On efficient optimal transport: an analysis of greedy and accelerated mirror descent algorithms

Tianyi Lin*, Nhat Ho*, Michael I. Jordan

University of California, Berkeley

June, 2019

Optimal transport (OT)

OT formulation:

$$\min \langle C, X \rangle \qquad X\mathbf{1} = r, \; X^\top \mathbf{1} = l, \; X \geq 0.$$

- $X \in \mathbb{R}_+^{n \times n}$: *transportation plan*
- $C \in \mathbb{R}_+^{n \times n}$: cost matrix comprised of nonnegative elements
- r and l: fixed vectors in the probability simplex Δ^n.

Entropic regularized OT

Entropic regularized OT:

$$\min_{X \in \mathbb{R}^{n \times n}} \langle C, X \rangle - \eta H(X) \quad X\mathbf{1} = r, \; X^\top \mathbf{1} = l.$$

- $\eta > 0$: *regularization parameter*
- $H(X)$: *entropic regularization*, given by

$$H(X) \; := \; -\sum_{i,j=1}^{n} X_{ij} \log(X_{ij}).$$

Main goal

Goal

Find ε-approximation transportation plan $\hat{X} \in \mathbb{R}_+^{n \times n}$ such that:

- $\hat{X}\mathbf{1} = r$ and $\hat{X}^\top \mathbf{1} = l$
- $\langle C, \hat{X} \rangle \ \leq \ \langle C, X^* \rangle + \varepsilon$ where X^*: optimal transportation plan.

Dual entropic regularized OT

Simple form:

$$\max_{u,v\in\mathbb{R}^n} \; -\sum_{i,j=1}^{n} e^{-\frac{C_{ij}}{\eta}+u_i+v_j} + \langle u, r\rangle + \langle v, l\rangle \,.$$

Matrix form:

$$\min_{u,v\in\mathbb{R}^n} \; f(u,v) \;:=\; \mathbf{1}^\top B(u,v)\mathbf{1} - \langle u, r\rangle - \langle v, l\rangle .$$

where $B(u,v) \;:=\; (e^u)\; e^{-\frac{C}{\eta}}\; (e^v)$.

- Popular algorithm for solving regularized OT is Sinkhorn algorithm

Greenkhorn algorithm

- $\rho(a, b) := b - a + a \log\left(\frac{a}{b}\right)$.

Algorithm 1 GREENKHORN$(C, \eta, r, l, \varepsilon')$

Input: $k = 0$ and $u^0 = v^0 = 0$.
while $E^k > \varepsilon'$ **do**
  $r(u^k, v^k) = B(u^k, v^k)\mathbf{1}$.
  $l(u^k, v^k) = B(u^k, v^k)^\top \mathbf{1}$.
  $I = \text{argmax}_{1 \le i \le n}\, \rho\left(r_i, r_i(u^k, v^k)\right)$.
  $J = \text{argmax}_{1 \le j \le n}\, \rho\left(l_j, l_j(u^k, v^k)\right)$.
  if $\rho\left(r_i, r_i(u^k, v^k)\right) > \rho\left(l_j, l_j(u^k, v^k)\right)$ **then**
    $u_I^{k+1} = u_I^k + \log\left(r_I\right) - \log\left(r_I(u^k, v^k)\right)$.
  else
    $v_J^{k+1} = v_J^k + \log\left(l_J\right) - \log\left(l_J(u^k, v^k)\right)$.
  end if
  $k = k + 1$.
end while
Output: $B(u^k, v^k)$.

Greenkhorn algorithm (Cont.)

Algorithm 2 Approximating OT by GREENKHORN

Input: $\eta = \frac{\varepsilon}{4\log(n)}$ and $\varepsilon' = \frac{\varepsilon}{8\|C\|_\infty}$.

Step 1: Let $\tilde{r} \in \Delta_n$ and $\tilde{l} \in \Delta_n$ be defined as

$$\left(\tilde{r}, \tilde{l}\right) = \left(1 - \frac{\varepsilon'}{8}\right)(r, l) + \frac{\varepsilon'}{8n}(\mathbf{1}, \mathbf{1}).$$

Step 2: Compute

$$\tilde{X} = \text{GREENKHORN}\left(C, \eta, \tilde{r}, \tilde{l}, \varepsilon'/2\right).$$

Step 3: Round $\tilde{X}$ to $\hat{X}$ by Algorithm 2 (Altschuler et al., 2017) such that $\hat{X}\mathbf{1} = r$ and $\hat{X}^\top \mathbf{1} = l$.

Output: $\hat{X}$.

Numerical experiments

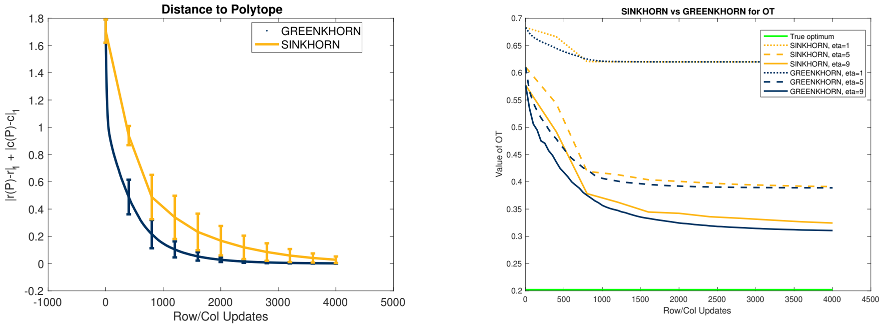

Figure: Comparison of Greenkhorn and Sinkhorn. Left panel: Distance to transportation polytope; Right panel: Different regularization parameter $\eta \in \{1, 5, 9\}$.

- Best known complexity of Sinkhorn is $\mathcal{O}(n^2/\varepsilon^2)$
- Best known complexity of Greenkhorn is $\mathcal{O}(n^2/\varepsilon^3)$ (Altschuler et. al. [2017])

Complexity analysis

- $E^k \;:=\; \|B(u^k, v^k)\mathbf{1} - r\|_1 + \|B(u^k, v^k)^\top \mathbf{1} - l\|_1$.

Theorem 1

The Greenkhorn algorithm returns $B(u^k, v^k)$ satisfying $E_k \leq \varepsilon'$ as long as $k \;\leq\; 2 + \frac{112nR}{\varepsilon'}$ where $R := \; \frac{\|C\|_\infty}{\eta} + \log(n) - 2\log\left(\min_{1\leq i,j\leq n}\{r_i, l_j\}\right)$.

Theorem 2

The Greenkhorn algorithm for approximating OT returns ε-approximation transportation plan $\hat{X} \in \mathbb{R}^{n\times n}$ in

$$\mathcal{O}\left(\frac{n^2\,\|C\|_\infty^2 \log(n)}{\varepsilon^2}\right)$$

arithmetic operations.

Future directions

- Is the complexity bound $\mathcal{O}(n^2/\varepsilon^2)$ of Greenkhorn algorithm tight?
- How to accelerate Sinkhorn and Greenkhorn algorithms for OT?

References

T. Lin*, N. Ho*, M. I. Jordan. On efficient optimal transport: an analysis of greedy and accelerated mirror descent algorithms. *ICML, 2019*.

T. Lin, N. Ho, M. I. Jordan. On the acceleration of the Sinkhorn and Greenkhorn algorithms for optimal transport. *ArXiv preprint arXiv: 1906.01437*.

J. Altschuler, J. Weed, P. Rigollet. Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. *NIPS, 2017*.