Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

SEQUENTIAL QUADRATIC PROGRAMMING FOR OPTIMAL TRANSPORT

Zihe Zhou[1]

[1]School of Industrial Eng., Purdue University, West Lafayette, IN, USA

ABSTRACT

The Monge optimal transport (OT) problem seeks to optimize the transportation cost between two probability measures. The optimization is over a function space and the transportation cost is defined by a cost functional of the maps. Many recent works focus on addressing the OT problem computationally using finite approximations. In this work, we present the infinite-dimensional OT problem over a Banach space. We provide explicit expressions for the first and second-order variation of the objective functional, and of the function form constraint. We propose a Sequential Quadratic Programming (SQP) framework and show that subject to reasonable regularity assumptions, our framework satisfies Alt's SQP condition, immediately yielding local convergence. Moreover, we demonstrate that a merit functional effectively serves as a step-size monitor, leading to global convergence towards a critical point. To the best of our knowledge, this is the first attempt at a globally convergent SQP operator recursion over infinite-dimensional spaces.

1 PROBLEM SETTING

Let $\mathscr{X} \subseteq \mathbb{R}$, $\mathscr{S} = \{s(\cdot) : \mathscr{X} \to \mathscr{X}\}$ be a normed space of transport maps, and $C(\cdot) : \mathbb{R} \to \mathbb{R}$ be a cost function. Denote $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ as two fixed probability density functions. Let $\mathscr{P}(\mathbb{R})$ be the space of probability density functions on $\mathbb{R}$. Consider the Frobenius-Perron (FP) operator $Pf : \mathscr{S} \to \mathscr{P}(\mathbb{R})$,

$$Pf(s) := \left\{ \frac{d}{dx} \int_{s^{-1}((-\infty,x])} f(y)dy, x \in \mathbb{R} \right\}.$$

Consider the optimal transport problem:

$$\begin{aligned} &\min_{s\in\mathscr{S}} J(s) = \int_{\mathscr{X}} C(x - s(x)) f(x) dx \qquad \text{(NLP)} \\ &\text{s.t.} \\ &Pf(s) - g = \theta_{\mathscr{M}}, \end{aligned}$$

where $\theta_{\mathscr{M}}$ is the zero element in $\mathscr{M}$. We can view this objective as minimizing the total transportation cost under the constraint that the transport map s pushes forward the source measure associated with f to exactly match the target measure associated with g". In other words, the transport map s is measure-preserving.

2 NOTATION

Let $\mathscr{X} \subseteq \mathbb{R}$ and $\mathscr{S} = \{s(\cdot) : \mathscr{X} \to \mathscr{X}\}$ be a normed space of transport maps. Denote $\mathscr{F}(\mathbb{R})$ as the space of real functions on $\mathbb{R}$, $\mathscr{B}(\mathscr{S}, \mathbb{R})$ the space of bounded linear functionals with domain $\mathscr{S}$. As introduced in the problem-setting, $J : \mathscr{S} \to \mathbb{R}$ is the objective functional. Its first and second variation operators are $J' : \mathscr{S} \to \mathscr{B}(\mathscr{S}, \mathbb{R}) := \mathscr{S}^*$ and $J'' : \mathscr{S} \to \mathscr{B}(\mathscr{S}, \mathscr{B}(\mathscr{S}, \mathbb{R})) := \mathscr{S}^{**}$ defined through maps $J' : s \in \mathscr{S} \mapsto J_s \in \mathscr{S}^*$ and $J'' : s \in \mathscr{S} \mapsto H_s \in \mathscr{S}^{**}$. Recall the FP operator $Pf : \mathscr{S} \to \mathscr{P}(\mathbb{R})$, its first variation operator is $Pf' : \mathscr{S} \to \mathscr{B}(\mathscr{S}, \mathscr{F}(\mathbb{R}))$ defined through the map $Pf' : s \mapsto Pf_s \in \mathscr{F}(\mathbb{R})$.

3 MAIN CONTRIBUTIONS

3.1 Sequential Quadratic Procedure over Infinite-dimensional Space

Consider a sequential quadratic procedure. Let s_k represent the transport map computed at the kth step of the procedure and let $h \in \mathscr{S}$ be a perturbation direction to the transport map. We approximate $J(s_k + h)$ up to the second-order term and the constraint up to a first-order term. We now propose to solve the following quadratic problem:

$$\begin{aligned} &\min_{h \in \mathscr{S}} J(s_k) + J_{s_k}(h) + \frac{1}{2} H_{s_k}(h)(h) \\ \text{s.t} \quad &Pf(s_k)(x) + Pf_{s_k}(h)(x) - g = 0. \end{aligned} \tag{QP$^{\text{k}}$}$$

Let h_k be a direction that satisfies the above first-order optimality conditions associated with the Lagrangian of (QP$^{\text{k}}$), the update rule is given by:

$$s_{k+1} = s_k + \alpha_k h_k,$$

where α_k is a suitable step size.

3.2 Local Convergence

Under reasonable regularity conditions on the smoothness of the cost function and densities f and g, we can analytically derive the first and second variations of the objective and the function-form constraint. This enables us to verify Alt's conditions for local convergence of SQP over function space (Alt 1990), yielding the following result:

Theorem 1 For $s \in S$, let ϕ defines the Lagrange multiplier for s in the quadratic sub-problem. Let s^* be a critical point of (NLP) and ϕ^* be the function that defines the Lagrange multiplier for s^*. Modulo some regularity conditions, there exists a neighborhood of (s^*, ϕ^*) such that (s_k, ϕ_k) is a unique sequence that converges quadratically to (s^*, ϕ^*) starting from any point in the neighborhood.

3.3 Global Convergence to Critical Points

Consider the merit functional:

$$M(s, \eta) = J(s) + \int_{\mathscr{X}} \phi(x)(Pf(s)(x) - g(x))dx + \frac{\eta}{2} \int_{\mathscr{X}} |Pf(s)(x) - g(x)|dx$$

where η is the penalty parameter of the merit functional. The merit functional serves as a step-size monitor, as the standard line-search methods can be implemented on this unconstrained objective. We can show that global convergence can be achieved with the help of merit functional.

Theorem 2 For a sufficiently large η, starting from any initial location, a critical point s^* of $M(s, \eta)$ can be reached through the SQP procedure. Moreover, s^* is a local min of (NLP) if and only if it is a local min of the merit functional $M(s, \eta)$.

REFERENCES

Alt, W. 1990. "The Lagrange-Newton method for infinite-dimensional optimization problems". *Numerical Functional Analysis and Optimization* 11(3-4):201–224.