

Numerical Computation of Martingale Optimal Transport on Real Line

Gaoyue Guo, University of Oxford

joint work with Jan Oblój

May 1, 2017 @ Oaxaca

OXFORD



Outline

Introduction

Numerical Counterpart

Numerical Computation: Primal Problem

Numerical Computation: Dual Problem (partial results)

OXFORD



Objective

We aim to solve the martingale optimal transport (MOT) problem:

$$P(\mu, \nu) := \sup_{\pi \in \mathcal{M}(\mu, \nu)} \mathbb{E}_{\pi} [c(X, Y)]$$

$$D(\mu, \nu) := \inf_{(\varphi, \psi, h) \in \mathcal{D}} \left\{ \int \varphi d\mu + \int \psi d\nu \right\}$$

Objective

We aim to solve the martingale optimal transport (MOT) problem:

$$P(\mu, \nu) := \sup_{\pi \in \mathcal{M}(\mu, \nu)} \mathbb{E}_{\pi} [c(X, Y)]$$

$$D(\mu, \nu) := \inf_{(\varphi, \psi, h) \in \mathcal{D}} \left\{ \int \varphi d\mu + \int \psi d\nu \right\}$$

- The first scheme considers the approximation of marginal distributions, *i.e.*
 $P(\mu, \nu) \rightsquigarrow P(\mu', \nu')$;

Objective

We aim to solve the martingale optimal transport (MOT) problem:

$$P(\mu, \nu) := \sup_{\pi \in \mathcal{M}(\mu, \nu)} \mathbb{E}_{\pi} [c(X, Y)]$$

$$D(\mu, \nu) := \inf_{(\varphi, \psi, h) \in \mathcal{D}} \left\{ \int \varphi d\mu + \int \psi d\nu \right\}$$

- The first scheme considers the approximation of marginal distributions, *i.e.* $P(\mu, \nu) \rightsquigarrow P(\mu', \nu')$;
- The second one consists of solving $D(\mu, \nu) = \inf_{\psi} J(\psi)$.

Primal problem

Let $X(x, y) := x$ and $Y(x, y) := y$ for all $(x, y) \in \mathbb{R}^2$. For (suitable) probability measures μ and ν , define

$$P(\mu, \nu) := \sup_{\pi \in \mathcal{M}(\mu, \nu)} \mathbb{E}_{\pi} [c(X, Y)],$$

where

$$\mathcal{M}(\mu, \nu) := \left\{ \pi : X \overset{\pi}{\sim} \mu, Y \overset{\pi}{\sim} \nu \text{ and } (X, Y) \text{ is } \pi - \text{martingale} \right\}.$$

Primal problem

Let $X(x, y) := x$ and $Y(x, y) := y$ for all $(x, y) \in \mathbb{R}^2$. For (suitable) probability measures μ and ν , define

$$P(\mu, \nu) := \sup_{\pi \in \mathcal{M}(\mu, \nu)} \mathbb{E}_{\pi} [c(X, Y)],$$

where

$$\mathcal{M}(\mu, \nu) := \left\{ \pi : X \overset{\pi}{\sim} \mu, Y \overset{\pi}{\sim} \nu \text{ and } (X, Y) \text{ is } \pi - \text{martingale} \right\}.$$

$\mathcal{M}(\mu, \nu) \neq \emptyset$ iff $\int |x| d\mu, \int |y| d\nu < +\infty$ and $\int \psi d\mu \leq \int \psi d\nu$ for all convex ψ .
Such a pair (μ, ν) is called PCOC.

Dual problem

Let Λ be the space of Lipschitz functions on \mathbb{R} , and $\mathcal{D} \subset \Lambda \times \Lambda \times \mathbb{L}^0(\mathbb{R})$ be the collection of triplets (φ, ψ, h) s.t.

$$\underbrace{\varphi(x) + \psi(y)}_{\text{static trading}} + \underbrace{h(x)(y-x)}_{\text{dynamic trading}} \geq c(x,y), \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

Define

$$D(\mu, \nu) := \inf_{(\varphi, \psi, h) \in \mathcal{D}} \left[\int \varphi d\mu + \int \psi d\nu \right].$$

Duality

Theorem (Beiglböeck, Henry-Labordère and Penkner)

Let $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ be u.s.c. and dominated from above by some affine function, i.e.

$$\sup_{(x,y) \in \mathbb{R}^2} \frac{c(x,y)}{1 + |x| + |y|} < +\infty.$$

Then

- (i) there exists $\pi^* \in \mathcal{M}(\mu, \nu)$ s.t. $P(\mu, \nu) = \mathbb{E}_{\pi^*} [c]$;
- (ii) the duality $P(\mu, \nu) = D(\mu, \nu)$ holds.

Questions

- Dependency $(\mu, \nu) \mapsto P(\mu, \nu)$. Continuous? Lipschitz?

Questions

- Dependency $(\mu, \nu) \mapsto P(\mu, \nu)$. Continuous? Lipschitz?
- Existence and characterization of the dual optimizer (φ^*, ψ^*, h^*) . Monge-Ampère equation?

Questions

- Dependency $(\mu, \nu) \mapsto P(\mu, \nu)$. Continuous? Lipschitz?
- Existence and characterization of the dual optimizer (φ^*, ψ^*, h^*) . Monge-Ampère equation?
- Numerical computation of $P(\mu, \nu) = D(\mu, \nu)$.

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

How to estimate $|P(\mu, \nu) - P(\mu', \nu')|$?

- Take an arbitrary $\pi \in \mathcal{M}(\mu, \nu)$;

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

How to estimate $|P(\mu, \nu) - P(\mu', \nu')|$?

- Take an arbitrary $\pi \in \mathcal{M}(\mu, \nu)$;
- Consider the optimal transport plans $T : \mu \rightsquigarrow \mu'$ and $S : \nu \rightsquigarrow \nu'$;

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

How to estimate $|P(\mu, \nu) - P(\mu', \nu')|$?

- Take an arbitrary $\pi \in \mathcal{M}(\mu, \nu)$;
- Consider the optimal transport plans $T : \mu \rightsquigarrow \mu'$ and $S : \nu \rightsquigarrow \nu'$;
- Construct π' by means of π, T and S ;

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

How to estimate $|P(\mu, \nu) - P(\mu', \nu')|$?

- Take an arbitrary $\pi \in \mathcal{M}(\mu, \nu)$;
- Consider the optimal transport plans $T : \mu \rightsquigarrow \mu'$ and $S : \nu \rightsquigarrow \nu'$;
- Construct π' by means of π, T and S ;
- $|\mathbb{E}_\pi[c] - \mathbb{E}_{\pi'}[c]|$ is “small” if (μ, ν) is “close” to (μ', ν') ;

A heuristic idea

$P(\mu, \nu)$ reduces to be a linear optimization problem if $\text{supp}(\mu)$ and $\text{supp}(\nu)$ are finite.

How to approximate (μ, ν) by another PCOC (μ', ν') ?

How to estimate $|P(\mu, \nu) - P(\mu', \nu')|$?

- Take an arbitrary $\pi \in \mathcal{M}(\mu, \nu)$;
- Consider the optimal transport plans $T : \mu \rightsquigarrow \mu'$ and $S : \nu \rightsquigarrow \nu'$;
- Construct π' by means of π, T and S ;
- $|\mathbb{E}_\pi[c] - \mathbb{E}_{\pi'}[c]|$ is “small” if (μ, ν) is “close” to (μ', ν') ;
- In general $\pi' \notin \mathcal{M}(\mu', \nu')$.

A relaxed optimization problem

For $\varepsilon \in \mathbb{R}_+$, define

$$\mathcal{M}_\varepsilon(\mu, \nu) := \left\{ \pi : X \stackrel{\pi}{\sim} \mu, Y \stackrel{\pi}{\sim} \nu \text{ and } \sup_{\|h\|_\infty \leq 1} \mathbb{E}_\pi [h(X)(Y - X)] \leq \varepsilon \right\},$$

and consider the corresponding optimization problem

$$P_\varepsilon(\mu, \nu) := \sup_{\pi \in \mathcal{M}_\varepsilon(\mu, \nu)} \mathbb{E}_\pi [c(X, Y)].$$

A relaxed optimization problem

For $\varepsilon \in \mathbb{R}_+$, define

$$\mathcal{M}_\varepsilon(\mu, \nu) := \left\{ \pi : X \overset{\pi}{\sim} \mu, Y \overset{\pi}{\sim} \nu \text{ and } \sup_{\|h\|_\infty \leq 1} \mathbb{E}_\pi [h(X)(Y - X)] \leq \varepsilon \right\},$$

and consider the corresponding optimization problem

$$P_\varepsilon(\mu, \nu) := \sup_{\pi \in \mathcal{M}_\varepsilon(\mu, \nu)} \mathbb{E}_\pi [c(X, Y)].$$

Denote $\mathcal{W}_1^\oplus((\mu, \nu), (\mu', \nu')) := \mathcal{W}_1(\mu, \mu') + \mathcal{W}_1(\nu, \nu')$.

A stability result

Proposition

Let (μ', ν') be another PCOC. If c is Lipschitz, then one has $C > 0$ s.t.

$$P_\varepsilon(\mu, \nu) \leq P_{\varepsilon+d}(\mu', \nu') + Cd, \quad \text{with } d := \mathcal{W}_1^\oplus((\mu, \nu), (\mu', \nu')).$$

A stability result

Proposition

Let (μ', ν') be another PCOC. If c is Lipschitz, then one has $C > 0$ s.t.

$$P_\varepsilon(\mu, \nu) \leq P_{\varepsilon+d}(\mu', \nu') + Cd, \quad \text{with } d := \mathcal{W}_1^\oplus((\mu, \nu), (\mu', \nu')).$$

Corollary

Let $((\mu_n, \nu_n))_{n \geq 1}$ be a sequence of PCOCs converging to (μ, ν) under \mathcal{W}_1^\oplus . Set $d_n := \mathcal{W}_1^\oplus((\mu_n, \nu_n), (\mu, \nu))$, then one has $C > 0$ s.t.

$$P(\mu, \nu) \leq P_{d_n}(\mu_n, \nu_n) + Cd_n \leq P_{2d_n}(\mu, \nu) + 2Cd_n.$$

Stability: continuation...

Proposition

- (i) Let c be u.s.c. Then $\lim_{\varepsilon \rightarrow 0} P_\varepsilon(\mu, \nu) = P(\mu, \nu)$.
- (ii) Let (μ, ν) be boundedly supported, with $a = \inf(\text{supp}(\mu))$ and $b = \sup(\text{supp}(\mu))$. Assume further $c \in C^2(\mathbb{R}^2)$ and

$$\int_{[a,b]} \left(\frac{1}{x-a} + \frac{1}{b-x} \right) d\mu < +\infty.$$

Then one has $C > 0$ s.t. $0 \leq P_\varepsilon(\mu, \nu) - P(\mu, \nu) \leq C\varepsilon$ and

$$|P(\mu, \nu) - P_{d_n}(\mu_n, \nu_n)| \leq Cd_n.$$

An explicit construction

Define

$$\begin{aligned}\mu_n(\{k/n\}) &:= \int_{[(k-1)/n, k/n)} (nx + 1 - k) d\mu + \int_{[k/n, (k+1)/n)} (1 + k - nx) d\mu, \\ \nu_n(\{k/n\}) &:= \int_{[(k-1)/n, k/n)} (nx + 1 - k) d\nu + \int_{[k/n, (k+1)/n)} (1 + k - nx) d\nu.\end{aligned}$$

An explicit construction

Define

$$\begin{aligned}\mu_n(\{k/n\}) &:= \int_{[(k-1)/n, k/n)} (nx + 1 - k) d\mu + \int_{[k/n, (k+1)/n)} (1 + k - nx) d\mu, \\ \nu_n(\{k/n\}) &:= \int_{[(k-1)/n, k/n)} (nx + 1 - k) d\nu + \int_{[k/n, (k+1)/n)} (1 + k - nx) d\nu.\end{aligned}$$

Lemma

- (i) μ_n and ν_n are probability measures.
- (ii) (μ_n, ν_n) are PCOCs and $d_n \leq 2/n$.

A linear optimization problem

Set $\alpha_k := \mu_n(\{k/n\})$ and $\beta_k := \nu_n(\{k/n\})$. Then

A linear optimization problem

Set $\alpha_k := \mu_n(\{k/n\})$ and $\beta_k := \nu_n(\{k/n\})$. Then

$$P_{d_n}(\mu_n, \nu_n) = \sup_{p=(p_{i,j})_{i,j \in \mathbb{Z}}} \sum_{i,j \in \mathbb{Z}} p_{i,j} c(i/n, j/n)$$

$$\text{s.t. } \sum_{i,j \in \mathbb{Z}} p_{i,j} = 1 \text{ and } p_{i,j} \geq 0, \text{ for all } i, j \in \mathbb{Z},$$

$$\sum_{j \in \mathbb{Z}} p_{k,j} = \alpha_k \text{ and } \sum_{i \in \mathbb{Z}} p_{i,k} = \beta_k, \text{ for all } k \in \mathbb{Z},$$

$$\sum_{j \in \mathbb{Z}} p_{k,j} j/n \leq (\geq) \left(\sum_{j \in \mathbb{Z}} p_{k,j} \right) (k/n \pm d_n), \text{ for all } k \in \mathbb{Z}.$$

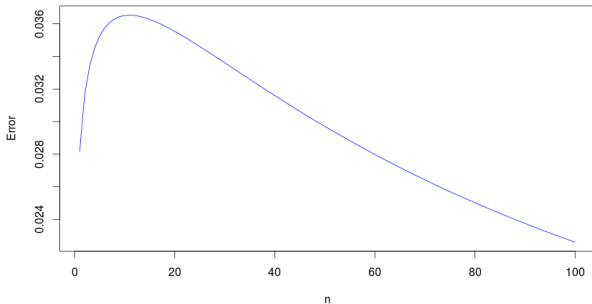
$\mu = \mathcal{U}([-1, 1])$, $\nu = \mathcal{U}([-2, 2])$, $c(x, y) = |x - y|$. Then $P(\mu, \nu) = 1$.

$\mu = \mathcal{U}([-1, 1]), \nu = \mathcal{U}([-2, 2]), c(x, y) = |x - y|$. Then $\mathbf{P}(\mu, \nu) = 1$.

- $\alpha_{-n} = \alpha_n = 1/4n, \alpha_k = 1/2n$, for $-n < k < n$;
- $\beta_{-2n} = \beta_{2n} = 1/8n, \beta_k = 1/4n$, for $-2n < k < 2n$;
- $d_n \leq 2/n$ (set w.l.o.g. $d_n = 2/n$).

$\mu = \mathcal{U}([-1, 1])$, $\nu = \mathcal{U}([-2, 2])$, $c(x, y) = |x - y|$. Then $P(\mu, \nu) = 1$.

- $\alpha_{-n} = \alpha_n = 1/4n$, $\alpha_k = 1/2n$, for $-n < k < n$;
- $\beta_{-2n} = \beta_{2n} = 1/8n$, $\beta_k = 1/4n$, for $-2n < k < 2n$;
- $d_n \leq 2/n$ (set w.l.o.g. $d_n = 2/n$).



Remarks

- Find the sufficient and necessary conditions for $\mathcal{M}_\varepsilon(\mu, \nu) \supset \mathcal{M}(\mu, \nu) \neq \emptyset$;

Remarks

- Find the sufficient and necessary conditions for $\mathcal{M}_\varepsilon(\mu, \nu) \supset \mathcal{M}(\mu, \nu) \neq \emptyset$;
- Generalize the numerical solver for optimal transport;

Remarks

- Find the sufficient and necessary conditions for $\mathcal{M}_\varepsilon(\mu, \nu) \supset \mathcal{M}(\mu, \nu) \neq \emptyset$;
- Generalize the numerical solver for optimal transport;
- Estimate the convergence rate $P_\varepsilon(\mu, \nu) - P(\mu, \nu)$ under more general conditions;

Remarks

- Find the sufficient and necessary conditions for $\mathcal{M}_\varepsilon(\mu, \nu) \supset \mathcal{M}(\mu, \nu) \neq \emptyset$;
- Generalize the numerical solver for optimal transport;
- Estimate the convergence rate $P_\varepsilon(\mu, \nu) - P(\mu, \nu)$ under more general conditions;
- This approach can apply for the multi-dimensional case.

An alternative formulation

Proposition

One has

$$D(\mu, \nu) = \inf_{\psi \in \Lambda} \left\{ \int v_\psi d\mu + \int \psi d\nu \right\}, \quad \text{with } v_\psi(x) := (c(x, \cdot) - \psi)^c(x).$$

Here $(c(x, \cdot) - \psi)^c$ denotes the *concave envelope* of $y \mapsto c(x, y) - \psi(y)$. In addition, see Obermann, $(c(x, \cdot) - \psi)^c$ is the viscosity solution of

$$\max \left(c(x, y) - \psi(y) - u(y), \quad u''(y) \right) = 0.$$

First approximation

Define

$$J(\psi) := \int v_\psi d\mu + \int \psi d\nu,$$

Set $D_L(\mu, \nu) := \inf_{\psi \in \Lambda_L} J(\psi)$, where $\Lambda_L \subset \Lambda$ consists of L -Lipschitz functions ψ with $\psi(0) = 0$. Then

Proposition

$J : \Lambda \rightarrow \mathbb{R}$ is convex and

$$D(\mu, \nu) = \lim_{L \rightarrow +\infty} D_L(\mu, \nu).$$

Further approximation

Let (μ, ν) be supported on a finite interval, e.g. $[0, 1]$. Consider the set $\Lambda_L(n) \subset \Lambda_L$ of functions ψ which are affine on $[(k-1)/n, k/n]$ for $k = 1, \dots, n$.

Further approximation

Let (μ, ν) be supported on a finite interval, e.g. $[0, 1]$. Consider the set $\Lambda_L(n) \subset \Lambda_L$ of functions ψ which are affine on $[(k-1)/n, k/n]$ for $k = 1, \dots, n$.

Remark

Let $U_L(n) \subset \mathbb{R}^n$ be the set of vectors $(u_k)_{1 \leq k \leq n}$ s.t. $|u_k| \leq L$, then there exists a bijection between $U_L(n)$ and $\Lambda_L(n)$. Denote by $\Phi : U_L(n) \rightarrow \Lambda_L(n)$ this bijection.

Further approximation

Let (μ, ν) be supported on a finite interval, e.g. $[0, 1]$. Consider the set $\Lambda_L(n) \subset \Lambda_L$ of functions ψ which are affine on $[(k-1)/n, k/n]$ for $k = 1, \dots, n$.

Remark

Let $U_L(n) \subset \mathbb{R}^n$ be the set of vectors $(u_k)_{1 \leq k \leq n}$ s.t. $|u_k| \leq L$, then there exists a bijection between $U_L(n)$ and $\Lambda_L(n)$. Denote by $\Phi : U_L(n) \rightarrow \Lambda_L(n)$ this bijection.

Define

$$c_n(x, y) := (1 + \lfloor ny \rfloor - ny)c(\lfloor nx \rfloor/n, \lfloor ny \rfloor/n) + (ny - \lfloor ny \rfloor)c(\lfloor nx \rfloor/n, (1 + \lfloor ny \rfloor)/n)$$

Define similarly $J_n(\psi)$ by replacing c by c_n .

OXFORD



Define similarly $J_n(\psi)$ by replacing c by c_n .

Lemma

If c is Lipschitz, then one has $C > 0$ s.t.

$$0 \leq \inf_{\psi \in \Lambda_L(n)} J(\psi) - D_L(\mu, \nu) \leq \frac{C}{n},$$
$$\left| \inf_{\psi \in \Lambda_L(n)} J_n(\psi) - \inf_{\psi \in \Lambda_L(n)} J(\psi) \right| \leq \frac{C}{n}.$$

Define similarly $J_n(\psi)$ by replacing c by c_n .

Lemma

If c is Lipschitz, then one has $C > 0$ s.t.

$$0 \leq \inf_{\psi \in \Lambda_L(n)} J(\psi) - D_L(\mu, \nu) \leq \frac{C}{n},$$
$$\left| \inf_{\psi \in \Lambda_L(n)} J_n(\psi) - \inf_{\psi \in \Lambda_L(n)} J(\psi) \right| \leq \frac{C}{n}.$$

We have

$$\inf_{\psi \in \Lambda_L(n)} J_n(\psi) = \inf_{u \in U_L(n)} \mathcal{J}_n(u), \quad \text{with } \mathcal{J}_n := J_n \circ \Phi.$$

Notice $U_L(n) \subset \mathbb{R}^n$ is convex and compact, and the map \mathcal{J}_n is convex.

More...

For any numerical solver for $\mathcal{J}_n(u)$, e.g. Boyd & Vandenberghe, we may compute $\inf_{u \in U_L(n)} \mathcal{J}_n(u)$ by the following gradient projection algorithm.

More...

For any numerical solver for $\mathcal{J}_n(u)$, e.g. Boyd & Vandenberghe, we may compute $\inf_{u \in U_L(n)} \mathcal{J}_n(u)$ by the following gradient projection algorithm.

Letting $(\gamma_i)_{i \geq 0} \subset \mathbb{R}_+$ be a sequence satisfying $\sum_{i \geq 0} \gamma_i = +\infty$:

1. Let $u^0 := 0$.
2. Given u^i , compute the sub-gradient $\nabla \mathcal{J}_n(u^i)$ of \mathcal{J}_n at u^i .
3. Let $u^{i+1} := \text{Proj}_{U_L(n)}(u^i + \gamma_i \nabla \mathcal{J}_n(u^i))$.
4. Go back to Step 2.

Remarks

- Estimate the convergence rate $D_L(\mu, \nu) - D(\mu, \nu)$;

Remarks

- Estimate the convergence rate $D_L(\mu, \nu) - D(\mu, \nu)$;
- Is there some necessary condition for the optimizer $\operatorname{argmin}_{\psi \in \Lambda} J(\psi)$ (if exists) or $\operatorname{argmin}_{\psi \in \Lambda_L} J(\psi)$?

Thank you very much!

O|X|F|O|R|D

