Martingale Optimal Transport

Kexin SHAO[1,2,3]

Journées Scientifiques des Jeunes Chercheurs du CERMICS (JSJC)

[1]CERMICS
École Nationale des Ponts et Chaussées

[2]CEREMADE
Université Paris Dauphine

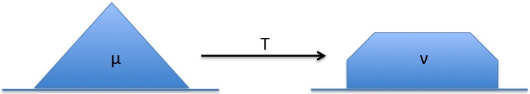
[3]INRIA
Équipe Mathrisk

The Optimal Transport problem

Monge Problem

- Introduced in 1781 by Monge in his work "Théorie des déblais et des remblais", the problem was concerned with transferring mass from one location to another in such a way as to minimize a given cost c.
- After normalization by the total mass, Monge was looking for transport maps $T : \mathbb{R}^d \to \mathbb{R}^d$ measurable such that the image $T\#\mu$ of μ by T is equal to ν :

$$(\text{MP}) := \inf_{T : T\#\mu = \nu} \int_{\mathbb{R}^d} c(x, T(x))\mu(dx).$$



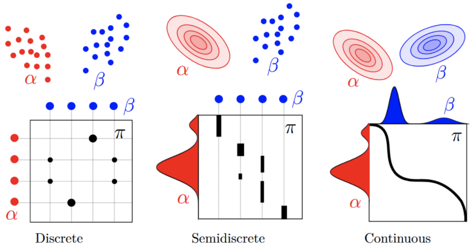
Kantorovich Problem

- After World War II, Kantorovich proposed a relaxation by considering probability couplings (or plans) on the product space
- Given two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ and a measurable mapping $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$:

$$(\mathrm{KP}) := \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\, \pi(dx,dy)$$

where $\Pi(\mu,\nu)$ denotes the set of couplings between μ and ν, that is

$$\Pi(\mu,\nu) = \{\pi \in \mathcal{P}_\rho(\mathbb{R}^d \times \mathbb{R}^d) \mid \pi(dx,\mathbb{R}^d) = \mu(dx) \text{ and } \pi(\mathbb{R}^d,dy) = \nu(dy)\}.$$



[1]

[1] Image by Gabriel Peyré and Marco Cuturi. Computational Optimal Transport.

The Martingale Optimal Transport problem

Let $\Pi^M(\mu,\nu)$ be the set of **martingale coupling** between μ and ν,

$$\Pi^M(\mu,\nu) = \left\{ M(dx,dy) = \mu(dx)m(x,dy) \in \Pi(\mu,\nu) \mid \mu(dx)\text{-a.e.}, \int_{\mathbb{R}^d} y\, m(x,dy) = x \right\},$$

where $(m(x,dy))_{x\in\mathbb{R}}$ is the disintegration of any $M \in \Pi(\mu,\nu)$ with respect to the initial distribution μ, and it is also called transition probability or Markov kernel.

More specifically, for $S_{T_1} \sim \mu$ and $S_{T_2} \sim \nu$, the martingale constraint is equivalent to

$$\mathbb{E}\left[S_{T_2} | S_{T_1}\right] = S_{T_1}$$

The **Martingale Optimal Transport (MOT)** problem states as:

$$\mathrm{MOT}(\mu,\nu,c) = \inf_{\pi\in\Pi^M(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\, M(dx,dy)$$

The existence of MOT

Strassen Theorem, 1965 [4]

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be with finite first moment, then

$$\Pi^M(\mu,\nu) \neq \emptyset \iff \forall f : \mathbb{R}^d \to \mathbb{R} \text{ convex}, \int_{\mathbb{R}^d} f(x)\,\mu(dx) \leq \int_{\mathbb{R}^d} f(y)\,\nu(dy)$$
$$\iff \mu \leq_{cx} \nu$$

- Introduced by Beiglböck, Henry-Labordère and Penkner [1] in a discrete time setting in 2013
- Introduced by Galichon, Henry-Labordère and Touzi [2] in 2014 in a continuous time setting
- Many contributions since : Acciaio, Alfonsi, Backhoff-Veraguas, Bayraktar, Beiglböck, Brückerhoff, Corbetta, Cox, De March, Galichon, Ghoussoub, Guo, Guyon, Henry-Labordère, Hobson, Huesmann, Juillet, Kim, Lim, Neufeld, Nutz, Obłój, Pagès, Pammer, Sester, Siorpaes, Stebegg, Tan, Touzi,...

The numerical representation of MOT

- When μ and ν are finitely supported, the MOT problem becomes a linear programming problem. Although it's still linear, the martingale constraint is non-local which explains why it is not so easy to handle.
- Let us consider the approximation of the probability measures μ and ν by probability measures with finite supports such that

$$\mu_I = \sum_{i=1}^{I} p_i \delta_{x_i} \quad \text{and} \quad \nu_J = \sum_{j=1}^{J} q_j \delta_{y_j},$$

then the approximation of the MOT problem becomes:

$$\min \ \sum_{i=1}^{I} \sum_{j=1}^{J} M_{i,j} c(x_i, y_j)$$

$$s.t. \ M_{i,j} \geq 0 \ \sum_{i=1}^{I} M_{i,j} = q_j, \ \sum_{j=1}^{J} M_{i,j} = p_i, \text{ and } \sum_{j=1}^{J} M_{i,j} y_j = p_i x_i.$$

Definition (Wasserstein distance)

Let $\rho \in [1, \infty)$, for $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ and $\mu \neq \nu$, the **Wasserstein distance** with index ρ is defined as

$$\mathcal{W}_\rho(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\rho \, \pi(dx, dy) \right)^{1/\rho}$$

Notation

For all $\rho \geq 1$, $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ such that $\mu \leq_{cx} \nu$, we define

$$\overline{\mathcal{M}}_\rho(\mu, \nu) = \left(\sup_{M \in \Pi^M(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\rho \, M(dx, dy) \right)^{1/\rho},$$

and

$$\underline{\mathcal{M}}_\rho(\mu, \nu) = \left(\inf_{M \in \Pi^M(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\rho \, M(dx, dy) \right)^{1/\rho}.$$

Financial application

Consider a financial market with

- trading is only allowed in two future time points, $0 < T_1 < T_2$;
- we can dynamically trade the stock whose future values will be S_{T_1} and S_{T_2};
- Let $c(S_{T_1}, S_{T_2})$ be the payoff of the exotic option, where $S_{T_1} \sim \mu$ and $S_{T_2} \sim \nu$, the robust bounds are provided by

$$\mathrm{MOT}(\mu, \nu, c) \leq \mathbb{E}[c(S_{T_1}, S_{T_2})] \leq -\mathrm{MOT}(\mu, \nu, -c).$$

- Let us consider an simple setting such that $c(S_{T_1}, S_{T_2}) = S_{T_2} - S_{T_1}$, then

$$\underline{\mathcal{M}}_1(\mu, \nu) \leq \mathbb{E}[c(S_{T_1}, S_{T_2})] \leq \overline{\mathcal{M}}_1(\mu, \nu).$$

- In practice, μ and ν are extrapolated from the noisy market data, it makes no sense to solve the MOT problem if the stability does not hold

Martingale Wasserstein inequality

When $d = 1$, a martingale Wasserstein inequality was obtained by Jourdain and Margheriti[3]:

Proposition

For all $\rho \geq 1$, there exists $C_\rho < \infty$, such that for all $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfying $\mu \leq_{cx} \nu$ and $\int_{\mathbb{R}} |y|^\rho \nu(dy) < \infty$,

$$\underline{\mathcal{M}}_\rho(\mu, \nu) \leq C_\rho \mathcal{W}_\rho(\mu, \nu) \sigma_\rho^{\rho-1}(\nu).$$

The central moment $\sigma_\rho(\nu)$ of order ρ of $\nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ is defined by

$$\sigma_\rho(\nu) = \inf_{c \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d} |y - c|^\rho \, \nu(dy) \right)^{1/\rho}.$$

Extend to higher dimensions such that $d \in \mathbb{N}^*$,

Proposition

For any $\rho \geq 2$ and $q \geq 1$, there exists some finite constant $C_{\rho,q}$ such that for all $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ with $\mu \leq_{cx} \nu$,

$$\overline{\mathcal{M}}_\rho^\rho(\mu, \nu) \leq C_{\rho,q} \mathcal{W}_q(\mu, \nu) \sigma_{\frac{q(\rho-1)}{q-1}}^{\rho-1}(\nu) \tag{1}$$

For any $1 \leq \rho \leq 2$, $\mu, \nu \in \mathcal{P}_{\frac{q}{q-1}}(\mathbb{R}^d)$

$$\overline{\mathcal{M}}_\rho(\mu, \nu) \leq \overline{\mathcal{M}}_2(\mu, \nu) \leq \sqrt{2\mathcal{W}_q(\mu, \nu) \sigma_{\frac{q}{q-1}}(\nu)} \tag{2}$$

References

M. Beiglböck, P. Henry-Labordère, and F. Penkner.
Model-independent bounds for option prices: A mass transport approach.
Finance and Stochastics, 17(3):477–501, 2013.

A. Galichon, P. Henry-Labordère, and N. Touzi.
A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options.
Annals of Applied Probability, 24(1):312–336, 2014.

Benjamin Jourdain and William Margheriti.
Martingale wasserstein inequality for probability measures in the convex order.
arXiv preprint arXiv:2011.11599, 2020.

V. Strassen.
The existence of probability measures with given marginals.
Annals of Mathematical Statistics, 36(2):423–439, 1965.