保证 $\tilde{y}(x)$ 可以通过两个固定端点。

这时我们可以说，$y(x)$ 所对应的泛函数 I 的值是泛函数 $\tilde{I}=\int_{x_1}^{x_2}F(\tilde{y},\tilde{y}';x)dx$ 的极值。我们可以进一步用 ϵ 表示 $\tilde{I}$：

$$\tilde{I}=\int_{x_1}^{x_2}F(\tilde{y},\tilde{y}';x)dx=\int_{x1}^{x2}F(y+\epsilon\eta,y'+\epsilon\eta';x)dx$$

虽然 $y(x)$ 未知，但是根据之前的合理假设，$y(x)$ 是一个存在的确定函数。所以根据上式，如果给定一个特定的 $\eta(x)$，$\tilde{I}$ 的变化只取决于 ϵ 的变化。所以我们现在可以把 $\tilde{I}$ 看做是 ϵ 的函数。用泰勒展开公式将 $\tilde{I}$ 在 $\epsilon=0$ 处展开得到：

$$\tilde{I}(\epsilon)=\tilde{I}|_{\epsilon=0}+(\frac{d\tilde{I}}{d\epsilon})\Big|_{\epsilon=0}\cdot\epsilon+(\frac{d^2\tilde{I}}{d\epsilon^2})\Big|_{\epsilon=0}\cdot\frac{\epsilon^2}{2!}+\cdots=\tilde{I}_0+\tilde{I}_1\epsilon+\tilde{I}_2\epsilon^2+\cdots$$

很明显，当 $\epsilon=0$ 时，$\tilde{I}|_{\epsilon=0}=I$,带入上式可得到：

$$\tilde{I}-I=\tilde{I}_1\epsilon+\tilde{I}_2\epsilon^2+\cdots$$

这里我们记 $\delta I=\tilde{I}_1\epsilon=\frac{d\tilde{I}}{d\epsilon}\Big|_{\epsilon=0}\cdot\epsilon$，并称之为一阶变分。同理二阶变分为 $\delta I^2=\tilde{I}_2\epsilon^2$。

（这里插一句变分和微分的区别。变分在上图的直观解释是 $\tilde{y}$ 和 y 在竖直方向上的距离，称之为 δy，所以这个差是在同一个x上计算的。而微分则是由于x的微小变动引起的y的变动。）

然后我们可以类比求函数极值时的做法。求函数极值时，我们会令函数的一阶导数为零。这里同样，为了求泛函数 $\tilde{I}$ 的极值，我们令一阶变分 $\delta I=0$。现在我们计算化简 δI：

$$\delta I=(\int_{x_1}^{x_2}\frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0}dx)\cdot\epsilon$$

$$\frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0}\cdot\epsilon=(\frac{\partial\tilde{F}}{\partial\tilde{y}}\cdot\frac{d\tilde{y}}{d\epsilon}+\frac{\partial\tilde{F}}{\partial\tilde{y}'}\cdot\frac{d\tilde{y}'}{d\epsilon})\Big|_{\epsilon=0}\cdot\epsilon$$

因为 $\tilde{y}(x)=y(x)+\epsilon\eta(x)$, 不难得到：$\frac{d\tilde{y}}{d\epsilon}=\eta,\frac{d\tilde{y}'}{d\epsilon}=\eta'$,另外我们有

$$\delta y=\epsilon\eta,\delta y'=\epsilon\eta'$$

又因为当 $\epsilon\to0$ 时，$\tilde{F}\to0,\tilde{y}\to y,\tilde{y}'\to y'$,将这些式子带入原式可以得到：

$$\delta I=\int_{x_1}^{x_2}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')dx$$

终于到最后一步啦，分部积分一下得到：

$$\delta I=\int_{x_1}^{x_2}(\frac{\partial F}{\partial y}-\frac{d}{dx}(\frac{\partial F}{\partial y'}))\delta ydx+\frac{\partial F}{\partial y'}\delta y\Big|_{x_1}^{x_2}$$

另 $\delta I=0$ 就可以解得最小化泛函数的y啦。我们注意到 δI 有两个部分。对于第一个积分部分，由于 δy 是任意的，所以要想使这个部分等于零，需要保证 [1] ：

这就是传说中的欧拉-拉格朗日方程(E-L equation)。

而第二部分等于零则是边界条件。

在Part 2.， 我们会以用这次介绍的内容和上述方程解决两点之间直线最短的问题为开头，继续介绍变分法。

注[1]:

假设 x_1 和 x_2 是给定的常数， $\phi(x)$ 是一个特定的在 $x_1 \leq x \leq x_2$ 上连续的函数，那么如果对于任意连续可微的函数 $\eta(x)$ 都成立 $\int_{x_1}^{x_2} \phi(x)\eta(x)dx = 0$ ，则 $\phi(x) = 0$（$x_1 \leq x \leq x_2$ ）。

（任意函数和一个非零的特定函数的乘积仍是任意函数，由于

这个特定函数必须在这个区间上恒等于零使得乘积为零，这样可以保证积分为零。）

编辑于 2016-10-21

有限元分析（FEA） 固体力学 应用数学

以上为精选评论 ?

TyLIU 1 年前

没来得及看，先赞一个。这个领域写的人少

1

TyLIU 1 年前

清晰~

1

Dr.Stein (作者) 回复 TyLIU 1 年前

多谢捧场:)

1 查看对话

TyLIU 1 年前

最后一步，分部积分之后，欧拉拉格朗日方程引出部分，论证不够详细。如果可以证明得更细，则算变分法很好的入门资料。

赞

Dr.Stein (作者) 回复 TyLIU 1 年前

谢谢你的建议，已做注释

1 查看对话

ray home 1 年前

“很明显，当时，,带入上式可得到” 这句里x=0是不是写错了?应该是ε=0？

1

Dr.Stein (作者) 回复 ray home 1 年前

谢谢指正，已更正

2 查看对话

快使用双截棍巴拉 1 年前

可以推荐下这方面的书吗

1

萨塔妮亚 回复 快使用双截棍巴拉 1 年前

老大中 变分法基础

3 查看对话

快使用双截棍巴拉 回复 萨塔妮亚 1 年前

thanks

1 查看对话

夏筠峥 1 年前

您讲的很清楚，尤其是很基本的概念有涉及到，我一直没明白变分具体含义是什么，在这里终于解了很多疑惑。有个可能比较基本的问题：

“然后我们可以类比求函数极值时的做法。求函数极值时，我们会令函数的一阶导数为零。” 后面的第二个方程式的步骤不是很懂（没有学过偏微分）；

能解释一下拉格朗日方程的物理意义么？而不是纯表达式的推导。

1

夏筠峥 1 年前

Dr.Stein (作者) 回复 夏筠峥 1 年前

你好，第二个方程式用到了链式法则(chain rule)。如果y是a和b的函数，a和b都是x的函数，那么dy/dx=(dy/da)*(da/dx)+(dy/db)*(db/dx)，形式上看有点类似分数的约分。维百上有例子，看一下应该就理解了。
欧拉拉格朗日方程从数学上的意义来说，就是方程的解是可以让对应原泛函数取得极值的函数。如果赋予F, y, x不同的物理意义，方程也会有不同的物理意义。我比较熟悉的一个应用是在力学上。zh.wikipedia.org/wiki/%... 可以参看这个维百链接，物理意义是广义牛顿第二定律。

2 查看对话

Dr.Stein (作者) 回复 夏筠峥 1 年前

这里没有把变分当做微分，变分一直没有动，只是用分部积分把微分转移了，评论不好打公式，有需要我可以手写拍照私信发你具体过程

1 查看对话

夏筠峥 回复 Dr.Stein (作者) 1 年前

我感觉我理解的chain rule 是 dy/dx = dy/dz*dz/dx,但是这里是关于dy和dy'两次chain rule并相加，我感觉相加这里不是很能理解，是对任意的两个variable都可以这样进行chain rule么，比如y和y'这里是有一定关系的，也可以这样分开求导再相加么。
我查了拉格朗日乘数的推导方式，有用到gradient平行的概念（constrain function和objective function在极值地方gradient平行），所以我想问的是这个欧拉-拉格朗日方程(E-L equation)是否也能有相关physical meaning （几何学上的意义）。抱歉可能我问题表述的不是很清楚，还有谢谢您的及时回复！

赞 查看对话

Dr.Stein (作者) 回复 夏筠峥 1 年前

维百chain rule里有多元复合函数求导法则zh.wikipedia.org/wiki/%... 举个例子，f(x)=x+x^2. 我们可以选择子函数为g(x)=x, h(x)=x^2, 则f=g+h。f'=(df/dg)*g'+(df/dh)*h'=1+2x。我们也可以选择子函数为g(x)=x, I(x)=(1+x) ,则f=g*I。f'=(df/dg)*g'+(df/dI)*I'=1+x*2。是同样的结果，所以子函数的选择不是唯一的，但是要保证所选的子函数能完整的表示出原函数。
不好意思，E-L equation的几何意义我不清楚，帮你at下学理的人试试，@学习ing

赞 查看对话

Dr.Stein (作者) 回复 夏筠峥 1 年前

我现在才明白你上一条问的是什么意思...就像你上次写的，这个例子中的v相当于δy。所以dv相当于δy'dx。相信你应该理解d(f(x))=f'(x)dx。你这次写的那个分部积分的公式没有错，但是如果想更容易理解，把它写成∫uv'dx=uv-∫vu'dx。这样直接将v=δy带入就可以了。

赞 查看对话

缪禧臻 1 年前

正文倒数第三条公式的积分区域是x_1 到 x_2 不是 x^2 吧？

1

Dr.Stein (作者) 回复 缪禧臻 1 年前

多谢~赞下认真阅读的亲

1 查看对话

1 2 下一页