Multi-dimensional Martingale Optimal Transport: Local structure and Numerics

Hadrien De March

CMAP, Ecole Polytechnique

September 6, 2018

CMAP
Résoudre autrement vos conflits
avec le Centre de Médiation et d'Arbitrage de Paris

Outline

Table of Contents

Motivation: the model risk

- Major stake of last century for financial mathematics: pricing exotic derivatives.
- In practice the price is given by $\mathbb{E}^{\mathbb{P}}[\text{payoff}]$, with $\mathbb{P}$ martingale model fitting the prices of vanilla derivatives.
- Problem: two models $\mathbb{P}_1$ and $\mathbb{P}_2$ give two prices!
- Question: what is the maximal possible price?
- Martingale Optimal Transport problem: $\sup_{\mathbb{P}} \mathbb{E}^{\mathbb{P}}[\text{payoff}]$.
- Understand the model risk and find riskless strategies.

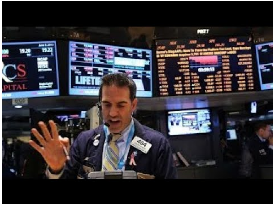


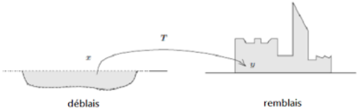
Figure: Jean-Michel, binary options trader.

Table of Contents

The Monge optimal transport problem

Originally a soil moving problem for building from Monge [22].

$$\inf_{\substack{T:\mathcal{X}\longrightarrow\mathcal{Y}\\ T\#\mu=\nu}} \int_{\mathcal{X}} |T(x)-x|\,\mu(dx).$$



(a) The Monge problem illustrated.

(b) The cost of moving bricks.

Figure: Mass moving and cost of moving it.

Probabilistic optimal transport

Kantorovitch [16] made this problem probabilistic.

- $\Omega = \mathbb{R}^d \times \mathbb{R}^d$
- X and Y the two canonical random variables $\Omega \to \mathbb{R}^d$, $X : (x, y) \mapsto x$ and $Y : (x, y) \mapsto y$.
- $\mathcal{P}(\mu, \nu) := \{\mathbb{P} \in \mathcal{P}(\Omega) : \mathbb{P} \circ X^{-1} = \mu, \mathbb{P} \circ Y^{-1} = \nu\}$, set of all coupling probability laws between μ and ν.

Definition

The optimal transport problem is:

$$\mathbf{P} = \inf_{\mathbb{P} \in \mathcal{P}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}[c(X, Y)]$$

Useful cost functions

- $(x, y) \mapsto d(x, y)$ where $d(\cdot, \cdot)$ is a distance.
- $(x, y) \mapsto |x - y|^2$
- $(x, y) \mapsto |x - y|$

Definition

We define the p-Wasserstein distance for $p \geq 1$: for $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$,

$$W^p(\mu, \nu) := \left(\inf_{\mathbb{P} \in \mathcal{P}(\mu, \nu)} \mathbb{E}^{\mathbb{P}}[|X - Y|^p] \right)^{\frac{1}{p}}$$

$(\mathcal{P}(\mathbb{R}^d), W^p)$ is a Polish space.

The dual problem

For $\varphi, \psi : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, let $\varphi \oplus \psi : (x, y) \longmapsto \varphi(x) + \psi(y)$.

Definition

The dual problem is

$$\mathbf{D} := \sup_{\varphi \oplus \psi \leq c} \mu[\varphi] + \nu[\psi]$$

Remark

If $\varphi \oplus \psi \leq c$ *and* $\mathbb{P} \in \mathcal{P}(\mu, \nu)$ *then*

$$\mu[\varphi] + \nu[\psi] = \mathbb{E}^{\mathbb{P}}[\varphi \oplus \psi] \leq \mathbb{E}^{\mathbb{P}}[c].$$

Therefore, $\mathbf{D} \leq \mathbf{P}$.

Kantorovitch Duality

The parallel study of these two problems is justified by the following theorem (see Villani 2008 [25])

Theorem

If the cost c is lower semicontinuous, then

- *Strong duality holds:* $\mathbf{P} = \mathbf{D}$.
- *There are optimizers* φ^*, ψ^* *for* $\mathbf{D}$ *and* $\mathbb{P}^* \in \mathcal{P}(\mu, \nu)$ *for* $\mathbf{P}$.

Proposition

$\mathbb{P} \in \mathcal{P}(\mu, \nu)$ *is concentrated on* $\Gamma := \{\varphi^*(X) + \psi^*(Y) = c(X, Y)\}$ *if and only if it is optimal.*

Structure of Optimal Transport

Theorem

If $y \mapsto \partial_x c(x, y)$ is injective,
then $\mathbb{P}^*[Y = T(X)] = 1$ with $T(x) = \partial_x c(x, \cdot)^{-1}(\nabla\varphi^*(x))$

Idea of the proof:

- $\Delta(x, y) := c(x, y) - \varphi^*(x) - \psi^*(y) \geq 0$,
- $\Delta(x, y) = 0$ if $(x, y) \in \Gamma = \operatorname{supp} \mathbb{P}^*$.
- Then $\partial_x \Delta(x, y) = 0$ if $(x, y) \in \Gamma$.
- $\partial_x c(x, y) = \nabla\varphi^*(x)$.

Table of Contents

Change in the Primal and the Dual Problems

Definition

The martingale optimal transport problem and its dual are:

$$\mathbf{P} = \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}\big[c(X,Y)\big],$$

and

$$\mathbf{D} := \inf_{\varphi\oplus\psi+h^{\otimes}\geq c} \mu[\varphi] + \nu[\psi].$$

With $\mathcal{M}(\mu,\nu) := \{\mathbb{P}\in\mathcal{P}(\mu,\nu) \text{ s.t. } \mathbb{E}^{\mathbb{P}}[Y|X] = X,\ a.s.\}$ and

$$h^{\otimes} := h(X)\cdot(Y-X), \quad \text{for} \quad h:\mathbb{R}^d \longrightarrow \mathbb{R}^d.$$

Binary optimal models for specific cost functions

Beiglböck-Juillet [3], and Henry-Labordère-Touzi [15].

Theorem

We suppose that $d = 1$, $\mu << Leb$ and that $\partial_{xyy} c > 0$. Then

$$Card(\operatorname{supp} \mathbb{P}_X) \leq 2.$$

$$\Gamma = \left\{(x, T_d(x)), (x, T_u(x)) / x \in \mathbb{R}\right\}$$

Numerical solving possible.

Graphical interpretation

- $\partial_x c(x,y) = \nabla h(x) y + \nabla\varphi(x) - h(x) - \nabla h(x) x$, for $(x,y) \in \Gamma$.
- y in the intersection of an affine map with the graph of $\partial_x c(x,\cdot)$.
- Martingale Spence-Mirless condition: $\partial_{xyy} c > 0$, i.e. $\partial_x c$ strictly convex.

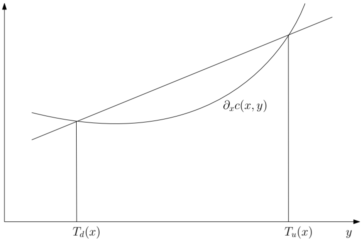


Figure: Structure of Γ_x.

The left-curtain coupling

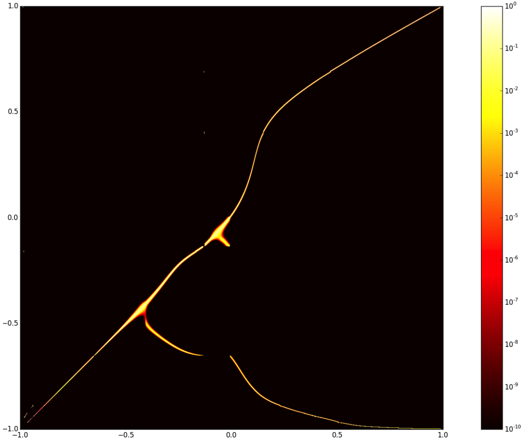


Figure: Left-curtain coupling.

Conjecture on higher dimension

- For $d = 1$, the conditional optimal transports are supported on $2 = d + 1$ points.
- As the optimal plans are extreme in $\mathcal{M}(\mu, \nu)$, we may conjecture that the optimal plans are supported on $d + 1$ points.
- $d + 1$ points convenient for using Partial Differential Equation methods.
- Ghoussoub-Kim-Lim [11] conjecture $d + 1$ points.
- Models for complete markets: $d + 1$ points.

Table of Contents

Structure theorem

DM 2018 [10].

Theorem

(i) *Under Assumption, we may find* $(A_x)_{x\in\mathbb{R}^d} \subset \mathrm{Aff}_d$ *such that for all* $\mathbb{P}^* \in \mathcal{M}(\mu,\nu)$ *optimal for MOT,*

$$\mathrm{supp}\ \mathbb{P}^*_x \subset \{\partial_x c(x,Y) = A_x(Y)\}, \text{ for } \mu - a.e.\ x \in \mathbb{R}^d.$$

(ii) *Conversely, let* $S_0 \subset \{\partial_x c(x_0,Y) = A(Y)\}$, *then under assumptions we may find* $\mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^d)$ *such that*

$$\mathbb{P}^*(dx,dy) := \mu_0(dx)\sum_{i=1}^k \lambda_i(x)\delta_{T_i(x)}(dy)$$

is the unique solution to MOT, and $S_0 = \{T_i(x_0) : 1 \le i \le k\}$.

Link with algebraic geometry

- Taylor: $\partial_x c(x_0, y) \approx \sum_{|j|\leq k} \frac{1}{|i|!}\partial_{x,y^j} c(x_0, y_0)(y - y_0)^j$.
- $\partial_x c(x_0, y) = A(y)$ is locally $P_1(y) = 0, ..., P_d(y) = 0$, with $P_i := \sum_{|j|\leq k} \frac{1}{|i|!}\partial_{x_i,y^j} c(x_0, y_0)(Y - y_0)^j - A_i(Y)$
- Algebraic geometry problem.
- No a priori information about A: $k \geq 2$, at infinity $P_i \approx \sum_{|j|=k} \frac{1}{|i|!}\partial_{x_i,y^j} c(x_0, y_0)(Y - y_0)^j$.
- Example: $\alpha Y^2 - A(Y)$ has 2 roots for all A if and only if $\alpha Y^2 \neq 0$.

Family of polynomial complete at infinity

Definition

We say that the family $(P_1, ..., P_d)$ is $\infty-$complete if

$$QP_i^{hom} \notin \langle P_1^{hom}, ..., P_{i-1}^{hom} \rangle, \text{ for all } Q \notin \langle P_1^{hom}, ..., P_{i-1}^{hom} \rangle.$$

where $< P_1^{hom}, ..., P_{i-1}^{hom} >$ is the ideal generated by the homogeneous part of $P_1, ..., P_{i-1}$.

Remark

Polynomial equation system $T \in \mathbb{R}\left[(X_{i,j})_{1\le i\le d, j\in(\mathbb{N}^*)^d:|j|\le k_i}\right]$ *such that for* $P_i = \sum_{j\in(\mathbb{N}^*)^d:|j|\le k_i} a_{i,j} X_1^{j_1}...X_d^{j_d}$, *we have*

$$T\left((a_{i,j})_{1\le i\le d, |j|\le k_i}\right) \neq 0 \iff (P_1, ..., P_d) \textit{ is } \infty-\textit{complete.}$$

Criterion of discreteness

Theorem (Bezout)

Let $d \in \mathbb{N}$ and $P_1, ..., P_d \in \mathbb{R}[X_1, ..., X_d]$ be $\infty-$complete. Then $|Z^{proj}(P_1, ..., P_d)| = deg(P_1)...deg(P_d)$, counted with multiplicity.

Theorem

If $\left(\sum_{\alpha\in\mathbb{N}^d,\,|\alpha|=k_i} \partial_{y^\alpha} c(x_0, y_0) Y^\alpha\right)_{1\leq i\leq d}$ is $\infty-$complete for all y_0, then $S_0 = \{\partial_x c(x_0, Y) = A(Y)\}$ is discrete.

Remark

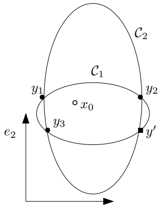
Warning: $c := f(|X - Y|)$ not second order $\infty-$complete.

An illustrative example

Example

Let $c : (x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \longmapsto x_1(y_1^2 + 2y_2^2) + x_2(2y_1^2 + y_2^2)$. Then $\partial_x c(x, y) = (y_1^2 + 2y_2^2)e_1 + (2y_1^2 + y_2^2)e_2$. Let $A \in \mathrm{Aff}(\mathbb{R}^2, \mathbb{R}^2)$, $A = A_1 e_1 + A_2 e_2$. The equation $\partial_x c(x_0, y) = A(y)$ can be written

$$\begin{cases} y_1^2 + 2y_2^2 = & A_1(e_1)y_1 + A_1(e_2)y_2 + A_1(0) \\ 2y_1^2 + y_2^2 = & A_2(e_1)y_1 + A_2(e_2)y_2 + A_2(0). \end{cases}$$



Minimal number of mappings for smooth cost

Theorem

Let $c : \Omega \longrightarrow \mathbb{R}$ be second order complete at infinity and $\mathrm{C}^{2,0} \cap \mathrm{C}^{1,2}$ in the neighborhood of (x_0, x_0) for some $x_0 \in \mathbb{R}^d$. Then, we may find $\mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^d)$ such that the optimal coupling $\mathbb{P}^* \in \mathcal{M}(\mu_0, \nu_0)$ is unique and

$$|\mathrm{supp}\ \mathbb{P}^*_X| = N_c(x_0),\ \mu - a.s.$$

$$N_c(x_0) := \sup_{P \in \mathbb{R}_1[Y_1, \ldots, Y_d]^d} \left| Z^1_{\mathbb{R}}(H_c(x_0) + P) \right|,$$

where $H_c(x_0) := \sum_{i,j} \partial_{x, y_i y_j} c(x_0, x_0) Y_i Y_j$.

Value of $N_c(x_0)$

Theorem

Let $x_0 \in \mathbb{R}^d$, and c second order $\infty-$complete in (x_0, x_0) and $\mathrm{C}^{1,2}$ at (x_0, x_0), then

$$d + 1 + \mathbf{1}_{\{d\ even\}} \leq N_c(x_0) \leq 2^d.$$

Conjecture: $N_c(x_0) = 2^d$ for all second order complete at infinity c.

The case of the distance cost

Theorem

Let $c(X,Y) := |X-Y|^p$. Let $S_0 := \{\partial_x c(x_0,Y) = A(Y)\}$, for some $x_0 \in \operatorname{int}\operatorname{conv} S_0$, and $A \in \operatorname{Aff}(\mathbb{R}^d,\mathbb{R}^d)$.
(i) $p \leq 1$: S_0 contains $2d$ possibly degenerate points counted with multiplicity;
(ii) $p > 1$, S_0 contains $2d+1$ possibly degenerate points counted with multiplicity.

Degenerate points consist in $2k$ points with the same distance to x_0, then we replace these points with a $k-1$ dimensional sphere containing these points.

Limit: disintegration of the Lebesgue measure

- Assumption the the local structure result: μ_I dominated by Lebesgue.
- Invalid argument: the Sudakov mistake and the Nikodym set.
- Further work: prove that this cannot happen.

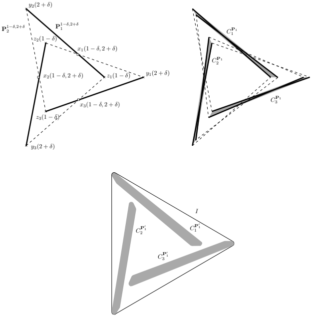


Figure: Three disjoint clusters in one single irreducible component.

Table of Contents

Existing methods in literature

- Linear programming: the Hungarian method [18], the auction algorithm [7], the network simplex [1], [12]...
- Polynomial time: only suited for small sized problems.
- Stochastic control: Benhamou & Brenier [4].
- Monge-Ampère: $\det D^2 u = \frac{g \circ c_x(X,\cdot)^{-1} \circ \nabla u}{f}$, where f is the density of μ and g is the density of ν: see [6] and [5].
- Semi-discrete approach Merigot [21], or Levy [20].
- Very fast but only works for specific (while relevant) costs.
- Entropic approach: Leonard [19], Marco Cuturi [8].

Numerical resolution: the entropic approach

- Adding regularity thanks to an entropic penalization.
- $\mathbf{P}^\varepsilon := \inf_{\mathbb{P}\in\mathcal{P}(\mu,\nu)} \mathbb{P}[c] + \varepsilon H(\mathbb{P})$,
  with $H(\mathbb{P}) := \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(\ln \mathbb{P}(x,y) - 1\big)\mathbb{P}(x,y)$.
- Gibbs shaped solution:
  $\mathbb{P} = \sum_{x,y} \exp\left(-\frac{c(x,y)-\varphi(x)-\psi(y)}{\varepsilon}\right)\delta_{(x,y)}$
- Dual problem:

$$\mathbf{D}^\varepsilon := \sup_{(\varphi,\psi)} \mu[\varphi]+\nu[\psi]-\varepsilon\sum_{x,y}\exp\left(-\frac{c(x,y)-\varphi(x)+\psi(y)}{\varepsilon}\right).$$

The Sinkhorn algorithm

- Smooth concave operator $V_\varepsilon(\varphi, \psi) := \mu[\varphi] + \nu[\psi] - \varepsilon \sum_{x,y} \exp\left(-\frac{c(x,y)-\varphi(x)-\psi(y)}{\varepsilon}\right)$.
- Euler-Lagrange equations $\partial_\varphi V_\varepsilon = 0$ (resp. $\partial_\psi V_\varepsilon = 0$) equivalent to the marginal relations $\mathbb{P} \circ X^{-1} = \mu$ (resp. $\mathbb{P} \circ Y^{-1} = \nu$).
- Marco Cuturi [8] gives the closed formulas: $\varphi(x) = -\varepsilon \ln\left(\frac{1}{\mu_x} \sum_y \exp\left(-\frac{c(x,y)-\psi(y)}{\varepsilon}\right)\right)$, and $\psi(y) = -\varepsilon \ln\left(\frac{1}{\nu_y} \sum_x \exp\left(-\frac{c(x,y)-\varphi(x)}{\varepsilon}\right)\right)$.
- Iterating these partial optimization: Sinkhorn algorithm [23], equivalent to a block optimization of V_ε.
- Converges exponentially fast, see Knight [17].

Existing litterature for numerical solving of MOT

- Henry-Labordère [14]: dual linear programming techniques for specific cost functions so that the dual constraints are much easier to check.
- Alfonsi, Corbetta & Jourdain [2]: difficulty to get a discrete approximation of continuous marginals in convex order, that are still in convex order in higher dimension.
- Guo & Oblój [13]: provide convergence results of the discrete problem to the continuous problem
- [13] provides an equivalent of the Sinkhorn algorithm for one dimensional MOT.
- Tan & Touzi [24]: used a dynamic programming approach to solve a continuous-time version of MOT.

Table of Contents

Entropic framework for MOT

DM 2018 [9].

$$\mathbf{P}^{\varepsilon} := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)}\mathbb{P}[c] - \varepsilon H(\mathbb{P})$$

$$= \mathbf{D}^{\varepsilon} := \inf_{(\varphi,\psi,h)} \mu[\varphi]+\nu[\psi]+\varepsilon\sum_{x,y}\exp\left(-\frac{\varphi\oplus\psi+h^{\otimes}-c}{\varepsilon}\right)(x,y).$$

- We denote $\Delta := \varphi\oplus\psi + h^{\otimes} - c$.
- The optimal coupling probability becomes $\mathbb{P} := \sum_{x,y}\exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right)\delta_{(x,y)}$.
- the convex function to minimize becomes $V_{\varepsilon}(\varphi,\psi,h) := \mu[\varphi]+\nu[\psi]+\varepsilon\sum_{x,y}\exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right)$.

The martingale Sinkhorn algorithm

- Sinkhorn algorithm completed by another step for the martingale relation: $0 = \frac{1}{\mu_x}\sum_y \exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right)(y-x)$ $= \frac{\varepsilon}{\mu_x}\frac{\partial}{\partial h(x)}\sum_y \exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right) = \frac{1}{\mu_x}\frac{\partial}{\partial h(x)}V_\varepsilon$.
- Martingale step not closed form: Newton algorithm for the smooth strongly convex function F_x of d variables given by

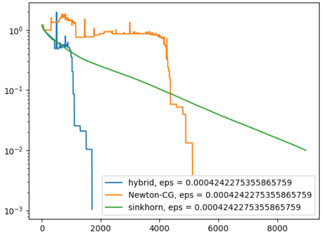
$$\begin{aligned}
F_x(h) &:= \frac{\varepsilon}{\mu_x}\sum_y \exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right),\\
\nabla F_x(h) &= \frac{1}{\mu_x}\sum_y \exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right)(y-x),\\
D^2F_x(h) &= \frac{1}{\mu_x\varepsilon}\sum_y \exp\left(-\frac{\Delta(x,y)}{\varepsilon}\right)(y-x)\otimes(y-x).
\end{aligned}$$

Implied conjugate gradient Newton algorithm

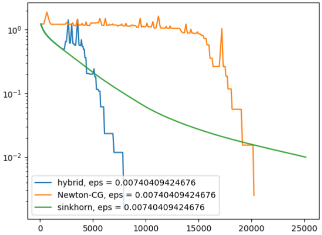
- The optimization in h is costly.
- Newton algorithm: $x_{n+1} - x_n = -D^2V(x_n)^{-1}\nabla V(x_n)$.
- Matrix inverting not adapted to high dimension.
- Conjugate gradient algorithm: find p such that $|D^2V(x_n)p - \nabla V(x_n)| \leq o(|\nabla V(x_n)|)$.
- Exploiting the closed formulas: if V_ε is $\alpha-$convex, so does $\widetilde{V}_\varepsilon(\psi) := \min_{\varphi,h} V_\varepsilon(\varphi,\psi,h)$.
- Optimization on $\widetilde{V}_\varepsilon$ by Newton conjugate gradient: much faster and much more stable.

Comparing the algorithms

- Figure 7a compares these algorithms, for a grid size 2500, with the cost function XY^2.
- Figure 7b compares then on a 160×160 grid, for $c : (x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \longmapsto x_1(y_1^2 + 2y_2^2) + x_2(2y_1^2 + y_2^2)$.
- Newton efficient close to the optimum, and Sinkhorn gets close by closed formulas: we design an hybrid algorithm.

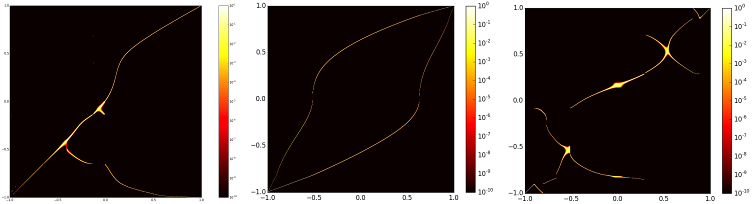


(a) Dimension 1.



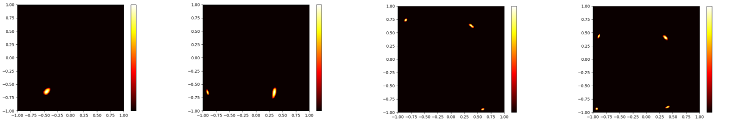
(b) Dimension 2.

The shapes of optimal transport



(a) $c = XY^2$. (b) $c = |X - Y|$. (c) $c = \sin(8XY)$.

Figure: Optimal coupling for different costs in dimension one.



(a) $x = -(0.45, 0.65)$ (b) $x = (0.3, -0.66)$ (c) $x = (0.2, 0.44)$ (d) $x = (0.13, 0.16)$

Default of convex order

- We consider the $\alpha-$penalized problem:

$$\min_{\psi\in\mathbb{R}^{\mathcal{Y}}} \widetilde{V}_\varepsilon(\psi) + \alpha f(\psi). \tag{6.1}$$

- f super-linear, strictly convex, homogeneous: enhances speed and stability.

Theorem

Let $(\mu,\nu)\in\mathcal{P}(\mathcal{X})\times\mathcal{P}(\mathcal{Y})$ not in convex order. Let $\nu_\alpha := \mathbb{P}_\alpha\circ Y^{-1}$, where $\mathbb{P}_\alpha$ is the optimal probability for Problem (6.1). Then $\nu_\alpha \longrightarrow \nu_l$ when $\alpha\longrightarrow 0$, for some $\nu_l \succeq_c \mu$ satisfying

$$f^*(\nu_l-\nu) = \min_{\tilde\nu\succeq_c\mu} f^*(\tilde\nu-\nu).$$

Entropy error

In the litterature the theoretical error linked to entropy is bad: $O(\varepsilon \log(\mathit{grid\ size}))$. Up to making uncheckable assumptions, we get a surprisingly universal and much better result.

Theorem

If

(i) $\Delta_\varepsilon := \varphi_\varepsilon \oplus \psi_\varepsilon + h_\varepsilon^\otimes - c$ *is* C^2;

(iii) $\mathrm{card} Z_x^\varepsilon = \dim Z_x^\varepsilon + 1$, *where* $Z_x^\varepsilon := \{\Delta_\varepsilon(x, Y) \approx 0\}$;

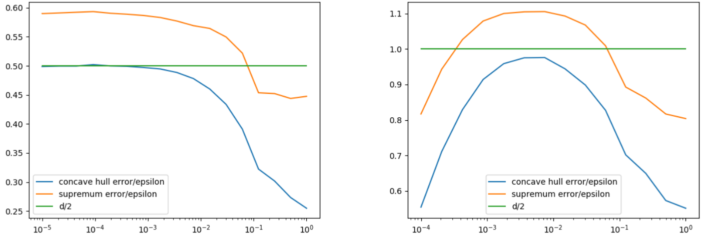
(vi) *the step of the* $Y-$*grid is* $o(\sqrt{\varepsilon})$.

Then $\frac{\mu_\varepsilon\left[(c(X,\cdot)-\psi_\varepsilon)_{conc}(X)\right]+\nu_\varepsilon[\psi_\varepsilon]-\mathbb{P}_\varepsilon[c]}{\varepsilon} \longrightarrow \frac{d}{2}$ *when* $\varepsilon \longrightarrow \infty$.

Corollary

Under the assumptions of Theorem 17, we have that $\mathbb{P}_\varepsilon[c] \geq \mathbf{S}_{\mu,\nu}(c) - \frac{d}{2}\varepsilon + o(\varepsilon)$, *when* $\varepsilon \longrightarrow 0$.

Numerical entropy error



(a) Dimension 1.

(b) Dimension 2.

Figure: Duality gap for the supremum, and the concave hull dual approximation vs ε.

References I

Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin.
Network flows.
1988.

Aurélien Alfonsi, Jacopo Corbetta, and Benjamin Jourdain.
Sampling of probability measures in the convex order and approximation of martingale optimal transport problems.
2017.

Mathias Beiglböck and Nicolas Juillet.
On a problem of optimal transport under marginal martingale constraints.
The Annals of Probability, 44(1):42–106, 2016.

Jean-David Benamou and Yann Brenier.
A computational fluid mechanics solution to the monge-kantorovich mass transfer problem.
Numerische Mathematik, 84(3):375–393, 2000.

Jean-David Benamou, Francis Collino, and Jean-Marie Mirebeau.
Monotone and consistent discretization of the monge-ampere operator.
Mathematics of computation, 85(302):2743–2775, 2016.

Jean-David Benamou, Brittany D Froese, and Adam M Oberman.
Numerical solution of the optimal transportation problem using the monge–ampere equation.
Journal of Computational Physics, 260:107–126, 2014.

Dimitri P Bertsekas.
The auction algorithm: A distributed relaxation method for the assignment problem.
Annals of operations research, 14(1):105–123, 1988.

References II

Marco Cuturi.
Sinkhorn distances: Lightspeed computation of optimal transport.
In Advances in neural information processing systems, pages 2292–2300, 2013.

Hadrien De March.
Entropic resolution for multi-dimensional optimal transport (work in progress).
2018.

Hadrien De March.
Local structure of the optimizer of multi-dimensional martingale optimal transport.
arXiv preprint arXiv:1805.09469, 2018.

Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim.
Structure of optimal martingale transport plans in general dimensions.
arXiv preprint arXiv:1508.01806, 2015.

Andrew V Goldberg and Robert E Tarjan.
Finding minimum-cost circulations by canceling negative cycles.
Journal of the ACM (JACM), 36(4):873–886, 1989.

Gaoyue Guo and Jan Obloj.
Computational methods for martingale optimal transport problems.
arXiv preprint arXiv:1710.07911, 2017.

Pierre Henry-Labordere.
Model-free Hedging: A Martingale Optimal Transport Viewpoint.
CRC Press, 2017.

References III

Pierre Henry-Labordère and Nizar Touzi.
An explicit martingale version of the one-dimensional brenier theorem.
Finance and Stochastics, 20(3):635–668, 2016.

Leonid V Kantorovich.
On a problem of monge.
Journal of Mathematical Sciences, 133(4):1383–1383, 2006.

Philip A Knight.
The sinkhorn–knopp algorithm: convergence and applications.
SIAM Journal on Matrix Analysis and Applications, 30(1):261–275, 2008.

Harold W Kuhn.
The hungarian method for the assignment problem.
Naval Research Logistics (NRL), 2(1-2):83–97, 1955.

Christian Léonard.
From the schrödinger problem to the monge–kantorovich problem.
Journal of Functional Analysis, 262(4):1879–1920, 2012.

Bruno Lévy.
A numerical algorithm for l2 semi-discrete optimal transport in 3d.
ESAIM: Mathematical Modelling and Numerical Analysis, 49(6):1693–1715, 2015.

Quentin Mérigot.
A multiscale approach to optimal transport.
In Computer Graphics Forum, volume 30, pages 1583–1592. Wiley Online Library, 2011.

References IV

Gaspard Monge.
Mémoire sur la théorie des déblais et des remblais.
De l'Imprimerie Royale, 1781.

Richard Sinkhorn and Paul Knopp.
Concerning nonnegative matrices and doubly stochastic matrices.
Pacific Journal of Mathematics, 21(2):343–348, 1967.

Xiaolu Tan, Nizar Touzi, et al.
Optimal transportation under controlled stochastic dynamics.
The annals of probability, 41(5):3201–3240, 2013.

Cédric Villani.
Optimal transport: old and new, volume 338.
Springer Science & Business Media, 2008.

Questions?

Figure: Optimal transport in practice.