Elastic Obstacle Problem

Andersen Ang
Dept. Combinatorics & Optimization, University of Waterloo, Canada
ms$\boldsymbol{x}$ang@uwaterloo.ca, where $\boldsymbol{x} = \lfloor \pi \rfloor$ Homepage: angms.science

First draft: November 18, 2022 Last update: November 23, 2022

Table of Contents

1. A minimization problem in function space
2. Elastic Obstacle Problem
3. Solving EOP

A minimization problem in function space

$$(\mathcal{P}_0) \;:\; \underset{u\left(x_1,\ldots,x_N\right)\geq\varphi(x_1,\ldots,x_N)}{\operatorname{argmin}} \int\cdots\int_\Omega L\Big(x_1,\ldots,x_N,\ u(x_1,\ldots,x_N),\ \frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_N}\Big)dx_1\ldots dx_N$$

- Task: find the functional $u(x_1,\ldots,x_N):\mathbb{R}^N\to\mathbb{R}$ of N variable that solves $\mathcal{P}_0$.
- $u(x_1,\ldots,x_N):\mathbb{R}^N\to\mathbb{R}$ is the optimization variable with coordinates $x_1,\ldots,x_N$
- $\varphi(x_1,\ldots,x_N):\mathbb{R}^N\to\mathbb{R}$ is a given function.
- $\Omega\subset\mathbb{R}^N$ is a given domain.
- $\dfrac{\partial u}{\partial x_i}$ is the ith partial derivative of u with respect to x_i.
- $L:\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}^N\to\mathbb{R}$ is a Lagrangian (for computing Euler-Lagrangian equation).
- $\mathcal{P}_0$ is a minimization over function space, so $u\in B$ for some appopriate Banach space B.

Be compact

$$(\mathcal{P}_0) : \underset{u(x_1,\ldots,x_N)\geq\varphi(x_1,\ldots,x_N)}{\operatorname{argmin}} \int\cdots\int_\Omega L\Big(x_1,\ldots,x_N,\ u(x_1,\ldots,x_N),\ \frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_N}\Big) dx_1\ldots dx_N$$

- Nobody write the inconvenient full expression.
- Let $\boldsymbol{x} = [x_1,\, x_2, \ldots,\, x_N]^\top$, we have

$$u = u(x_1,\ldots,x_N) = u(\boldsymbol{x}), \qquad \nabla u = \nabla_{\boldsymbol{x}} u(\boldsymbol{x}) = \Big[\frac{\partial u(\boldsymbol{x})}{\partial x_1},\ldots,\frac{\partial u(\boldsymbol{x})}{\partial x_N}\Big]^\top.$$

- Let $\mathrm{d}\boldsymbol{x}$ denotes the total differential[1], then $\mathcal{P}_0$ written compactly is

$$(\mathcal{P}) \ : \ \underset{u\geq\varphi}{\operatorname{argmin}} \int_\Omega L(\boldsymbol{x}, u, \nabla u)\,\mathrm{d}\boldsymbol{x}.$$

[1]Note that it is $\mathrm{d}\boldsymbol{x}$ not $d\boldsymbol{x}$, the d is upright and not italic.

A constrained minimization problem

$$(\mathcal{P}) \;:\; \operatorname*{argmin}_{u \geq \varphi} \int_\Omega L(\boldsymbol{x}, u, \nabla u)\, \mathrm{d}\boldsymbol{x}.$$

- Let $F(u) = \int_\Omega L(\boldsymbol{x}, u, \nabla u)\, \mathrm{d}\boldsymbol{x}$, then $\mathcal{P}$ is equivalent to

$$(\mathcal{P}') \;:\; \operatorname*{argmin}_{u \geq \varphi} F(u).$$

  I.e., $\mathcal{P}$ is a constrained problem: minimize F under the inequality constraint $u \geq \varphi$.

- $F(u)$ is convex $\implies$ all local minima of $\mathcal{P}'$ are global minima.
- $F(u)$ is strictly convex $\implies$ $\mathcal{P}'$ has exactly one unique global minima.
- $F(u)$ is nonconvex $\implies$ $\mathcal{P}'$ may have more than one global minima.

Table of Contents

1. A minimization problem in function space
2. Elastic Obstacle Problem
3. Solving EOP

Elastic surface potential energy

- A special class of problem in $\mathcal{P}$ is to minimize the elastic potential energy of a surface.
- Physics assumption: the elastic potential energy of the deformed membrane is proportional to the increase in the area of its surface.
- Consider 2-dimensional case, recall vector calculus 101, the surface area integral is

$$\int_\Omega \sqrt{1 + \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right)} dxdy = \int_\Omega \sqrt{1 + \|\nabla u\|^2_{L^2(\Omega)}} dxdy, \tag{$\dagger$}$$

  where $\|\cdot\|_{L^2(\Omega)}$ is the L^2-norm for function.
- Now we are solving

$$\operatorname*{argmin}_{u \geq \varphi} \int_\Omega c\sqrt{1 + \|\nabla u\|^2_{L^2(\Omega)}}\, \mathrm{d}\boldsymbol{x}.$$

  where c is a coefficient relating surface growth and potential energy. For us we ignore c by assuming $c = 1$.

1st-order Taylor approximation of surface area

- Recall that, near $z = 0$,

$$\begin{aligned} \sqrt{1+z^2} &= 1 + \tfrac{z^2}{2} - \tfrac{z^4}{8} + \tfrac{z^6}{16} - \tfrac{5z^8}{128} + o(z^9) \\ &= 1 + \tfrac{z^2}{2} + o(z^4) \end{aligned}$$

- For small deformations, we neglect higher order derivatives, hence

$$\sqrt{1+z^2} \approx 1 + \frac{1}{2} z^2, \quad \text{if } z \approx 0. \qquad (!)$$

$$\begin{aligned} \int_\Omega \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} dxdy &\overset{(!)}{\approx} \int_\Omega \left(1 + \frac{1}{2}\left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right)\right) dxdy \\ &= \int_\Omega dxdy + \int_\Omega \frac{1}{2}\left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right) dxdy \\ &= \text{constant} + \int_\Omega \frac{1}{2} \|\nabla u\|^2_{L^2(\Omega)} dxdy \end{aligned}$$

Elastic Obstacle Problem (EOP)

- Ignoring the constant, and consider zero boundary condition, we have

$$\underset{u \geq \varphi}{\operatorname{argmin}} \int_{\Omega} \frac{1}{2} \|\nabla u\|_{L^2(\Omega)}^2 \, \mathrm{d}\boldsymbol{x} \ \text{ s.t. } \ u = 0 \text{ on } \partial\Omega, \tag{EOP}$$

  this problem is called *Elastic Obstacle Problem* (EOP), where φ is called the *obstacle*.

- Note that here $L(\boldsymbol{x}, u, \nabla u) = \frac{1}{2}\|\nabla u\|_{L^2(\Omega)}^2$
  - L does not explicitly depend on $(\boldsymbol{x}, u)$ but ∇u
  - L implicitly depends on $(\boldsymbol{x}, u)$
  - L is only the Taylor approximation of $\sqrt{1 + \|\nabla u\|_{L^2(\Omega)}^2}$ near $\nabla u = 0$.

Understanding EOP

$$\underset{u \geq \varphi}{\operatorname{argmin}} \int_\Omega \frac{1}{2} \|\nabla u\|_{L^2(\Omega)}^2 \, \mathrm{d}\boldsymbol{x} \ \text{ s.t. } \ u = 0 \text{ on } \partial\Omega, \qquad \text{(EOP)}$$

- EOP can be interpreted as follows: given an obstacle φ within a domain Ω, find a elastic membrane that
  1. minimizes the approximate elastic potential energy $\frac{1}{2}\|\nabla u\|_{L^2(\Omega)}^2$
     - the elastic potential energy is approximately the total amount of variation
     - ∇u is the amount of variation
     - $\|\nabla u\|_{L^2(\Omega)}$ is the total amount of variation
     - $\|\nabla u\|_{L^2(\Omega)}^2$ is the squared total amount of variation
  2. subject to the non-penetration constraint $u \geq \varphi$
     - u has to be on top of φ
  3. Zero boundary condition
- In other words, find the equilibrium state a membrane wrapping the obstacle φ.

Applications of EOP

$$\underset{u \geq \varphi}{\operatorname{argmin}} \int_\Omega \frac{1}{2} \|\nabla u\|^2_{L^2(\Omega)} \, \mathrm{d}\boldsymbol{x} \ \text{ s.t. } \ u = 0 \text{ on } \partial\Omega, \tag{EOP}$$

- PDEs
  - Membrane application
  - Material science
  - Biology
  - Two-phase fluid structure
- Control: optimal stopping.
  - Probability.
  - Financial mathematics: American options pricing.

Other things about EOP

I am not a PDE researcher, the following topics will not be the focus here

- The theoretical characterization of EOP, such as
  - Existence of a solution
  - Uniqueness of a solution
  - The optimality condition at the solution in function space.
  - Regularity property of solutions

Table of Contents

1. A minimization problem in function space
2. Elastic Obstacle Problem
3. Solving EOP

Discretizing the problem

$$\operatorname*{argmin}_{u \geq \varphi} \int_{\Omega} \frac{1}{2} \|\nabla u\|^2_{L^2(\Omega)} \, \mathrm{d}\boldsymbol{x} \ \text{ s.t. } \ u = 0 \text{ on } \partial\Omega, \qquad \text{(EOP)}$$

- The Lagrangian-Euler equation of EOP

$$-\nabla^2 u = 0, \ u \geq \varphi, \ u = 0 \text{ on } \partial\Omega, \qquad \text{(EOP-opt)}$$

  (recall that divergence is the negative transpose of grad so we have a negative sign in front of $\nabla^2 u$)

- (EOP-opt) can be approximate by the following constrained quadratic program

$$\operatorname*{argmin}_{\boldsymbol{u} \geq \boldsymbol{\phi}} \frac{1}{2} \langle \boldsymbol{Q}\boldsymbol{u}, \boldsymbol{u} \rangle,$$

  where $\boldsymbol{Q}$ is a Laplacian matrix. Details here.

Shifting / translation

$$\underset{\boldsymbol{u}}{\operatorname{argmin}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{u}, \boldsymbol{u}\rangle \quad \text{s.t.} \quad \boldsymbol{u} \geq \boldsymbol{\phi}$$

- Note that $\boldsymbol{u} \geq \boldsymbol{\phi}$ is the same as $\boldsymbol{v} \coloneqq \boldsymbol{u} - \boldsymbol{\phi} \geq \boldsymbol{0}$, so we have

  $$\underset{\boldsymbol{v}}{\operatorname{argmin}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle \quad \text{s.t.} \quad \boldsymbol{v} \geq \boldsymbol{0}$$

  where

  $$\boldsymbol{p} = \frac{\boldsymbol{Q} + \boldsymbol{Q}^\top}{2}\boldsymbol{\phi}.$$

- Shifting $\boldsymbol{u}$ to $\boldsymbol{v}$ simplfy the constraint.

$\boldsymbol{v} \geq \mathbf{0}$ as hard constraint: indicator function

$$\operatorname*{argmin}_{\boldsymbol{v} \geq \mathbf{0}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle$$

- Let i_+ denotes the indicator function of nonnegative orthant

$$i_+(a) = \begin{cases} 0 & a \geq 0 \\ \infty & a < 0 \end{cases}$$

- Using i_+ element-wise, the problem in unconstrained form is now

$$\operatorname*{argmin}_{\boldsymbol{v} \geq \mathbf{0}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + i_+(\boldsymbol{v}),$$

i.e. if any $v_i < 0$, then $i_+(\boldsymbol{v}) = +\infty$.

$\boldsymbol{v} \geq \mathbf{0}$ as soft constraint: penalty function

$$\operatorname*{argmin}_{\boldsymbol{v} \geq \mathbf{0}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle$$

- Let $\mu > 0$ be a pre-defined penalty parameter. Introduce the penalty function

$$\sum_i (-v_i)_+ = \|(-\boldsymbol{v})_+\|_1,$$

  where $(a)_+ = \max\{0, a\}$ and $\|\cdot\|_1$ is ℓ_1 norm. Note that $\max$ is nonnegative so we can remove the ℓ_1 norm.

- The penalty form of the problem is then

$$\operatorname*{argmin}_{\boldsymbol{v}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + \mu\|(-\boldsymbol{v})_+\|_1. \qquad (\star)$$

- It can be shown that, for sufficiently large μ, the solution of $(\star)$ is the same as the solution of the original problem.

How to solve EOP: proximal gradient iteration

- Both

$$\operatorname*{argmin}_{\boldsymbol{v}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + i_+(\boldsymbol{v}) \qquad \operatorname*{argmin}_{\boldsymbol{v}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + \mu\|(-\boldsymbol{v})_+\|_1$$

  can be solved by proximal gradient descent.

- **Proximal gradient**: for $\min f(\boldsymbol{x}) + g(\boldsymbol{x})$ where
  - f is convex differentiable
  - g convex and possibly nondifferentiable, lower semi-continuous

  the proximal gradient iteration with a stepsize $\alpha > 0$ is

$$\boldsymbol{x}^+ = \operatorname{prox}_{\alpha g}\Big(\boldsymbol{x} - \alpha \nabla f(\boldsymbol{x})\Big).$$

  See here for more about proximal gradient.

Proximal operator

- Given a constant $\alpha > 0$ and a function $f : \mathbb{R}^n \to \mathbb{R}$ that is lower semi-continuous and convex, the proximal operator associated with αf at a point $\boldsymbol{z}$ is

$$\operatorname{prox}_{\alpha f}(\boldsymbol{z}) := \underset{\boldsymbol{\zeta}}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{\zeta} - \boldsymbol{z}\|_2^2 + \alpha f(\boldsymbol{\zeta}).$$

  - $\operatorname{prox}_{\alpha f}(\boldsymbol{z})$ is itself an optimization problem
  - $\operatorname{prox}_{\alpha f}(\boldsymbol{z}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$: prox is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ and the mapping is possibly non-unique
    - If f is convex, such problem is a minimization of a strongly-convex function, thus such problem has a unique solution. I.e. we have $\operatorname{prox}_{\alpha f}(\boldsymbol{z}) : \mathbb{R}^n \to \mathbb{R}^n$
    - If f is non-convex, such problem may have multiple solutions.

- Why we discuss proximal operator: it generalizes projection and therefore able to handle the indicator function g and also the penalty function $\|(-\boldsymbol{v})_+\|_1$.

Proximal gradient descent

$$\operatorname*{argmin}_{\boldsymbol{v}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + i_+(\boldsymbol{v}) \quad \operatorname*{argmin}_{\boldsymbol{v}} \frac{1}{2}\langle \boldsymbol{Q}\boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{p}, \boldsymbol{v}\rangle + \mu\|(-\boldsymbol{v})_+\|_1$$

- The gradient step for both problem is

$$\boldsymbol{v} - \frac{\boldsymbol{Q}\boldsymbol{v} - \boldsymbol{p}}{L}, \quad L = \|\boldsymbol{Q}\|_2$$

- The proximal operator of $\alpha i_+(\boldsymbol{z})$ at a point $\boldsymbol{z}$

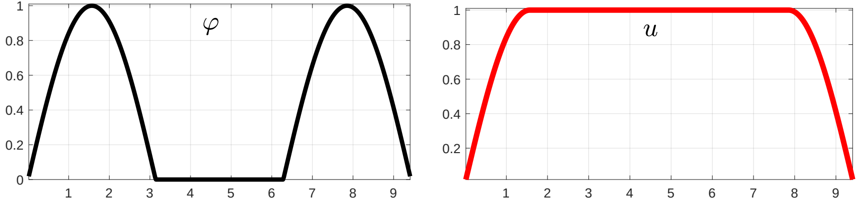
$$\left[\operatorname{prox}_{\alpha i_+(\,\cdot\,)}(\boldsymbol{z})\right]_i = \begin{cases} z_i & z_i > 0 \\ 0 & z_i \leq 0 \end{cases}$$

- The proximal operator of $\alpha\|(-\boldsymbol{z})_+\|_1$ at a point $\boldsymbol{z}$

$$\left[\operatorname{prox}_{\alpha\|(-\,\cdot\,)_+\|_1}(\boldsymbol{z})\right]_i = \begin{cases} z_i + \alpha & 0 > z_i + \alpha \\ 0 & z_i \leq 0 \leq z_i + \alpha \\ z_i & z_i > 0 \end{cases}$$

1-dimensional example

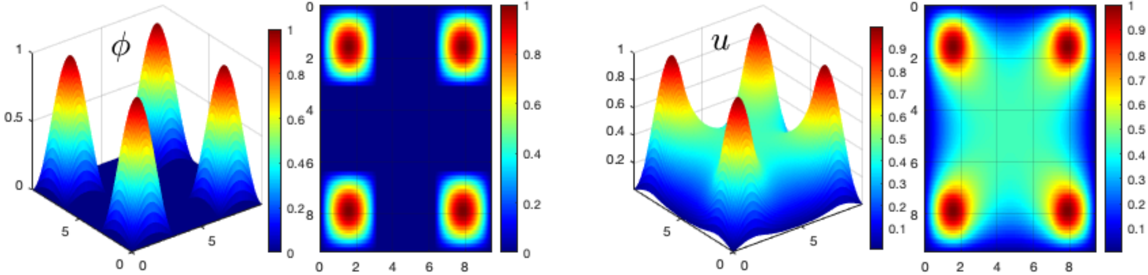
- $x \in \Omega := [0, 3\pi]$
- $\varphi(x) = \max\{0, \sin x\}$
- Solution



- You can prove that such u is the optimal u
  - it is feasible: it sits on top of φ
  - it has the smallest amount of variation

2-dimensional example

- $(x, y) \in \Omega := [0, 3\pi]^2$
- $\varphi(x, y) = \max\{0, \sin x\} \max\{0, \sin y\}$
- Solution



Last page - summary

Discussed

- Derivation of the approximate EOP problem
- Solving the approximate EOP problem in discretized vector space
- Indicator function and penalty function
- Projected gradient and proximal gradient method

Not discussed

- What about solving the original EOP problem
- Theoretical properties of EOP

End of document