Approximation of OT problems with marginal moments contraints

Aurélien Alfonsi[1] Rafaël Coyaud[1] Virginie Ehrlacher[2]
Damiano Lombardi[3]

[1]CERMICS, École des Ponts ParisTech & MATHRISK project, INRIA

[2]CERMICS, École des Ponts ParisTech & MATHERIALS project, INRIA

[3]SERENA project, INRIA

Banff - Optimal Transport Methods in Density Functional Theory - Thursday January the 31st 2019

Sommaire

The Moment Constraint Optimal Transport Problem

In the following presentation $\mathcal{X}$ and $\mathcal{Y}$ are compact subsets of $\mathbb{R}^d$, $d \in \mathbb{N}^*$.

- The Optimal Transport Problem

$$I^* = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y), \tag{1}$$

  where $\Pi(\mu,\nu) = \{\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}) \text{ s.t. } \int_{\mathcal{X}} \mathrm{d}\pi = \mathrm{d}\nu,\ \int_{\mathcal{Y}} \mathrm{d}\pi = \mathrm{d}\mu\}$.

- The Moment Constraint Optimal Transport (MCOT) Problem

$$I^N = \min_{\substack{\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}) \\ \forall 1\leq m\leq N,\ \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\pi(x,y) = \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\mu(x) \\ \forall 1\leq n\leq N,\ \int_{\mathcal{Y}} \psi_n(y)\mathrm{d}\pi(x,y) = \int_{\mathcal{Y}} \psi_n(y)\mathrm{d}\nu(y)}} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y), \tag{2}$$

  where for all $1 \leq m, n \leq N$, ϕ_m, ψ_n are given continuous integrable functions.

Main Results

- One can characterize a minimizer of the MCOT Problem (which is numerically computable).
- With well-chosen sets of test functions $(\phi_m)_{m\in\mathbb{N}}$ and $(\psi_n)_{n\in\mathbb{N}}$ the MCOT problem converges towards the OT problem.
- The convergence speed depends on the choice of test functions.

Sommaire

The Tchakaloff Theorem

Proposition (Tchakaloff [Bayer & Teichmann, 2006], [Berschneider & Sasvári, 2012])

Let π be a positive measure on the space $\mathbb{R}^d$, with the Borel σ-algebra $\mathcal{F}$, concentrated in $A \in \mathcal{F}$, i.e. $\pi(\mathbb{R}^d \setminus A) = 0$, and $\Xi : \mathbb{R}^d \to \mathbb{R}^{N_0}$ a Borel measurable map.
Assume that the first moments of $\Xi\#\pi$ exist, i.e.

$$\int_{\mathbb{R}^N} \|u\| \mathrm{d}\Xi\#\pi(u) < \infty.$$

Then, there exist an integer $1 \leq K \leq N_0$, points $z_1, ..., z_K \in A$ and weights $w_1, ..., w_K > 0$ such that

$$\int_{\Omega} \Xi_i(z) \mathrm{d}\pi(z) = \sum_{k=1}^{K} w_k \Xi_i(z_k)$$

for all $1 \leq i \leq N$, where Ξ_i denotes the i-th component of Ξ.

Characterization of a minimizer of the MCOT Problem

Proposition

For any l.s.c. cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+ \cup \{\infty\}$, *we consider problems of the form*

$$I^N = \min_{\substack{\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}) \\ \forall 1\le m\le N, \int_{\mathcal{X}\times\mathcal{Y}} \phi_m \mathrm{d}\pi = \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\mu(x) \\ \forall 1\le n\le N, \int_{\mathcal{X}\times\mathcal{Y}} \psi_n \mathrm{d}\pi = \int_{\mathcal{Y}} \psi_n(y)\mathrm{d}\nu(y)}} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y). \right\}, \quad (3)$$

With appropriate additional conditions on the test functions $(\phi_m)_{1\le m\le N}$ *and* $(\psi_n)_{1\le n\le N}$, I^N **is finite and is a minimum**.
Moreover, there exists a **finite discrete probability measure** $\gamma = \sum_{k=1}^K w_k \delta_{x_k,y_k}$ *(where for all* k, $x_k \in \mathcal{X}$, $y_k \in \mathcal{Y}$ *and* $w_k \in \mathbb{R}_+^*$ *and* $\sum_{k=1}^K w_k = 1$, *and all points* (x_k, y_k) *are different) with* $\mathbf{0 < K \le 2N + 2}$, *which* **is a minimizer**.

Characterization of a minimizer of the MCOT Problem, remarks

Remark

One can formulate such a Moment Constraint Optimal Transport Problem even in the case where $\mathcal{X}$ and $\mathcal{Y}$ are non compact sets, with some additional technicalities. Thus it **can be applied to DFT**.

Remark

The numerical interest of the characterization of such a minimizer is that, for N given test functions on each set, it is computable by a particle algorithm needing only $2N+2$ points and weights, and in a multimarginal case, with D marginal laws, $DN+2$ points and weights.

Sommaire

Density Condition on test functions

The density condition needed to establish the convergence on the compact sets $\mathcal{X}$ and $\mathcal{Y}$ is for the continuous bounded functions and for the L^∞ norm:

$$\forall f \in C_c^0(\mathcal{X}), \forall \epsilon > 0, \exists M \in \mathbb{N}, \lambda_1, ..., \lambda_M \in \mathbb{R} \mid \sup_{x \in \mathcal{X}} \left| f(x) - \sum_{i=1}^{M} \lambda_i \phi_i(x) \right| \leq \epsilon \tag{4}$$

and

$$\forall f \in C_c^0(\mathcal{Y}), \forall \epsilon > 0, \exists M \in \mathbb{N}, \lambda_1, ..., \lambda_M \in \mathbb{R} \mid \sup_{y \in \mathcal{Y}} \left| f(y) - \sum_{i=1}^{M} \lambda_i \psi_i(y) \right| \leq \epsilon \tag{5}$$

Notations

Recall

- The Optimal Transport Problem

$$I^* = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y). \tag{6}$$

- The Moment Constraint Optimal Transport (MCOT) Problem

$$I^N = \min_{\substack{\pi \in \mathcal{P}(\mathcal{X}\times\mathcal{Y}) \\ \forall 1\leq m\leq N,\ \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\pi(x,y) = \int_{\mathcal{X}} \phi_m(x)\mathrm{d}\mu(x) \\ \forall 1\leq n\leq N,\ \int_{\mathcal{Y}} \psi_n(y)\mathrm{d}\pi(x,y) = \int_{\mathcal{Y}} \psi_n(y)\mathrm{d}\nu(y)}} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y). \tag{7}$$

Convergence of the MCOT Pb towards the OT Pb

Proposition

Let us consider sequences of continuous test functions $(\phi_m)_{m\in\mathbb{N}}$ and $(\psi_n)_{n\in\mathbb{N}}$ defined on $\mathcal{X}$ (resp. $\mathcal{Y}$) and valued on $\mathbb{R}$, and verifying the density conditions (4) and (5).
Then, using the previous notations where c is a l.s.c. cost function valued on $\mathbb{R}_+ \cup \{+\infty\}$, one has that

$$I^N \xrightarrow[N\to\infty]{} I^*$$

and that from every sequence $(\pi^N)_{N\in\mathbb{N}}$ such that for all N, π^N is a minimizer of the MCOT Problem with N moments, one can extract a subsequence $(\pi^{\varphi(N)})_{N\in\mathbb{N}}$ which **converges towards**

$$\pi^* \in \underset{\pi\in\Pi(\mu,\nu)}{\arg\min} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y) \right\}.$$

Convergence of the MCOT Pb towards the OT Pb, remark

Remark

This result can be extended to non-compact sets $\mathcal{X}$ and $\mathcal{Y}$ with some more technical conditions on the test functions, for a DFT application.

Convergence speed for piecewise constant test functions

Let us define the intervals

$$\forall 1 \leq i \leq N-1,\ T_i^N = [\frac{i-1}{N}, \frac{i}{N}) \text{ and } T_N^N = [\frac{N-1}{N}, 1]. \tag{8}$$

On compact sets in dimension 1, analogous MCOT problems[1] with piecewise constant test functions $\phi_i^N = \mathbf{1}_{T_i^N}$ converge towards the OT Problem at a $1/N$ speed.

Proposition

Let $\mu, \nu \in \mathcal{P}([0,1])$ and $c : [0,1]^2 \to \mathbb{R}_+$ a function with Lipschitz constant $K > 0$. Then, for all $N \in \mathbb{N}^*$,

$$I^N \leq I^* \leq I^N + \frac{K}{N}. \tag{9}$$

[1] Here the MCOT Problem is a infimum and not a minimum.

Convergence speed for piecewise affine test functions

On compact sets in dimension 1, MCOT problems with continuous piecewise affine test functions

$$\phi_i(x) = \begin{cases} N\left(x - \frac{i-1}{N}\right) & \text{if} \quad x \in T_{i-1}^N \\ 1 - N\left(x - \frac{i}{N}\right) & \text{if} \quad x \in T_i^N \\ 0 & \text{elsewhere.} \end{cases}$$

converge towards the Wasserstein-1 distance at a $1/N^2$ speed.

Proposition

Consider two marginal laws $\mu \in \mathcal{P}([0,1])$ and $\nu \in \mathcal{P}([0,1])$ with density ρ_μ and ρ_ν and cumulative distribution functions F_μ and F_ν respectively. Then

$$I^N \leq W_1(\mu,\nu) \leq I^N + 2\sup_{[0,1]} |\rho_\mu - \rho_\nu| \frac{M}{N^2}, \tag{10}$$

where M is the number of intervals T_i^N $(1 \leq i \leq N)$ on which $(F_\mu - F_\nu)$ changes of sign.

Sommaire

Convergence speed for piecewise affine test functions in dimension 1

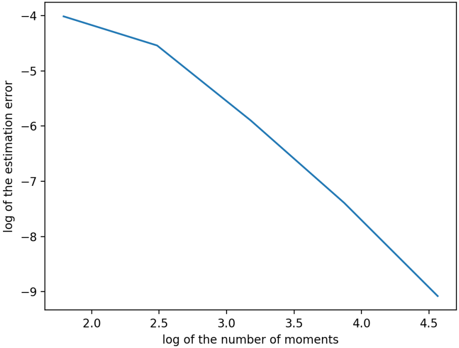
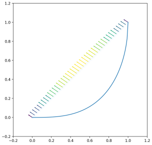
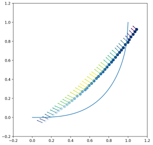


Figure: Convergence speed for piecewise affine test functions in dimension 1 in log-log scale.

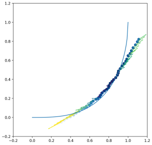
dimension 1



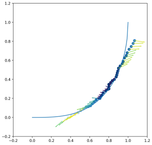
(a) iteration 0



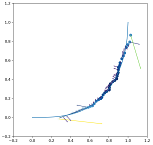
(b) iteration 20



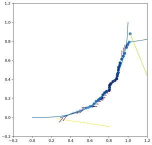
(c) iteration 140



(d) iteration 360

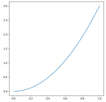


(e) iteration 3000

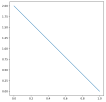


(f) iteration 5000

Figure: Convergence for two 1D marginal laws with 20 test functions on each set



(a) μ



(b) ν

Figure: Marginal laws

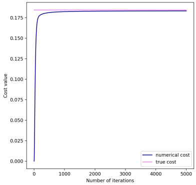
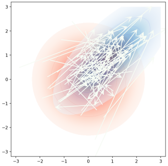
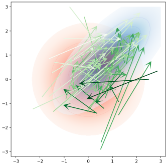


Figure: Cost

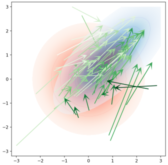
dimension 2



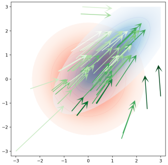
(a) iteration 0



(b) iteration 400



(c) iteration 1400



(d) iteration 9000

Figure: Convergence for two 2D marginal laws with 36 test functions on each set

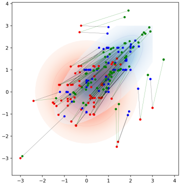


Figure: Transport map

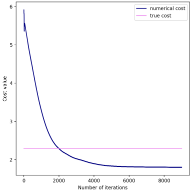


Figure: Cost

Further work

- Explore other possibilities of particle algorithms.
- Study of a symmetric Tchakaloff theorem in order to treat the symmetrical multimarginal case more efficiently.
- Develop an efficient (perhaps multilevel) particle algorithm for dimension 3 in the multimarginal case for a Coulomb cost and in the martingale case.
- Proof of more general rates of convergences.

References

- C. Bayer and Josef Teichmann, *The proof of tchakaloff's theorem*, Proceedings of the American mathematical society, 134(10):3035–3040, 2006.
- J-D Benamou, G. Carlier, M. Cuturi, G. Peyré and L. Nenna *Iterative Bregman Projections for regularizd transportatino problems*, SIAM Journal on Scientific Computing, 37(2):A1111–A1138, 2015.
- G. Berschneider and Z. Sasári *On a theorem of karhunen and related moment problems and quadrature formulae* Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equation, pages 173-187, Springer, 2012.
- P. Billingslay *Probability and measure*, John Wiley & Sons, 2008
- B. Jourdain and J. Reygner, *Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation*, Stochastic Partial Differential Equations: Analysis and Computations, 1(3):455–506, Sep 2013.
- J. Kitagawa, Q. Mérigot and B. Thibert *Convergence of a newton algorithm for semi-discrete optimal transport*, arXiv preprint arXiv:1603.05579, 2016.
- B. Lévy *A numerical algorithm for l2 semi-discrete optimal transport in 3d*, ESAIM: Mathematical Modelling and Numerical Analysis, 49(6):1693–1715, 2015.
- Q. Mérigot *A multiscale approach to optimal transport*, In Computer Graphics Forum, volume 30, pages 1583–1592. Wiley Online Library, 2011.
- B. Pass *Multi-marginal optimal transport: theory and applications*, ESAIM: Mathematical Modelling and Numerical Analysis, 49(6):1771–1790, 2015.
- F. Santambrogio *Optimal transport for applied mathematicians*, Birkäuser, NY, 2015.
- B. Schmitzer *A sparse multiscale algorithm for dense optimal transport*, Journal of Mathematical Imaging and Vision, 56(2):238–259, Oct 2016.
- C. Villani *Optimal transport: old and new*, volume 38, Springer Science & Business Media, 2008.

Thank you for your attention.

Idea of proof of the Tchakaloof Theorem

- For a given measure $\pi \in \mathcal{P}(\mathbb{R}^d)$, $\int_{\mathbb{R}^d} z \mathrm{d}\pi(z)$ lies in $\mathrm{cone}(A)$, where A is such that $\pi(\mathbb{R}^d \setminus A) = 0$.
- The Caratheodory Theorem states that for a set B in dimension N, a point in $\mathrm{cone}(B)$ a positive combination of at most N points of B. Thus, $\exists z_1, .., z_N \in \mathbb{R}^d$, $w_1, ...w_N > 0$,

$$\int_{\mathbb{R}^d} z \mathrm{d}\pi(z) = \sum_{i=1}^{N} w_i z_i. \tag{11}$$

- One can apply the previous result to the measure $\Xi \# \pi$ wich yields to: $\exists z_1, .., z_N \in \mathbb{R}^d$, $w_1, ...w_N > 0$,

$$\int_{\mathbb{R}^d} z \mathrm{d}\Xi \# \pi(z) = \sum_{i=1}^{N} w_i \Xi(z_i). \tag{12}$$