泛函变分与极值

王杰 肖诗麒 2020.9.22

一、从多元函数的微分入手

首先，为了方便理解泛函变分和极值的概念，有必要回顾一下函数的微分和极值。

对于多元函数 $f(\vec{x})=f(x_1,x_2,\dots,x_n)$，若满足如下条件：

$$\Delta f=f(\vec{x}+\Delta\vec{x})-f(\vec{x})=A(\vec{x},\Delta\vec{x})+o(|\Delta\vec{x}|)$$

其中 A 为Δx 的线性函数，则称 f 可微。若 $\exists\vec{x}_0\Rightarrow\forall\Delta\vec{x},A(\vec{x}_0,\Delta\vec{x})=0$，则称 x0 为 f 的一个驻点。如果驻点 x0 还满足加强条件 $\forall\Delta\vec{x},\Delta^2 f=f(\vec{x}_0+2\Delta\vec{x})-2f(\vec{x}_0+\Delta\vec{x})+f(\vec{x}_0)>0$，那么此处函数取到极小值；如果以上条件为小于 0，那么此处取到极小值。

注意到，多元函数 $f:R^n\rightarrow R$，是一个从 n 维向量空间到实数的映射。事实上，定义在某区间上的函数是一个无穷维向量空间[1]。泛函是一个函数到实数的映射，那么可以认为泛函是一个从无穷维向量空间到实数的映射。下面将过渡到泛函的变分和极值。

以下如果没有特别说明，“函数 f 的泛函”所指代的泛函最多涉及到 f 的一阶导数。

设 f(x)为区间[x1,x2]上的函数[2]，考虑 f 的泛函 $J[f]:f\rightarrow R$。如果进行类比，将此处的 x 比作上面的下标 i，x 处的函数值 f(x)比作上面的 xi，将 J 看成“无穷元”函数，则可以类比出泛函的可微和极值条件。若 J 满足如下条件：

$$\delta J=J[f+\delta f]-J[f]=F[f,\delta f]+R[f,\delta f]$$

其中 F 为δf 的线性泛函；且 $R=o(\min\{|\delta f|,|\delta f'|\})$，其中 $|\delta f|=\sqrt{\int_{x_1}^{x_2}\delta f^2(x)dx}$ 为函数δf 的“模”，则称 J 可微。

若 $\exists f_0(x)\Rightarrow\forall\delta f(x),F[f_0,\delta f]=0$，则泛函 J 达到“驻点”，并称 f0 为 J 的驻定曲线[3]。

如果驻定曲线 f0 满足加强的条件 $\forall\delta f,\delta^2 J=J[f_0+2\delta f]-2J[f_0+\delta f]+J[f_0]>0$，那么称此时泛函取极小值；如果以上条件为大于 0，那么称此时泛函取极大值。

二、泛函驻点的两种等价表述

阿诺尔德：若 $\exists f_0(x)\Rightarrow\forall\delta f(x),F[f_0,\delta f]=0$，则此时泛函 J 达到驻点。

袁业飞：若 $\exists f_0(x)\Rightarrow\forall\delta f(x),\dfrac{d}{d\varepsilon}J[f_0+\varepsilon\delta f]|_{\varepsilon=0}=0$，则此时泛函 J 达到驻点。

以上两种表述以姓名首字母排序。下面证明两种表述等价。

由阿推出袁：

如果阿诺尔德表述成立，则存在 f0 对于任意δf 有 $J[f_0+\delta f]-J[f_0]=R$

考虑到 R 是δf 的二阶或以上小量，则当ε较小时有 $J[f_0+\varepsilon\delta f]-J[f_0]=\varepsilon^2 R+a_3\varepsilon^3+\dots$

所以$\frac{d}{d\varepsilon}J[f_0+\varepsilon\delta f]|_{\varepsilon=0}=0$，袁业飞表述成立。

由袁推出阿：

如果袁业飞表述成立，则存在 f0 对于任意δf 有$\frac{d}{d\varepsilon}J[f_0+\varepsilon\delta f]|_{\varepsilon=0}=0$，因此

$$0=\frac{d}{d\varepsilon}J[f_0+\varepsilon\delta f]|_{\varepsilon=0}=\frac{d}{d\varepsilon}\{J[f_0]+F[f_0,\varepsilon\delta f]+R[f_0,\varepsilon\delta f]\}|_{\varepsilon=0}$$
$$=\frac{d}{d\varepsilon}\{\varepsilon F[f_0,\delta f]+R[f_0,\delta f]\varepsilon^2+a_3\varepsilon^3+\ldots)\}|_{\varepsilon=0}=F[f_0,\delta f]$$

所以阿诺尔德表述成立。

三、有关泛函驻点的定理

定理 1：区间[t1,t2]上函数 x 的泛函$J=\int_{t_1}^{t_2}L(x(t),\frac{dx}{dt},t)dt$，符合条件$x(t_1)=x_1,x(t_2)=x_2$（即端点处变分为 0）的驻定曲线满足欧拉-拉格朗日方程$\frac{d}{dt}[\frac{\partial L}{\partial(dx/dt)}]-\frac{\partial L}{\partial x}=0$。

该定理以及多宗量的形式讲义已经证明，不在赘述。但是一个额外命题值得考虑。

命题 1：如果$\forall\delta q_1,\ldots,\delta q_n\Rightarrow\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_i>0$，其中对 i，j 求和，那么泛函当达到驻点时取得极小值。

证明：考虑驻点处二阶变分

$$\delta^2 J=\int_{t_1}^{t_2}\delta^2 Ldt=\int_{t_1}^{t_2}(\frac{\partial^2 L}{\partial q_i\partial q_j}\delta q_i\delta q_j+2\frac{\partial^2 L}{\partial q_i\partial\dot{q}_j}\delta q_i\delta\dot{q}_j+\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_j)dt$$
$$=\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial q_i\partial q_j}\delta q_i\delta q_j dt+\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial q_i\partial\dot{q}_j}d(\delta q_i\delta q_j)+\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_j dt$$
$$=\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial q_i\partial q_j}\delta q_i\delta q_j dt-\int_{t_1}^{t_2}\frac{d}{dt}(\frac{\partial^2 L}{\partial q_i\partial\dot{q}_j})\delta q_i\delta q_j dt+\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_j dt$$
$$=\int_{t_1}^{t_2}\frac{\partial}{\partial q_i}[\frac{\partial L}{\partial q_j}-\frac{d}{dt}(\frac{\partial L}{\partial\dot{q}_j})]\delta q_i\delta q_j dt+\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_j dt=\int_{t_1}^{t_2}\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_j dt$$

所以，对于任意的 q 的变动，都有$\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\delta\dot{q}_i\delta\dot{q}_i>0$时，二阶变分恒正，因此取得极小值。

下面介绍定理 1 的两个推广。

定理 2：多元函数为宗量的泛函变分。闭区域 D 上的 n 元函数$\phi(x_1,\ldots,x_n)$的泛函$J=\int_D L(\phi,\frac{\partial\phi}{\partial x_i},x_i)dx_1\ldots dx_n$，若边界处函数值给定（即边界处变分为 0），则 J 的驻定函数

满足$\frac{\partial}{\partial x_i}(\frac{\partial L}{\partial \phi / \partial x_i})=\frac{\partial L}{\partial \phi}$，其中对 i 求和。

证明：记区域体积为 v，边界面积为 s。

$$\delta J=\int_D L(\phi+\delta\phi,\frac{\partial}{\partial x_i}(\phi+\delta\phi),x_i)dv-\int_D L(\phi,\frac{\partial \phi}{\partial x_i},x_i)dv$$

$$=\int_D \frac{\partial L}{\partial \phi}\delta\phi dv+\int_D \frac{\partial L}{\partial(\partial\phi / \partial x_i)}\frac{\partial \delta\phi}{\partial x_i}dv+R$$

$$=\int_D \frac{\partial L}{\partial \phi}\delta\phi dv+\int_D \frac{\partial}{\partial x_i}[\frac{\partial L}{\partial(\partial\phi / \partial x_i)}\delta\phi]dv-\int_D [\frac{\partial}{\partial x_i}\frac{\partial L}{\partial(\partial\phi / \partial x_i)}]\delta\phi dv+R$$

$$=\int_D \frac{\partial L}{\partial \phi}\delta\phi dv+\int_{\partial D} \frac{\partial L}{\partial(\partial\phi / \partial x_i)}\delta\phi ds_i-\int_D [\frac{\partial}{\partial x_i}\frac{\partial L}{\partial(\partial\phi / \partial x_i)}]\delta\phi dv+R$$

$$=\int_D [\frac{\partial L}{\partial \phi}-\frac{\partial}{\partial x_i}\frac{\partial L}{\partial(\partial\phi / \partial x_i)}]\delta\phi dv+R=F[\phi,\delta\phi]+R$$

如果存在ϕ对于任意的δϕ都有 F=0，那么必有$\frac{\partial}{\partial x_i}(\frac{\partial L}{\partial \phi / \partial x_i})=\frac{\partial L}{\partial \phi}$，即多元的欧拉-拉格朗日方程。对于多宗量情形，每个ϕ分别满足该方程。

定理 3：带有约束的变分。[t1,t2]上 n 个函数为宗量的泛函$J=\int_{t_1}^{t_2} L(x_i,\frac{dx_i}{dt},t)dt$，其中 i 在 1 到 n 之间取值，且各宗量端点处变分为 0。如果存在 m<n 个约束$f_j(x_i)$ [4]，其中 j 在 1 到 m 之间取值，那么使得 J 取极值的多个函数分别满足$\frac{d}{dt}[\frac{\partial L}{\partial(dx_i / dt)}]-\frac{\partial L}{\partial x_i}-\lambda_j\frac{\partial f_j}{\partial x_i}=0$，其中λ为拉格朗日乘子，且对指标 j 求和。

证明这个定理之前先回顾一下多元微积分条件极值的拉格朗日乘子法。该方法表明对于多元函数$F(x_i)$，若存在 m 个条件$f_j(x_i)$，则可以构造含有乘子的函数$F_1=F+\lambda_j f_j$，而驻点满足$\frac{\partial F_1}{\partial x_i}=\frac{\partial F}{\partial x_i}+\lambda_j\frac{\partial f_j}{\partial x_i}=0,\frac{\partial F_1}{\partial \lambda_j}=f_j=0$。下面仿造此方法证明定理 3。

证明：设泛函$J_1=J+\lambda_j\int_{t_1}^{t_2} f_j dt$，下面求 J1 的驻定曲线。

首先计算出变分$\delta J_1=\int_{t_1}^{t_2}\{\frac{\partial L}{\partial x_i}-\frac{d}{dt}[\frac{\partial L}{\partial(dx_i / dt)}]+\lambda_j\frac{\partial f_j}{\partial x_i}\}\delta x_i dt+R=F[x,\delta x]+R$，其中对 i，j 求和。

如果存在 x 使得 F 对于任意的δx 都为 0，那么$\frac{d}{dt}[\frac{\partial L}{\partial(dx_i / dt)}]-\frac{\partial L}{\partial x_i}-\lambda_j\frac{\partial f_j}{\partial x_i}=0$成立。这里

类似$\frac{\partial F_1}{\partial x_i}=\frac{\partial F}{\partial x_i}+\lambda_j\frac{\partial f_j}{\partial x_i}=0$。

最后$\frac{\partial J_1}{\partial \lambda_j}=\int_{t_1}^{t_1} f_j dt=0$自然满足，这里类似$\frac{\partial F_1}{\partial \lambda_j}=f_j=0$。

此外，可以验证命题 1 对有约束的情形也成立，唯一的区别在于取得极小值的是泛函 J1。

[1]可以根据线性空间的定义验证$f(R^n)$是一个无穷维向量空间。

[2]如果 f 是某个闭区域上的多元函数，也可以得到同样结论。
[3]表述参考阿诺尔德《经典力学的数学方法》，有一些符号和措辞上的改动。

[4]对于不可积约束或者叫非完整约束$c_{ji}\delta x_i=0$，也有类似的结论：

$$\frac{d}{dt}[\frac{\partial L}{\partial(dx_i/dt)}]-\frac{\partial L}{\partial x_i}-\lambda_j c_{ji}=0$$