Sampling of probability measures in the convex order and approximation of Martingale Optimal Transport problems

Benjamin Jourdain

CERMICS, Ecole des Ponts, University Paris-Est
Joint work with Aurélien Alfonsi and Jacopo Corbetta

MCQMC 2018

5 july 2018

Structure of the talk

The convex order

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d) = \{$probability measures on $\mathbb{R}^d\}$. We say that μ is smaller than ν for the convex order and we write $\mu \leq_{cx} \nu$ if

$$\forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex }, \int_{\mathbb{R}^d} \phi(x)\mu(dx) \leq \int_{\mathbb{R}^d} \phi(y)\nu(dy),$$

when the integrals are defined. For $\phi(x) = \pm x$, we obtain that

$$\int_{\mathbb{R}^d} |y|\nu(dy) < \infty \text{ and } \mu \leq_{cx} \nu \Rightarrow \int_{\mathbb{R}^d} x\mu(dx) = \int_{\mathbb{R}^d} y\nu(dy).$$

Strassen's theorem : (1965) Assume $\int_{\mathbb{R}^d} |y|\nu(dy) < \infty$. $\mu \leq_{cx} \nu$ iff $\exists$ a martingale Markov kernel $R(x, dy)$ ($\forall x \in \mathbb{R}^d, \int_{\mathbb{R}^d} yR(x, dy) = x$) such that $\int \mu(dx)R(x, dy) = \nu(dy)$ i.e. $\mu R = \nu$.

The convex order

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d) = \{$probability measures on $\mathbb{R}^d\}$. We say that μ is smaller than ν for the convex order and we write $\mu \leq_{cx} \nu$ if

$$\forall \phi : \mathbb{R}^d \to \mathbb{R} \text{ convex }, \int_{\mathbb{R}^d} \phi(x)\mu(dx) \leq \int_{\mathbb{R}^d} \phi(y)\nu(dy),$$

when the integrals are defined. For $\phi(x) = \pm x$, we obtain that

$$\int_{\mathbb{R}^d} |y|\nu(dy) < \infty \text{ and } \mu \leq_{cx} \nu \Rightarrow \int_{\mathbb{R}^d} x\mu(dx) = \int_{\mathbb{R}^d} y\nu(dy).$$

Strassen's theorem : (1965) Assume $\int_{\mathbb{R}^d} |y|\nu(dy) < \infty$. $\mu \leq_{cx} \nu$ iff $\exists$ a martingale Markov kernel $R(x, dy)$ ($\forall x \in \mathbb{R}^d$, $\int_{\mathbb{R}^d} yR(x, dy) = x$) such that $\int \mu(dx)R(x, dy) = \nu(dy)$ i.e. $\mu R = \nu$.

Martingale Optimal Transport in Finance

We assume $r = 0$. $(S_t)_{t\geq 0}$: price process of d assets. Suppose that we know for $0 < T_1 < T_2$ the law of S_{T_1} and S_{T_2} (denoted by μ and ν), and that we want to price an option that pays $c(S_{T_1}, S_{T_2})$ at time T_2, with $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$.

Price bounds for the option :

$$\left[\inf_{R\ \text{mart}:\mu R=\nu} \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\mu(dx)R(x,dy), \sup_R \int_{\mathbb{R}^d\times\mathbb{R}^d} c\mu R\right].$$

Multi-marginal case : payoff $c(S_{T_1}, \dots, S_{T_n})$ with $c : (\mathbb{R}^d)^n \to \mathbb{R}$.
Beiglböck, Henry-Labordère, Penkner (2013) : Duality and connection with super/subhedging strategies. Many theoretical contributions since.

Sampling in the Convex order : motivation

When $\mu \leq_{cx} \nu$ are approximated by probability measures $\mu_I = \sum_{i=1}^{I} p_i \delta_{x_i}$ and $\nu_J = \sum_{j=1}^{J} q_j \delta_{y_j}$ with finite supports such that $\mu_I \leq_{cx} \nu_J$, then one can approximate $MOT(\mu, \nu, c)$ by $MOT(\mu_I, \nu_J, c)$ a finite dimensional linear programming problem for which there exist efficient solvers (simplex, interior points,...) :

$$\begin{cases} \text{mini/maximisation of } \sum_{i=1}^{I} \sum_{j=1}^{J} p_i r_{ij} c(x_i, y_j) \text{ under constraints} \\ r_{ij} \geq 0,\ \sum_{i=1}^{I} p_i r_{ij} = q_j,\ \sum_{j=1}^{J} r_{ij} = 1 \text{ et } \sum_{j=1}^{J} r_{ij}(x_i - y_j) = 0 \end{cases}.$$

Monte-Carlo : if $(X_i)_{i\geq 1}$ i.i.d. $\sim \mu$ and $(Y_j)_{j\geq 1}$ i.i.d. $\sim \nu$, in general $\mu_I = \frac{1}{I} \sum_{i=1}^{I} \delta_{X_i}$ is not smaller than $\nu_J = \frac{1}{J} \sum_{j=1}^{J} \delta_{Y_j}$ in the cvx order. In general, $\frac{1}{I} \sum_{i=1}^{I} X_i \neq \frac{1}{J} \sum_{j=1}^{J} Y_j$.

Sampling in the Convex order : motivation

When $\mu \leq_{cx} \nu$ are approximated by probability measures $\mu_I = \sum_{i=1}^{I} p_i \delta_{x_i}$ and $\nu_J = \sum_{j=1}^{J} q_j \delta_{y_j}$ with finite supports such that $\mu_I \leq_{cx} \nu_J$, then one can approximate $MOT(\mu, \nu, c)$ by $MOT(\mu_I, \nu_J, c)$ a finite dimensional linear programming problem for which there exist efficient solvers (simplex, interior points,...) :

$$\begin{cases} \text{mini/maximisation of } \sum_{i=1}^{I} \sum_{j=1}^{J} p_i r_{ij} c(x_i, y_j) \text{ under constraints} \\ r_{ij} \geq 0,\ \sum_{i=1}^{I} p_i r_{ij} = q_j,\ \sum_{j=1}^{J} r_{ij} = 1 \text{ et } \sum_{j=1}^{J} r_{ij}(x_i - y_j) = 0 \end{cases}.$$

Monte-Carlo : if $(X_i)_{i\geq 1}$ i.i.d. $\sim \mu$ and $(Y_j)_{j\geq 1}$ i.i.d. $\sim \nu$, in general $\mu_I = \frac{1}{I}\sum_{i=1}^{I} \delta_{X_i}$ is not smaller than $\nu_J = \frac{1}{J}\sum_{j=1}^{J} \delta_{Y_j}$ in the cvx order. In general, $\frac{1}{I}\sum_{i=1}^{I} X_i \neq \frac{1}{J}\sum_{j=1}^{J} Y_j$.

Approximation techniques preserving the convex order

Dimension $d = 1$ **:** For $\eta \in \mathcal{P}(\mathbb{R})$, we denote $F_\eta(x) = \eta(]-\infty, x])$ and $F_\eta^{-1}(u) = \inf\{x \in \mathbb{R} : F_\eta(x) \geq u\}$ the cumulative distribution function and the quantile function of η.
If $\mu \leq_{cx} \nu$, then (PhD thesis of David Baker UPMC 2012),

$$\frac{1}{I}\sum_{i=1}^{I} \delta_{I\int_{\frac{i-1}{I}}^{\frac{i}{I}} F_\mu^{-1}(u)du} \leq_{cx} \frac{1}{I}\sum_{i=1}^{I} \delta_{I\int_{\frac{i-1}{I}}^{\frac{i}{I}} F_\nu^{-1}(u)du}, \ \forall I \in \mathbb{N}^*.$$

Quantization :

- The dual quantization (Pagès Wilbertz 2012) preserves the convex order when $d = 1$. Whatever d, if ν compactly supported, it gives a probability measure $\hat{\nu}$ with finite support s.t. $\nu \leq_{cx} \hat{\nu}$.
- Stationary primal quantization gives a probability measure $\check{\mu}$ with finite support s.t. $\check{\mu} \leq_{cx} \mu$. So $\check{\mu} \leq_{cx} \mu \leq_{cx} \nu \leq_{cx} \hat{\nu}$.
- Limitations :
  - ν and therefore μ compactly supported,
  - only 2 marginals if $d \geq 2$.

Approximation techniques preserving the convex order

Dimension $d = 1$ **:** For $\eta \in \mathcal{P}(\mathbb{R})$, we denote $F_\eta(x) = \eta(]-\infty, x])$ and $F_\eta^{-1}(u) = \inf\{x \in \mathbb{R} : F_\eta(x) \geq u\}$ the cumulative distribution function and the quantile function of η.
If $\mu \leq_{cx} \nu$, then (PhD thesis of David Baker UPMC 2012),

$$\frac{1}{I}\sum_{i=1}^{I} \delta_{I\int_{\frac{i-1}{I}}^{\frac{i}{I}} F_\mu^{-1}(u)du} \leq_{cx} \frac{1}{I}\sum_{i=1}^{I} \delta_{I\int_{\frac{i-1}{I}}^{\frac{i}{I}} F_\nu^{-1}(u)du}, \ \forall I \in \mathbb{N}^*.$$

Quantization :

- The dual quantization (Pagès Wilbertz 2012) preserves the convex order when $d = 1$. Whatever d, if ν compactly supported, it gives a probability measure $\hat{\nu}$ with finite support s.t. $\nu \leq_{cx} \hat{\nu}$.
- Stationary primal quantization gives a probability measure $\check{\mu}$ with finite support s.t. $\check{\mu} \leq_{cx} \mu$. So $\check{\mu} \leq_{cx} \mu \leq_{cx} \nu \leq_{cx} \hat{\nu}$.
- Limitations :
  - ν and therefore μ compactly supported,
  - only 2 marginals if $d > 2$.

A first idea : equalizing the means

Suppose $\mu \leq_{cx} \nu$, $X_1, \dots, X_I$ i.i.d. $\sim \mu$ and $Y_1, \dots, Y_J$ i.i.d. $\sim \nu$. We set $\bar{X}_I = \frac{1}{I}\sum_{i=1}^{I} X_i$ and $\bar{Y}_J = \frac{1}{J}\sum_{j=1}^{J} Y_j$, and

$$\tilde{\mu}_I = \frac{1}{I}\sum_{i=1}^{I} \delta_{X_i+m-\bar{X}_I}, \; \tilde{\nu}_J = \frac{1}{J}\sum_{j=1}^{J} \delta_{Y_j+m-\bar{Y}_J},$$

with $m = \int x\mu(dx)$ if it is known explicitly (like in finance) or $\bar{X}_I$ otherwise.

- Under conditions slightly stronger than $\mu \leq_{cx} \nu$, a.s., $\exists M, \forall I, J \geq M, \tilde{\mu}_I \leq_{cx} \tilde{\nu}_J$.
- For $\mu = \mathcal{L}(exp(\sigma_\mu G - \frac{\sigma_\mu^2}{2}))$, $\nu = \mathcal{L}(exp(\sigma_\nu G - \frac{\sigma_\nu^2}{2}))$ with $G \sim \mathcal{N}_1(0,1)$, $\sigma_\mu = 0.24$, $\sigma_\nu = 0.28$, $\mathbb{P}(\tilde{\mu}_{100} \leq_{cx} \tilde{\nu}_{100}) \approx 0.45$.
  $\Longrightarrow$ **need for a non asymptotic approach.**

Characterisation of the convex order when $d = 1$

For $\mu \in \mathcal{P}_1(\mathbb{R}) = \{\eta \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |x|\eta(dx) < \infty\}$, we consider the potential function

$$\forall t \in \mathbb{R},\ P_\mu(t) = \int_{\mathbb{R}} (t-x)^+ \mu(dx) = \int_{-\infty}^t F_\mu(x)dx$$

Theorem

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$. One has $\mu \le_{cx} \nu$ iff $\int_{\mathbb{R}} x\mu(dx) = \int_{\mathbb{R}} y\nu(dy)$ and one of the following equivalent conditions hold

(i) $\forall t \in \mathbb{R}, P_\mu(t) \le P_\nu(t)$,

(ii) $\forall q \in [0,1],\ \int_q^1 F_\mu^{-1}(p)dp \le \int_q^1 F_\nu^{-1}(p)dp$.

- $\mathbb{R} \ni t \mapsto P_\mu(t)$ is convex,
- $[0,1] \ni q \mapsto \int_q^1 F_\mu^{-1}(p)dp$ is concave,
- (i) $\Rightarrow$ preservation of the convex order by Baker's approximation.

Infimum and Supremum

According to Kertz et Rösler 1992, 2000, for $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ s.t. $\int_{\mathbb{R}} x\mu(dx) = \int_{\mathbb{R}} y\nu(dy)$, one can define $\mu \vee \nu$ (smallest probability measure larger than μ and ν for the convex order) and $\mu \wedge \nu$ by

$$\forall t \in \mathbb{R}, \int_{-\infty}^{t} F_{\mu\vee\nu}(t)dt = P_{\mu\vee\nu}(t) = P_\mu \vee P_\nu(t)$$

$$\forall t \in \mathbb{R}, \int_{-\infty}^{t} F_{\mu\wedge\nu}(t)dt = P_{\mu\wedge\nu}(t) = \mathrm{Conv}(P_\mu \wedge P_\nu)(t)$$

Infimum and Supremum

According to Kertz et Rösler 1992, 2000, for $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ s.t. $\int_{\mathbb{R}} x\mu(dx) = \int_{\mathbb{R}} y\nu(dy)$, one can define $\mu \vee \nu$ (smallest probability measure larger than μ and ν for the convex order) and $\mu \wedge \nu$ by

$$\forall t \in \mathbb{R},\ \int_{-\infty}^{t} F_{\mu\vee\nu}(t)dt = P_{\mu\vee\nu}(t) = P_\mu \vee P_\nu(t)$$

$$\forall q \in (0,1),\ \int_q^1 F^{-1}_{\mu\wedge\nu}(p)dp = \left(\int_q^1 F^{-1}_\mu(p)dp\right) \wedge \left(\int_q^1 F^{-1}_\nu(p)dp\right).$$

Infimum and Supremum

According to Kertz et Rösler 1992, 2000, for $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ s.t. $\int_{\mathbb{R}} x\mu(dx) = \int_{\mathbb{R}} y\nu(dy)$, one can define $\mu \vee \nu$ (smallest probability measure larger than μ and ν for the convex order) and $\mu \wedge \nu$ by

$$\forall t \in \mathbb{R}, \int_{-\infty}^{t} F_{\mu\vee\nu}(t)dt = P_{\mu\vee\nu}(t) = P_\mu \vee P_\nu(t)$$

$$\forall q \in (0,1), \int_q^1 F^{-1}_{\mu\wedge\nu}(p)dp = \left(\int_q^1 F^{-1}_\mu(p)dp\right) \wedge \left(\int_q^1 F^{-1}_\nu(p)dp\right).$$

Explicit computation when μ and ν have finite supports.

$$\text{In particular for } \tilde{\mu}_I = \frac{1}{I}\sum_{i=1}^{I} \delta_{X_i+m-\bar{X}_I},\ \tilde{\nu}_J = \frac{1}{J}\sum_{j=1}^{J} \delta_{Y_j+m-\bar{Y}_J},$$

computation of $\tilde{\mu}_I \wedge \tilde{\nu}_J$ and $\tilde{\mu}_I \vee \tilde{\nu}_J$ for a cost $\mathcal{O}(I\ln(I) + J\ln(J))$.
When $\mu \leq_{cx} \nu$, a.s. $\tilde{\mu}_I \wedge \tilde{\nu}_J \to \mu$ and $\tilde{\mu}_I \vee \tilde{\nu}_J \to \nu$ weakly as $I, J \to \infty$.

A quadratic minimisation problem

- No nice characterization of the convex order through potential functions,
- According to Müller Scarsini 2006, one cannot define $\mu \vee \nu$ and $\mu \wedge \nu$ for all $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ s.t. $\int_{\mathbb{R}^d} x\mu(dx) = \int_{\mathbb{R}^d} y\nu(dy)$.

For $X_1, \dots, X_I$ i.i.d. $\sim \mu$ and $Y_1, \dots, Y_J$ i.i.d. $\sim \nu$, quadratic minimisation problem with linear constraints

$$\begin{cases} \text{minimise } \frac{1}{I}\sum_{i=1}^I \left| X_i - \sum_{j=1}^J r_{ij} Y_j \right|^2 \\ \text{constraints } \forall i,j,\ r_{ij} \geq 0, \forall i,\ \sum_{j=1}^J r_{ij} = 1 \text{ and } \forall j,\ \frac{1}{I}\sum_{i=1}^I r_{ij} = \frac{1}{J}. \end{cases}$$

- $\exists$ a minimiser $r^\star$, which can be computed by efficient solvers.
- $\frac{1}{I}\sum_{i=1}^I \delta_{\sum_{j=1}^J r^\star_{ij} Y_j}$ does not depend on the minimiser $r^\star$ and $\leq_{cx} \nu_J$

$$\frac{1}{I}\sum_{i=1}^I \phi\Big(\sum_{j=1}^J r_{ij} Y_j\Big) \overset{Jensen}{\leq} \frac{1}{I}\sum_{i=1}^I \sum_{j=1}^J r_{ij}\phi(Y_j) = \frac{1}{J}\sum_{j=1}^J \phi(Y_j).$$

A quadratic minimisation problem

- No nice characterization of the convex order through potential functions,
- According to Müller Scarsini 2006, one cannot define $\mu \vee \nu$ and $\mu \wedge \nu$ for all $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ s.t. $\int_{\mathbb{R}^d} x\mu(dx) = \int_{\mathbb{R}^d} y\nu(dy)$.

For $X_1, \dots, X_I$ i.i.d. $\sim \mu$ and $Y_1, \dots, Y_J$ i.i.d. $\sim \nu$, quadratic minimisation problem with linear constraints

$$\begin{cases} \text{minimise } \frac{1}{I}\sum_{i=1}^{I} \left| X_i - \sum_{j=1}^{J} r_{ij} Y_j \right|^2 \\ \text{constraints } \forall i,j,\ r_{ij} \geq 0, \forall i,\ \sum_{j=1}^{J} r_{ij} = 1 \text{ and } \forall j,\ \frac{1}{I}\sum_{i=1}^{I} r_{ij} = \frac{1}{J}. \end{cases}$$

- $\exists$ a minimiser $r^\star$, which can be computed by efficient solvers.
- $\frac{1}{I}\sum_{i=1}^{I} \delta_{\sum_{j=1}^{J} r^\star_{ij} Y_j}$ does not depend on the minimiser $r^\star$ and $\leq_{cx} \nu_J$

$$\frac{1}{I}\sum_{i=1}^{I} \phi\left(\sum_{j=1}^{J} r_{ij} Y_j\right) \overset{Jensen}{\leq} \frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}\phi(Y_j) = \frac{1}{J}\sum_{j=1}^{J} \phi(Y_j).$$

A quadratic minimisation problem

- No nice characterization of the convex order through potential functions,
- According to Müller Scarsini 2006, one cannot define $\mu \vee \nu$ and $\mu \wedge \nu$ for all $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ s.t. $\int_{\mathbb{R}^d} x\mu(dx) = \int_{\mathbb{R}^d} y\nu(dy)$.

For $X_1, \dots, X_I$ i.i.d. $\sim \mu$ and $Y_1, \dots, Y_J$ i.i.d. $\sim \nu$, quadratic minimisation problem with linear constraints

$$\begin{cases} \text{minimise } \frac{1}{I}\sum_{i=1}^{I} \left| X_i - \sum_{j=1}^{J} r_{ij} Y_j \right|^2 \\ \text{constraints } \forall i,j,\ r_{ij} \geq 0, \forall i,\ \sum_{j=1}^{J} r_{ij} = 1 \text{ and } \forall j,\ \frac{1}{I}\sum_{i=1}^{I} r_{ij} = \frac{1}{J}. \end{cases}$$

- $\exists$ a minimiser $r^\star$, which can be computed by efficient solvers.
- $\frac{1}{I}\sum_{i=1}^{I} \delta_{\sum_{j=1}^{J} r^\star_{ij} Y_j}$ does not depend on the minimiser $r^\star$ and $\leq_{cx} \nu_J$

$$\frac{1}{I}\sum_{i=1}^{I} \phi\left(\sum_{j=1}^{J} r_{ij} Y_j\right) \overset{Jensen}{\leq} \frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}\phi(Y_j) = \frac{1}{J}\sum_{j=1}^{J} \phi(Y_j).$$

Generalisation

For a Markov kernel R on $\mathbb{R}^d$, we set $m_R(x) = \int_{\mathbb{R}^d} yR(x, dy)$. For $\rho \geq 1$ and $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$, we want to minimise

$$\mathcal{J}_\rho(R) := \int_{\mathbb{R}^d} |x - m_R(x)|^\rho \mu(dx) \text{ on } R \text{ kernel s.t. } \mu R = \nu.$$

Wasserstein distance

$$W_\rho^\rho(\mu, \eta) = \inf_{\pi <_\eta^\mu} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\rho \pi(dx, dy) \overset{d=1}{=} \int_0^1 |F_\mu^{-1} - F_\eta^{-1}|^\rho(p)dp.$$

Theorem

$\inf_{R:\mu R=\nu} \mathcal{J}_\rho(R) = \inf_{\eta \leq_{cx} \nu} W_\rho^\rho(\mu, \eta)$ with the infima attained by $R_\star, \eta_\star$. If $\rho > 1$, $m_{R_\star}$ is unique μ a.e.., $\eta_\star = m_{R_\star}\#\mu$, $\pi_\star = \delta_{m_{R_\star}(x)}(dy)\mu(dx)$.

$\mu^\rho_{\underline{\mathcal{P}}(\nu)} := \eta_\star$ Wasserstein projection of μ on the set $\underline{\mathcal{P}}(\nu)$ of probability measures dominated by ν for the convex order.

Generalisation

For a Markov kernel R on $\mathbb{R}^d$, we set $m_R(x) = \int_{\mathbb{R}^d} yR(x,dy)$. For $\rho \geq 1$ and $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$, we want to minimise

$$\mathcal{J}_\rho(R) := \int_{\mathbb{R}^d} |x - m_R(x)|^\rho \mu(dx) \text{ on } R \text{ kernel s.t. } \mu R = \nu.$$

Wasserstein distance

$$W_\rho^\rho(\mu,\eta) = \inf_{\pi <^\mu_\eta} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\rho \pi(dx,dy) \overset{d=1}{=} \int_0^1 |F_\mu^{-1} - F_\eta^{-1}|^\rho(p)dp.$$

Theorem

$\inf_{R:\mu R=\nu} \mathcal{J}_\rho(R) = \inf_{\eta \leq_{cx} \nu} W_\rho^\rho(\mu,\eta)$ with the infima attained by $R_\star, \eta_\star$.
If $\rho > 1$, $m_{R_\star}$ is unique μ a.e.., $\eta_\star = m_{R_\star}\#\mu,\ \pi_\star = \delta_{m_{R_\star}(x)}(dy)\mu(dx)$.

$\mu^\rho_{\underline{\mathcal{P}}(\nu)} := \eta_\star$ Wasserstein projection of μ on the set $\underline{\mathcal{P}}(\nu)$ of probability measures dominated by ν for the convex order.

Projection error

Proposition

Let $\rho \geq 1$, $\mu, \nu, \mu_I, \nu_J \in \mathcal{P}_\rho(\mathbb{R}^d)$ with $\mu \leq_{cx} \nu$. Then

$$W_\rho(\mu_I, (\mu_I)^\rho_{\mathcal{P}(\nu_J)}) \leq W_\rho(\mu, \mu_I) + W_\rho(\nu, \nu_J),$$
$$W_\rho(\mu, (\mu_I)^\rho_{\mathcal{P}(\nu_J)}) \leq 2W_\rho(\mu, \mu_I) + W_\rho(\nu, \nu_J).$$

Corollary

If $\mu \leq_{cx} \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ and $(X_i)_{i\geq 1}$ i.i.d. $\sim \mu$, $(Y_j)_{j\geq 1}$ i.i.d. $\sim \nu$, $\lim_{I,J\to\infty} W_\rho(\mu, (\frac{1}{I}\sum_{i=1}^I \delta_{X_i})^\rho_{\mathcal{P}(\frac{1}{J}\sum_{j=1}^J \delta_{Y_j})}) = 0$

- Rate of cv of $W_\rho(\mu, \frac{1}{I}\sum_{i=1}^I \delta_{X_i})$ as $I \to \infty$: Fournier and Guillin 2015,
- Extension to the multi-marginals case going backwards in time.

Projection error

Proposition

Let $\rho \geq 1$, $\mu, \nu, \mu_I, \nu_J \in \mathcal{P}_\rho(\mathbb{R}^d)$ with $\mu \leq_{cx} \nu$. Then

$$W_\rho(\mu_I, (\mu_I)^\rho_{\mathcal{P}(\nu_J)}) \leq W_\rho(\mu, \mu_I) + W_\rho(\nu, \nu_J),$$
$$W_\rho(\mu, (\mu_I)^\rho_{\mathcal{P}(\nu_J)}) \leq 2W_\rho(\mu, \mu_I) + W_\rho(\nu, \nu_J).$$

Corollary

If $\mu \leq_{cx} \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ and $(X_i)_{i\geq 1}$ i.i.d. $\sim \mu$, $(Y_j)_{j\geq 1}$ i.i.d. $\sim \nu$, $\lim_{I,J\to\infty} W_\rho(\mu, (\frac{1}{I}\sum_{i=1}^I \delta_{X_i})^\rho_{\mathcal{P}(\frac{1}{J}\sum_{j=1}^J \delta_{Y_j})}) = 0$

- Rate of cv of $W_\rho(\mu, \frac{1}{I}\sum_{i=1}^I \delta_{X_i})$ as $I \to \infty$: Fournier and Guillin 2015,
- Extension to the multi-marginals case going backwards in time.

Non dependence on ρ when $d = 1$

Theorem

If $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$, $\exists$ a probability measure $\mu_{\underline{\mathcal{P}}(\nu)}$ defined by : $\forall q \in [0,1]$,

$$\int_0^q F^{-1}_{\mu_{\underline{\mathcal{P}}(\nu)}}(p)dp = \int_0^q F^{-1}_\mu(p)dp - \mathrm{Conv}\left(\int_0^\cdot F^{-1}_\mu(p) - F^{-1}_\nu(p)dp\right)(q).$$

If, for $\rho > 1$, $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R})$, then $\mu^\rho_{\underline{\mathcal{P}}(\nu)} = \mu_{\underline{\mathcal{P}}(\nu)}$.

$\left(\frac{1}{I}\sum_{i=1}^I \delta_{X_i}\right)_{\underline{\mathcal{P}}\left(\frac{1}{J}\sum_{j=1}^J \delta_{Y_j}\right)}$ can be computed with cost $\mathcal{O}(I\ln(I) + J\ln(J))$.

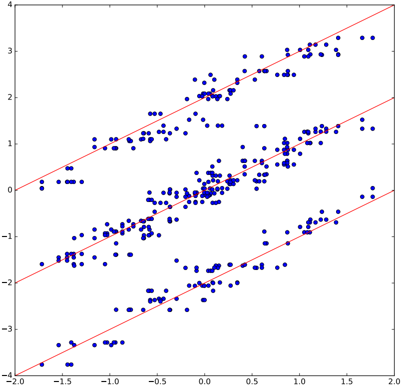
$\nu^\rho_{\overline{\mathcal{P}}(\mu)}$ Wasserstein proj. of ν on the set $\overline{\mathcal{P}}(\mu)$ of probab. meas. larger than μ in the cvx order not easy to compute unless $d = 1$

$$\int_0^q F^{-1}_{\nu_{\overline{\mathcal{P}}(\mu)}}(p)dp = \int_0^q F^{-1}_\nu(p)dp + \mathrm{Conv}\left(\int_0^\cdot F^{-1}_\mu(p) - F^{-1}_\nu(p)dp\right)(q).$$

Example with explicit MOT in dimension $d = 2$

- μ, ν uniform laws on $[-1, 1]^2$ and $[-2, 2]^2$.
- Cost function to minimize : $c(x, y) = |x^1 - y^1|^\rho + |x^2 - y^2|^\rho$, with $\rho > 2$.
- Optimal coupling : (X, Y) where $X \sim \mathcal{U}([-1, 1]^2)$, $Y = X + Z$, with $Z = (Z^1, Z^2)$ an independent couple of independent Rademacher r.v. $\mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = -1) = 1/2$.
  Optimal cost : 2.
- For $I = 100$, we have computed $(\mu_I)^2_{\underline{\mathcal{P}}(\nu_I)}$ and the MOT between $(\mu_I)^2_{\underline{\mathcal{P}}(\nu_I)}$ and ν_I on 100 independent runs $\to$ 95% confidence interval : $[1.9631, 2.0498]$.
- For the optimal coupling, $Y^2 - Y^1 = X^2 - X^1 + Z^2 - Z^1$. Thus, we draw $y_i^2 - y_i^1$ in function of $x_i^2 - x_i^1$ for the points (x_i, y_i) with positive probability in the MOT, and the lines $y = x - 2$, $y = x$ and $y = x + 2$.

Martingale Optimal transport



Financial example in dimension $d = 2$

- (G^1, G^2) centered Gaussian vector with covariance matrix $\Sigma = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$.
- $\mu = \mathcal{L}(X^1, X^2)$ where $X^\ell = \exp(G^\ell - \Sigma_{\ell\ell}/2)$, $\ell \in \{1, 2\}$.
- $\nu = \mathcal{L}(Y^1, Y^2)$ where $Y^\ell = \exp(\sqrt{2}G^\ell - \Sigma_{\ell\ell})$, $\ell \in \{1, 2\}$.
- Payoff : $\max(Y^1 - X^1, Y^2 - X^2, 0)$ (positive part of the best performance). Black-Scholes price ≈ 0.345
- $\tilde{\mu}_I = \frac{1}{I}\sum_{i=1}^I \delta_{(X_i^1+1-\bar{X}_I^1, X_i^2+1-\bar{X}_I^2)}$, $\tilde{\nu}_I = \frac{1}{I}\sum_{i=1}^I \delta_{(Y_i^1+1-\bar{Y}_I^1, Y_i^2+1-\bar{Y}_I^2)}$
- Lower bound (on 100 indep runs of $((\tilde{\mu}_{100})^2_{\underline{\mathcal{P}}(\tilde{\nu}_{100})}, \tilde{\nu}_{100})$: mean 0.2293, 95% confidence interval half-width 0.017,
- Upper bound : mean 0.4111, 95% CI half-width 0.0284

Example with three marginals

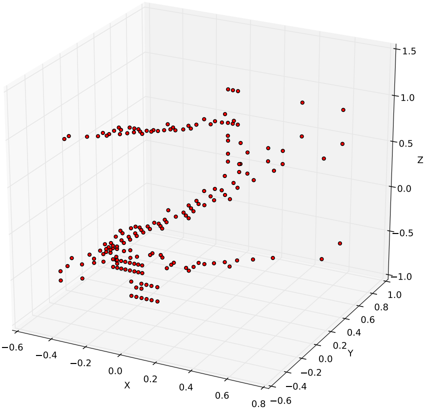
- marginals : $\mu = \mathcal{L}(\exp\left(\sigma_X G - \frac{1}{2}\sigma_X^2\right) - 1)$,
  $\nu = \mathcal{L}(\exp\left(\sigma_Y G - \frac{1}{2}\sigma_Y^2\right) - 1)$ et $\eta = \mathcal{L}(\exp\left(\sigma_Z G - \frac{1}{2}\sigma_Z^2\right) - 1)$,
  with$G \sim \mathcal{N}(0,1)$, $\sigma_X = 0.24$, $\sigma_Y = 0.28$, $\sigma_Z = 0.32$.
- Payoff : $c(x,y,z) = (z - \frac{x+y}{2})^+$, Black-Scholes price ≈ 0.0681.
- lower bound : 0.0303, upper bound 0.0856 obtained with
  $(\hat{\mu}_{25}, \hat{\nu}_{25}, \hat{\eta}_{25})$ (Baker $((\tilde{\mu}_{2500})^2_{\underline{\mathcal{P}}\left((\tilde{\nu}_{2500})^2_{\underline{\mathcal{P}}(\tilde{\eta}_{2500})}\right)}, (\tilde{\nu}_{2500})^2_{\underline{\mathcal{P}}(\tilde{\eta}_{2500})}, \tilde{\eta}_{2500})$).
- Minimisation/maximisation of

$$\sum_{i=1}^{I} p_i \sum_{j=1}^{J} \sum_{k=1}^{K} r_{ijk} c(x_i, y_j, z_k) \text{ under the constraints}$$

$$\forall i,j,k,\ r_{ijk} \geq 0,\ \forall i, \sum_{j=1}^{J}\sum_{k=1}^{K} r_{ijk} = 1,\ \forall j, \sum_{i=1}^{I} p_i \sum_{k=1}^{K} r_{ijk} = q_j,\ \forall k, \sum_{i=1}^{I} p_i \sum_{j=1}^{J} r_{ijk} = s_k,$$

$$\forall i, \sum_{j=1}^{J}\sum_{k=1}^{K} r_{ijk}(y_j - x_i) = 0,\ \forall i,j, \sum_{k=1}^{K} r_{ijk}(z_k - y_j) = 0.$$

MOT for $(\hat{\mu}_{25}, \hat{\nu}_{25}, \hat{\eta}_{25})$ minimisation



Conclusion

- The methods that we have presented, enable to calculate with a MC method at the same time option prices, and their bounds on all other models sharing the same marginal laws.
- The accuracy of the price bounds (maybe not so important in practice) is limited by the dimension of the linear programming problem.
- Possible directions to overcome this limitation : entropic regularization and iterated Bregman projection (Benamou, Carlier, Cuturi, Nenna, 2015 in the OT case), relaxation of the martingale constraints and dual formulation (preprint of Guo and Obloj 2017).