CONSTRAINED VARIATIONAL PROBLEMS

LONG CHEN

ABSTRACT.

CONTENTS

1. CONSTRAINED OPTIMIZATION PROBLEMS

As an introdutory example, let us consider a two-dimensional constraint optimization problem

$$\min_{x\in\mathbb{R}^2} f(x) \quad \text{s.t. } g(x)=0 \tag{1}$$

We introduce the Lagrangian $L(x,\lambda)=f(x)+\lambda g(x)$ and find the critical points of L:

$$\nabla_x L = \nabla f(x) + \lambda \nabla g(x) = 0,$$
$$\partial_\lambda L = g(x) = 0.$$

In the non-constrained case, where $x=(x_1,x_2)$ is free to move in $\mathbb{R}^2$. The equation $g(x_1,x_2)=0$ defines a curve and $x=x(t)$ can only move along this curve. Suppose $x_0=x(0)$ and $x'(0)\neq 0$ is a local minimum. Then, we have:

$$\left.\frac{\mathrm{d}f(x(t))}{\mathrm{d}t}\right|_{t=0} = \nabla f(x_0)\cdot x'(0) = 0.$$

Moreover, since $g(x(t))=0$ for all t near 0, taking derivative leads to

$$\nabla g(x_0)\cdot x'(0)=0.$$

Therefore it implies that $\nabla f(x_0)$ is parallel to $\nabla g(x_0)$, i.e., $\exists\lambda\in\mathbb{R}$ such that

$$\nabla f(x_0)+\lambda\nabla g(x_0)=0.$$

See Fig. 1 for an illustration. When $\nabla g(x_0)\neq 0$, we can calculate λ as follows:

$$\lambda = -\frac{(\nabla f(x_0),\nabla g(x_0))}{\|\nabla g(x_0)\|^2}.$$

By introducing a parameter t, we can transform the problem into a one-dimensional non-constrained problem.

Date: June 3, 2024.

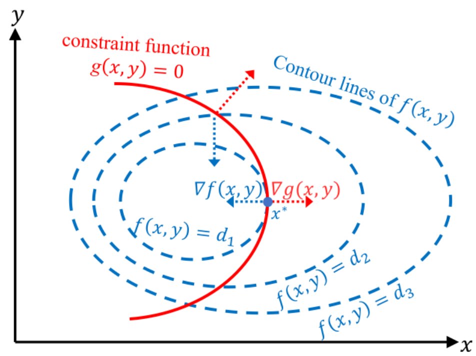

FIGURE 1. Minimization of function $f(x, y)$ subject to the constraint $g(x, y) = 0$. At the constrained local optimum, the gradients of f and g are parallel, i.e., $\nabla f + \lambda \nabla g = 0$.

2. INTEGRAL CONSTRAINT

We follow the book [1, Lecture 7] but simplify the presentation by introducing a parameterization. Given $L, G \in C^2\left(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN}\right), \rho \in C^1\left(\partial\Omega, \mathbb{R}^N\right)$, and

$$\mathcal{M} = \left\{u \in C^1\left(\bar{\Omega}, \mathbb{R}^N\right) \mid u|_{\partial\Omega} = \rho\right\},$$

consider the following constrained variational problem:

$$\begin{aligned} &\min_{u \in \mathcal{M}} I(u), \quad I(u) = \int_\Omega L(x, u, \nabla u)\mathrm{d}x, \\ &\text{s.t. } N(u) = 0, \quad N(u) = \int_\Omega G(x, u, \nabla u)\mathrm{d}x. \end{aligned} \tag{2}$$

Let $u \in \mathcal{M} \cap N^{-1}(0)$ and $\phi \in H_0^1(\Omega)$. The variation $u + \varepsilon\phi$ may not satisfy the constraint. To address this, we introduce variation in another direction $u + \varepsilon\phi + \tau\psi$. With a slight abuse of notation, we define

$$I(\varepsilon, \tau) = I(u + \varepsilon\phi + \tau\psi), \quad N(\varepsilon, \tau) = N(u + \varepsilon\phi + \tau\psi).$$

Now we face a situation similar to the 2D calculus example in Section 1. The two variables (ε, τ) are not free to choose due to the constraint. To satisfy the constraint, we need to eliminate one variable.

Assuming we can find a parameterization $(\varepsilon(t), \tau(t))$ such that $N(\varepsilon(t), \tau(t)) = 0$ for t near 0, and the minimum is achieved at $t = 0$ and $(\varepsilon(0), \tau(0)) = (0, 0)$, we then obtain a linear system:

$$\begin{cases} \left.\dfrac{\mathrm{d}}{\mathrm{d}t} I(\varepsilon(t), \tau(t))\right|_{t=0} = \delta I(u, \phi)\varepsilon'(0) + \delta I(u, \psi)\tau'(0) = 0, \\ \left.\dfrac{\mathrm{d}}{\mathrm{d}t} N(\varepsilon(t), \tau(t))\right|_{t=0} = \delta N(u, \phi)\varepsilon'(0) + \delta N(u, \psi)\tau'(0) = 0. \end{cases} \tag{3}$$

Assuming $\delta N(u,\psi) \neq 0$, we can solve $\tau'(0)$ from the second equation and substitute it back into the first equation to obtain:

$$\left[\delta I(u,\phi) - \frac{\delta I(u,\psi)}{\delta N(u,\psi)}\delta N(u,\phi)\right]\varepsilon'(0) = 0.$$

Assuming $\varepsilon'(0) \neq 0$, this implies that

$$\delta I(u,\phi) + \lambda\,\delta N(u,\phi) = 0,$$

where $\lambda = -\delta I(u,\psi)/\delta N(u,\psi)$.

Let us verify the assumption on the parameterization. We have $N(0,0) = 0$ and $\partial_\tau N(0,0) = \delta N(u,\psi) \neq 0$. Therefore, by the implicit function theorem, locally, i.e. for $|\varepsilon|$ sufficiently small, we can find a function $\tau = \tau(\varepsilon)$ such that $\tau(0) = 0$ and $N(\varepsilon, \tau(\varepsilon)) = 0$. The parameterization is given by $\varepsilon = t$, $\tau(t) = \tau(\varepsilon)$, and the derivative $\varepsilon'(0) = 1 \neq 0$.

Note that ψ is fixed, while ϕ is arbitrary. Thus, we arrive at the following result:

Theorem 2.1. *Suppose $N^{-1}(0) \cap \mathcal{M} \neq \emptyset$. Let $u \in \mathcal{M}$ be a weak minimum of $I(u)$ under the constraint $N(u) = 0$, i.e.,*

$$I(u) = \min_{w \in \mathcal{M} \cap N^{-1}(0)} I(w).$$

If there exists $\psi \in H_0^1(\Omega, \mathbb{R}^N)$ such that $\delta N(u,\psi) \neq 0$, then there exists $\lambda \in \mathbb{R}^1$ satisfying

$$\delta I(u,\phi) + \lambda\,\delta N(u,\phi) = 0, \quad \forall \phi \in H_0^1(\Omega, \mathbb{R}^N). \tag{4}$$

The first order necessary condition (4) can be derived by introducing a Lagrangian with multiplier λ

$$\mathcal{L}(u,\lambda) = L(x,u,\nabla u) + \lambda G(x,u,\nabla u)$$

and consider the inf-sup problem

$$\inf_{u \in \mathcal{M}} \sup_{\lambda \in \mathbb{R}} \int_\Omega \mathcal{L}(u,\lambda)\,\mathrm{d}x.$$

We include the existence result from Evan's book [2, Chapter 8]. Consider an integral constraint involving function only:

$$N(w) := \int_\Omega G(w) dx = 0$$

where $G : \mathbb{R} \to \mathbb{R}$ is a given, smooth function. Let us introduce as well the appropriate admissible class

$$\mathcal{A} := \left\{ w \in H_0^1(\Omega) \mid N(w) = 0 \right\}$$

Theorem 2.2 (Existence of constrained minimizer)**.** *Assume that L satisfies the coercivity inequality and is convex in the variable p. Assume the admissible set $\mathcal{A}$ is nonempty and the constraint satisfies*

$$|G'(z)| \leq C(|z| + 1)$$

for some constant C. Then there exists $u \in \mathcal{A}$ satisfying

$$I(u) = \min_{w \in \mathcal{A}} I(w).$$

Proof. We can choose a minimizing sequence $\{u_k\}_{k=1}^{\infty} \subset \mathcal{A}$ with

$$I\left[u_k\right] \rightarrow \inf_{w \in \mathcal{A}} I(w).$$

Then extract a subsequence

$$u_{k_j} \rightharpoonup u \text{ weakly in } H_0^1(U), \quad u_{k_j} \rightarrow u \text{ in } L^2(U).$$

The H^1-norm of $\{u_k\}$ is uniformly bounded and $I(u) \leq \inf_{w \in \mathcal{A}} I(w) = \liminf I(u_k)$.

We only need to verify $N(u) = 0$ so that $u \in \mathcal{A}$.

$$\begin{aligned}
|N(u)| = |N(u) - N(u_k)| &\leq \int_\Omega |G(u) - G(u_k)| \,\mathrm{d}x \\
&\leq C \int_\Omega |u - u_k| \left(1 + |u| + |u_k|\right) \mathrm{d}x \\
&\rightarrow 0 \quad \text{as } k \rightarrow \infty.
\end{aligned}$$

□

3. Pointwise Constraints

Now we consider the pointwise constraint $N(u(x)) = 0$ for all $x \in \Omega$, where $u : \Omega \rightarrow \mathbb{R}^n$ is a vector function and $N : \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth function. We still consider the following constrained variational problem:

$$\begin{aligned}
&\min_{u \in \mathcal{M}} I(u), \quad I(u) = \int_\Omega L(x, u, \nabla u) \mathrm{d}x, \\
&\text{s.t. } N(u) = 0.
\end{aligned} \tag{5}$$

with the Dirichlet boundary condition

$$\mathcal{M} = \left\{u \in C^1\left(\bar{\Omega}, \mathbb{R}^N\right) \mid u|_{\partial\Omega} = \rho\right\}.$$

Again we follow the book [1, Lecture 7] but simplify the presentation. Notice that u is a vector function and ∇u is a matrix as illustrated below

$$\underset{\mathbb{R}^d}{L(x,} \underset{\mathbb{R}^n}{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}}, \underset{\mathbb{R}^{n \times d}.}{\begin{pmatrix} -p_1- \\ -p_2- \\ \vdots \\ -p_n- \end{pmatrix}})$$

Let $u \in \mathcal{M} \cap N^{-1}(0)$ and $\phi \in H_0^1(\Omega)$. We first consider the non-constraint case with the following variation

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \rightarrow u + \varepsilon \phi_1 \vec{e}_1 := \begin{pmatrix} u_1 + \varepsilon \phi_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then we compute the varation

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} I\left(u + \varepsilon \phi_1 \vec{e}_1\right)\right|_{\varepsilon=0} = \int_\Omega L_{u_1} \phi_1 + L_{p_1} \cdot \nabla \phi_1 \,\mathrm{d}x = \int_\Omega \left(-\nabla \cdot L_{p_1} + L_{u_1}\right) \phi_1 \,\mathrm{d}x$$

So $\frac{d}{d\varepsilon} I\left(u+\varepsilon \phi_1 \vec{e}_1\right)\Big|_{\varepsilon=0}=0 \quad \forall \phi_1 \in H_0^1$ implies the Euler-Lagrange equation in multi-dimensions

$$-\nabla \cdot L_{p_i}+L_{u_i}=0, \quad i=1, \cdots, n,$$

where we change the index 1 to any index from $1: n$. It can be further simplified to

$$-\nabla \cdot L_p+L_u=0,$$

where L_p is a $n \times d$ matrix function and L_u is $n \times 1$ vector function and the divergence operator is applied row-wise.

In the constraint case, the variation $u+\varepsilon \phi_1 \vec{e}_1$ may not satisfy the constraint. To address this, we introduce a projection operator $P: H^1(\Omega) \rightarrow H^1(\Omega)$ s.t. $N(P(v))=0$ and $P(u)=u$ for $u \in \mathcal{M} \cap N^{-1}(0)$. Then we consider the first order condition

$$\frac{\mathrm{d}}{\mathrm{d} \varepsilon} I(x, P(u+\varepsilon \phi), \nabla_x P(u+\varepsilon \phi))\left.\right|_{\varepsilon=0}=0. \tag{6}$$

Here we use ∇_x to denote the derivative w.r.t to x and use ∇_u for the derivative w.r.t. u.

It remains to figure out the projection and its derivative. We assume the constraint is non-degenerate in the sense that $\nabla_x N(u(x)) \neq 0$, then locally one variable can be eliminated. More precisely, $\forall x_0 \in \Omega$, there exists a ball $B_r\left(x_0\right) \subset \Omega$ such that

$$\nabla_x(N(u(x)))=\nabla_u M(u(x)) \cdot \nabla u(x) \neq 0, \quad \forall x \in B_r\left(x_0\right).$$

Without loss of generality, we may assume $N_{u_n}(u(x)) \neq 0, \forall x \in B_r\left(x_0\right)$. Then by the implicit function theorem, there exists a C^2 function U s.t.

$$N\left(u_1+\varepsilon \phi_1, u_2, \cdots, U\left(u_1+\varepsilon \phi_1, \cdots, u_{n-1}\right)\right)=0. \tag{7}$$

Namely the projection P is given by

$$P((v_1, v_2, \ldots, v_n))=(v_1, v_2, \ldots, v_{n-1}, U(v^1, \ldots, v^{n-1})),$$

$$L(x,\left(\begin{array}{c}
u_1+\varepsilon \phi_1 \\
u_2 \\
\vdots \\
U\left(u_1+\varepsilon \phi_1, \cdots, u_{n-1}\right)
\end{array}\right),\left(\begin{array}{c}
\nabla_x u_1+\varepsilon \nabla_x \phi_1 \\
\nabla_x u_2 \\
\vdots \\
\nabla_x U\left(u_1+\varepsilon \phi_1, \cdots\right)
\end{array}\right))$$

The variation (6) is

$$\begin{aligned}
& L_{u_1} \phi_1+L_{p_1} \nabla \phi_1 \\
& +L_{u_n} \partial_{u_1} U \phi_1+L_{p_n} \partial_{\varepsilon} \nabla_x U\left(u_1+\varepsilon \phi_1, \cdots\right)
\end{aligned}$$

We can switch $\partial_{\varepsilon} \nabla_x$ to $\nabla_x \partial_{\varepsilon}$. So the last term is $L_{p_n} \cdot \nabla_x\left(\partial_{u_1} U \phi_1\right)$. Apply integration by parts, we get

$$-\nabla \cdot L_{p_1}+L u_1+\left(L_{u_n}-\nabla \cdot L_{p_n}\right) \partial_{u_1} U=0. \tag{8}$$

What is the partial derivative $\partial_{u_i} U$? We take derivative of (7) and obtain

$$N_{u_1}+N_{u_n} \partial_{u_1} U=0 \quad \Longrightarrow \quad \partial_{u_1} U=-\frac{N_{u_1}}{N_{u_n}}.$$

Plugging back to (8) and let

$$\lambda=\frac{1}{N_{u_n}}\left(\nabla \cdot L_{p_n}-L_{u_n}\right)$$

we arrive the following theorem.

Theorem 3.1. *Let* $\bar{\Omega} \subset \mathbb{R}^n$ *be a closed and bounded set. Let* $L \in C^2\left(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN}, \mathbb{R}^1\right), N \in C^2\left(\mathbb{R}^N, \mathbb{R}^1\right), \rho \in C^1\left(\partial\Omega, \mathbb{R}^N\right)$*, and*

$$\mathcal{M} = \left\{u \in H^1\left(\bar{\Omega}, \mathbb{R}^N\right) \mid u|_{\partial\Omega} = \rho\right\}.$$

Suppose $u \in \mathcal{M}$ *is a local minimum under the above constraint and it is* C^2 *outside finitely many* $(n-1)$ *dimensional piecewise* C^1 *hypersurfaces. If* $\forall x \in \bar{\Omega}$*,* $\nabla N\left(u(x)\right) \neq 0$*, then there exists a continuous function* $\lambda \in C(\bar{\Omega})$ *such that* u *satisfies the* E*-*L *equation of the adjusted Lagrangian* $\mathcal{L}(u, \lambda) = L + \lambda N$*:*

$$-\nabla \cdot L_{p_i} + L_{u^i} + \lambda N_{u^i} = 0, \quad 1 \leq i \leq n.$$

REFERENCES

[1] K.-C. Chang. *Lecture Notes on Calculus of Variations*, volume 6. World Scientific, 2016. 2, 4

[2] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1997. 3