Non-anticipative optimal transport: a powerful tool in stochastic optimization

Beatrice Acciaio
London School of Economics

based on several projects with
J. Backhoff, R. Carmona, A. Zalashko

17th Winter school on Mathematical Finance
Congrescentrum De Werelt, Lunteren
January 22-24, 2018

Outline

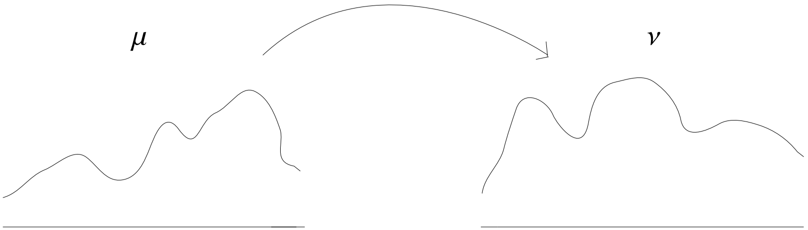
Classical and Causal Optimal Transport

Classical Monge-Kantorovich optimal transport

Given two Polish probability spaces $(\mathcal{X}, \mu), (\mathcal{Y}, \nu)$, move the mass from μ to ν minimizing the cost of transportation $c : \mathcal{X} \times \mathcal{Y} \to [0, \infty]$:

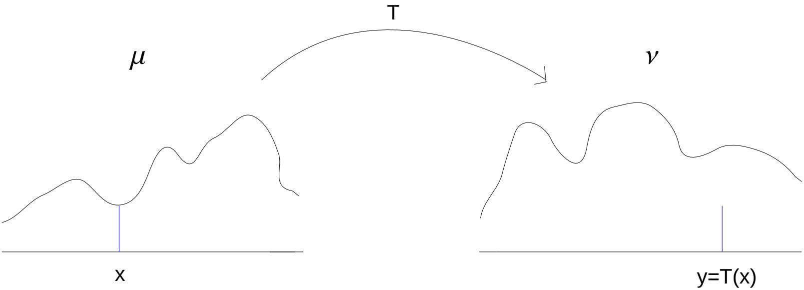
$$\mathrm{OT}(\mu, \nu, c) := \inf \left\{ \mathbb{E}^{\pi}[c(x, y)] : \pi \in \Pi(\mu, \nu) \right\},$$

$\Pi(\mu, \nu)$: probability measures on $\mathcal{X} \times \mathcal{Y}$ with marginals μ and ν.



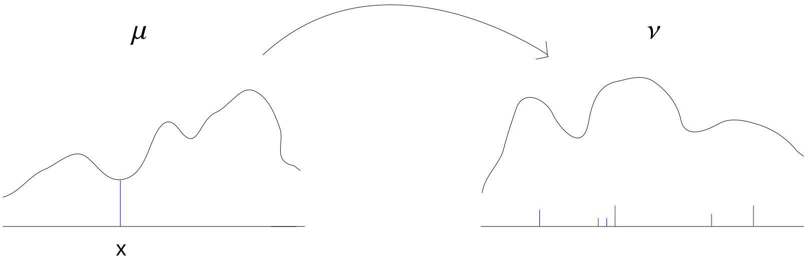
Classical Monge-Kantorovich optimal transport

Monge transport: all mass sitting on x is transported into $y = T(x)$.



Classical Monge-Kantorovich optimal transport

Kantorovich transport: mass can split.



From Monge-Kantorovich to causal optimal transport

Some literature on OT:

- G. Monge (1781)
- L.V. Kantorovich (1942, '48)
- L. Ambrosio, Y. Brenier, L. Caffarelli, A. Figalli, N. Gigli, R. McCann, F. Otto, F. Santabrogio, K.T. Sturm, C. Villani ...

From Monge-Kantorovich to causal optimal transport

Some literature on OT:

- G. Monge (1781)
- L.V. Kantorovich (1942, '48)
- L. Ambrosio, Y. Brenier, L. Caffarelli, A. Figalli, N. Gigli, R. McCann, F. Otto, F. Santabrogio, K.T. Sturm, C. Villani ...

$\rightarrow$ We consider a **dynamic setting**: we have the time component (mathematically: spaces $\mathcal{X}$ and $\mathcal{Y}$ endowed with filtrations)

$\rightarrow$ **Idea**: move the mass in a non-anticipative way: what is transported into the 2^{nd} coordinate at time t, depends on the 1^{st} coordinate only up to t (+ possibly on sth independent)

$\Longrightarrow$ causal (non-anticipatice) transport

Causal optimal transport

Let $\mathcal{F}^{\mathcal{X}} = \left(\mathcal{F}_t^{\mathcal{X}}\right)_t$ on $\mathcal{X}$, $\mathcal{F}^{\mathcal{Y}} = \left(\mathcal{F}_t^{\mathcal{Y}}\right)_t$ on $\mathcal{Y}$ be right-cont. filtrations.

Definition (Causal transport plans $\Pi^{\mathcal{F}^{\mathcal{X}},\mathcal{F}^{\mathcal{Y}}}(\mu,\nu)$)

A transport plan $\pi \in \Pi(\mu,\nu)$ is called causal between $(\mathcal{X},\mathcal{F}^{\mathcal{X}},\mu)$ and $(\mathcal{Y},\mathcal{F}^{\mathcal{Y}},\nu)$ if, for all t and $D \in \mathcal{F}_t^{\mathcal{Y}}$, the map $\mathcal{X} \ni x \mapsto \pi^x(D)$ is measurable w.t.to $\mathcal{F}_t^{\mathcal{X}}$ (π^x regular conditional kernel w.r.to $\mathcal{X}$).

The concept goes back to T. Yamada and S. Watanabe (1971); see also R. Lassalle (2013), J. Backhoff et al. (2016)

Causal optimal transport

Let $\mathcal{F}^{\mathcal{X}} = \left(\mathcal{F}_t^{\mathcal{X}}\right)_t$ on $\mathcal{X}$, $\mathcal{F}^{\mathcal{Y}} = \left(\mathcal{F}_t^{\mathcal{Y}}\right)_t$ on $\mathcal{Y}$ be right-cont. filtrations.

Definition (Causal transport plans $\Pi^{\mathcal{F}^{\mathcal{X}},\mathcal{F}^{\mathcal{Y}}}(\mu,\nu)$)

A transport plan $\pi \in \Pi(\mu,\nu)$ is called causal between $(\mathcal{X},\mathcal{F}^{\mathcal{X}},\mu)$ and $(\mathcal{Y},\mathcal{F}^{\mathcal{Y}},\nu)$ if, for all t and $D \in \mathcal{F}_t^{\mathcal{Y}}$, the map $\mathcal{X} \ni x \mapsto \pi^x(D)$ is measurable w.t.to $\mathcal{F}_t^{\mathcal{X}}$ (π^x regular conditional kernel w.r.to $\mathcal{X}$).

The concept goes back to T. Yamada and S. Watanabe (1971); see also R. Lassalle (2013), J. Backhoff et al. (2016)

Causal optimal transport problem:

$$\mathrm{COT}(\mu,\nu,c) := \inf\left\{\mathbb{E}^{\pi}[c(X,Y)] : \pi \in \Pi_c(\mu,\nu)\right\},$$

where $\Pi_c(\mu,\nu)$ = set of causal transports with marginals μ and ν

Example: weak-solutions of SDEs

- $\mathcal{X} = \mathcal{Y} = C_0[0, \infty)$
- $\mathcal{F}$ right-continuous canonical filtration

Example (Yamada-Watanabe'71)

Assume weak-existence of the solution to the SDE:

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt, \quad b, \sigma \text{ Borel measurable.}$$

Then $\mathcal{L}(B, Y)$ is a causal transport plan between the spaces $(C_0[0, \infty), \mathcal{F}, \gamma)$ and $(C_0[0, \infty), \mathcal{F}, \mathcal{L}(Y))$, γ = Wiener measure.

Example: weak-solutions of SDEs

- $\mathcal{X} = \mathcal{Y} = C_0[0, \infty)$
- $\mathcal{F}$ right-continuous canonical filtration

Example (Yamada-Watanabe'71)

Assume weak-existence of the solution to the SDE:

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt, \quad b, \sigma \text{ Borel measurable.}$$

Then $\mathcal{L}(B, Y)$ is a causal transport plan between the spaces $(C_0[0, \infty), \mathcal{F}, \gamma)$ and $(C_0[0, \infty), \mathcal{F}, \mathcal{L}(Y))$, γ = Wiener measure.

- **Transport perspective:** from an observed trajectory of B, the mass can be split at each moment of time into Y only based on the information available up to that time.

Example: weak-solutions of SDEs

- $\mathcal{X} = \mathcal{Y} = C_0[0,\infty)$
- $\mathcal{F}$ right-continuous canonical filtration

Example (Yamada-Watanabe'71)

Assume weak-existence of the solution to the SDE:

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt, \quad b, \sigma \text{ Borel measurable.}$$

Then $\mathcal{L}(B,Y)$ is a causal transport plan between the spaces $(C_0[0,\infty), \mathcal{F}, \gamma)$ and $(C_0[0,\infty), \mathcal{F}, \mathcal{L}(Y))$, γ = Wiener measure.

- **Transport perspective:** from an observed trajectory of B, the mass can be split at each moment of time into Y only based on the information available up to that time.
- **No splitting of mass:**

$$\text{Monge transport} \iff \text{strong solution } Y = F(B).$$

Semimartingale preservation in enlargement of filtrations

Semimartingale preservation

Problem formulation:

- given two filtrations $\mathcal{F} \subset \mathcal{G}$ on a space of events Ω
- and X semimartingale in $(\Omega, \mathcal{F}, \mathbb{P})$

$\rightarrow$ when is X going to remain a semimartingale in $(\Omega, \mathcal{G}, \mathbb{P})$?

Semimartingale preservation

Problem formulation:

- given two filtrations $\mathcal{F} \subset \mathcal{G}$ on a space of events Ω
- and X semimartingale in $(\Omega, \mathcal{F}, \mathbb{P})$

$\rightarrow$ when is X going to remain a semimartingale in $(\Omega, \mathcal{G}, \mathbb{P})$?

Why is this interesting?

- Semimartingales are the processes for which classical stochastic integration works: $\int H dX$ (e.g. asset price proc.)
- Agents have access to different sets of information

Semimartingale preservation

Problem formulation:

- given two filtrations $\mathcal{F} \subset \mathcal{G}$ on a space of events Ω
- and X semimartingale in $(\Omega, \mathcal{F}, \mathbb{P})$

$\rightarrow$ when is X going to remain a semimartingale in $(\Omega, \mathcal{G}, \mathbb{P})$?

Why is this interesting?

- Semimartingales are the processes for which classical stochastic integration works: $\int H dX$ (e.g. asset price proc.)
- Agents have access to different sets of information

Today: $X = B$ Brownian motion in its own filtration $\mathcal{F}^B \subset \mathcal{G}^B$:

- when is B semimartingale w.r.t. $\mathcal{G}^B$? $B_t = \tilde{B}_t + A_t$
- in particular, when is FV $\ll$ Leb? $B_t = \tilde{B}_t + \int_0^t a_s ds$

Filtration enlargement

→ Two most studied types of filtration enlargement:

- **initial enlargement**: $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(L)$
- **progressive enlargement** with a random time: $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau \wedge t)$

Some literature:

T. Jeulin and M. Yor (1978), P. Brémaud and M. Yor (1978), T. Jeulin (1980), J. Jacod (1985), R. Mansuy and M. Yor (2006)

Filtration enlargement

→ Two most studied types of filtration enlargement:

- **initial enlargement**: $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(L)$
- **progressive enlargement** with a random time: $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau \wedge t)$

Some literature:
T. Jeulin and M. Yor (1978), P. Brémaud and M. Yor (1978),
T. Jeulin (1980), J. Jacod (1985), R. Mansuy and M. Yor (2006)

→ We consider a **general enlargement** of a (right cont.) filtration $(\mathcal{F}_t)_{t\in[0,T]}$ to a (right cont.) filtration $(\mathcal{G}_t)_{t\in[0,T]}$:

$$\mathcal{F}_t \subseteq \mathcal{G}_t \ \ \forall t \in [0,T], \ \ , \mathcal{F}_T = \mathcal{G}_T$$

and characterize semim. preservation via causal transport: $\tilde{B} \to B$

Filtration enlargement

- $\mathcal{X} = \mathcal{Y} = C_0[0,\infty)$, W coordinate process: $W_t(\omega) = \omega_t$
- $\mathcal{F}^{\mathcal{X}} = \mathcal{F}$ filtration generated by W
- $\mathcal{F}^{\mathcal{Y}} = \mathcal{G} \supset \mathcal{F}$
- $\mathcal{F}^B = B^{-1}(\mathcal{F})$, $\mathcal{G}^B = B^{-1}(\mathcal{G})$

Example

Let B be a Brownian motion on $(\Omega, \mathcal{F}^B, \mathbb{P})$, which remains a semi-martingale w.r.to the enlarged filtration $\mathcal{G}^B$, with decomposition

$$dB_t = d\tilde{B}_t + dA_t.$$

Then $\mathcal{L}(\tilde{B}, B)$ is a causal transport plan between the spaces $(C_0[0,\infty), \mathcal{F}, \gamma)$ and $(C_0[0,\infty), \mathcal{G}, \gamma)$.

Semimartingale preservation: characterization via COT

Theorem

For any fixed anticipation $\mathcal{G}$, **TFAE:**

i. *any process B which is a Brownian motion on some $(\Omega, \mathbb{P})$, remains a semimartingale in the enlarged filtration $\mathcal{G}^B$;*

ii. *for some $\nu \sim \gamma$, the following causal transport problem is finite*

$$\inf_{\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\nu)} \mathbb{E}^{\pi}[V_T(\overline{\omega} - \omega)].$$

Optimal $\hat{\pi} := (\xi, id)_{\#}\nu$, where $\xi_t(\overline{\omega}) := \overline{\omega}_t - A_t(\overline{\omega})$, with A dual pred. proj. of $(\overline{\omega} - \omega)$ w.r.t. $(\pi, \{\emptyset, C_0[0,T]\} \times \mathcal{G})$, $\forall\ \pi$ with finite cost.

Notation. $(\omega, \overline{\omega})$: generic element in $C[0,T] \times C[0,T]$

V_T: total variation up to T

The absolutely continuous case

We have given a characterization of BM remaining semimartingale in a bigger filtration ($B_t = \tilde{B}_t + A_t$). Now we want to answer:

→ When does it have an absolutely continuous finite variation part? $\left(B_t = \tilde{B}_t + \int_0^t \alpha_s ds \text{ information drift}\right)$

The absolutely continuous case

We have given a characterization of BM remaining semimartingale in a bigger filtration ($B_t = \tilde{B}_t + A_t$). Now we want to answer:

→ When does it have an absolutely continuous finite variation part? $\left(B_t = \tilde{B}_t + \int_0^t \alpha_s ds \right.$ information drift$\left.\right)$

- Brownian bridge: $dB_t = d\tilde{B}_t + \frac{B_T - B_t}{T-t} dt$
- Initial enlargement under Jacod's condition
- Progressive enlargement with a random τ (Jeulin-Yor formula)
- Enlargement with $J_t := \inf_{s \geq t} R_s$, where $dR_t = \frac{1}{R_t} dt + dB_t$: $dB_t = d\tilde{B}_t + 2dJ_t - \frac{1}{R_t} dt$

The absolutely continuous case

We have given a characterization of BM remaining semimartingale in a bigger filtration ($B_t = \tilde{B}_t + A_t$). Now we want to answer:

→ When does it have an absolutely continuous finite variation part? $\left(B_t = \tilde{B}_t + \int_0^t \alpha_s ds \text{ information drift}\right)$

- Brownian bridge: $dB_t = d\tilde{B}_t + \frac{B_T - B_t}{T-t} dt$
- Initial enlargement under Jacod's condition
- Progressive enlargement with a random τ (Jeulin-Yor formula)
- Enlargement with $J_t := \inf_{s \geq t} R_s$, where $dR_t = \frac{1}{R_t} dt + dB_t$: $dB_t = d\tilde{B}_t + 2dJ_t - \frac{1}{R_t} dt$

→ This question can be answered in an analogous way.
[see slides at the end]

Extensions

Our results have natural extensions in two directions:

- → **Multidimensional processes.**
- → **General continuous semimartingales:** for non-Brownian processes, generalization of the definition of causality:

$$\mathbb{E}^{\pi}[(\omega_t - \omega_s) f_s(\overline{\omega})] = 0, \qquad 0 \leq s < t \leq T,\ f_s \in L^{\infty}(C, \mathcal{G}_s, \nu),$$

  which leads to analogous results.

  In particular, if X continuous semimartingale which remains a semimartingale in the enlarged filtration $\mathcal{G}^X$, with $X = \widetilde{X} + N \Rightarrow$ the transport plan $\mathcal{L}(\widetilde{X}, X)$ satisfies the condition above.

MkKean-Vlasov optimal control

N-player stochastic differential game

→ N players with private state processes evolving as

$$dX_t^{N,i} = b_t(X_t^{N,i}, \alpha_t^{N,i}, \bar{\nu}_t^{N,-i})dt + dW_t^i, \quad i = 1, ..., N$$

- $W^1, ..., W^N$ independent Wiener processes
- $\alpha^{N,1}, ..., \alpha^{N,N}$ controls of the N players
- $\bar{\nu}_t^{N,-i} = \frac{1}{N-1} \sum_{j \neq i} \delta_{X_t^{N,j}}$ empirical distrib. states of the other players

N-player stochastic differential game

→ N players with private state processes evolving as

$$dX_t^{N,i} = b_t(X_t^{N,i}, \alpha_t^{N,i}, \bar{\nu}_t^{N,-i})dt + dW_t^i, \quad i = 1, ..., N$$

- $W^1, ..., W^N$ independent Wiener processes
- $\alpha^{N,1}, ..., \alpha^{N,N}$ controls of the N players
- $\bar{\nu}_t^{N,-i} = \frac{1}{N-1}\sum_{j\neq i}\delta_{X_t^{N,j}}$ empirical distrib. states of the other players

→ Objective of player i: to choose a control $\alpha^{N,i}$ that minimizes

$$\mathbb{E}\left[\int_0^T f_t(X_t^{N,i}, \alpha_t^{N,i}, \bar{\eta}_t^{N,-i})dt + g(X_T^{N,i}, \bar{\nu}_T^{N,-i})\right]$$

- $\bar{\eta}_t^{N,-i} = \frac{1}{N-1}\sum_{j\neq i}\delta_{(X_t^{N,j},\alpha_t^{N,j})}$ empirical distrib. of states & controls

→ Statistically identical players: same functions b_t, f_t, g

N-player stochastic differential game

Problems:

- search for equilibria: very difficult
- even when they exist, difficult to characterize

N-player stochastic differential game

Problems:

- search for equilibria: very difficult
- even when they exist, difficult to characterize

Idea:

- for large symmetric games, some averaging is expected when the number of players tends to infinity
- resort to approximation by **asymptotic arguments:**

N-player game – – – – – – – – – > $N \to \infty$

N-player stochastic differential game

Problems:

- search for equilibria: very difficult
- even when they exist, difficult to characterize

Idea:

- for large symmetric games, some averaging is expected when the number of players tends to infinity
- resort to approximation by **asymptotic arguments:**

N-player game	– – – – – – – –– >	$N \to \infty$
Nash equilibrium (non-coop)	– – –– >	Mean Field Game
Social planner (cooperative)	– – –– >	McKean Vlasov

McKean-Vlasov control problem

$\rightarrow$ Asymptotic formulation in the case of cooperative equilibria, as well as for non-cooperative equilibria in the potential case:

McKean-Vlasov control problem:

$$\inf_{\alpha} \mathbb{E}\left[\int_0^T f_t\left(X_t, \alpha_t, \mathcal{L}(X_t, \alpha_t)\right) dt + g\left(X_T, \mathcal{L}(X_T)\right)\right]$$

$$\text{subject to} \quad dX_t = b_t\left(X_t, \alpha_t, \mathcal{L}(X_t)\right) dt + dW_t, \quad X_0 = 0$$

Some literature on MFG and MKV:

- J.M. Lasry and P.L. Lions (2006, ’07)
- M. Huang, P.E. Caines, and R.P. Malhamé (2006, ’07)
- P. Cardaliaguet, R. Carmona, F. Delarue, M. Fischer, J.P. Fouque, A. Lachapelle, D. Lacker, C.A. Lehalle, H. Pham, X. Wei ...

McKean-Vlasov control problem

Classical approaches for MFG/MKV:

- analytic (Lasry-Lions): HJB, forward-backward system of PDEs
- probabilistic: Pontryagin maximum principle, FBSDEs

McKean-Vlasov control problem

Classical approaches for MFG/MKV:

- analytic (Lasry-Lions): HJB, forward-backward system of PDEs
- probabilistic: Pontryagin maximum principle, FBSDEs

Our approach: use causal transport: $W \to X$, with the aim of:

$\hookrightarrow$ providing different existence results

$\hookrightarrow$ finding explicit solutions

McKean-Vlasov control problem and Causal Transport

→ McKean-Vlasov control problem:

$$\inf_{\alpha} \mathbb{E}\left[\int_0^T f_t\left(X_t, \alpha_t, \mathcal{L}(X_t, \alpha_t)\right) dt + g\left(X_T, \mathcal{L}(X_T)\right)\right]$$

subject to

$$dX_t = b_t\left(X_t, \alpha_t, \mathcal{L}(X_t)\right) dt + dW_t, \quad X_0 = 0$$

→ The joint distribution $\mathcal{L}(W, X)$ is a causal transport plan between $(C_0[0,T], \mathcal{F}, \gamma)$ and $(C_0[0,T], \mathcal{F}, \mathcal{L}(X))$:

$$\gamma \longrightarrow ? = \text{distribution of the state}$$

McKean-Vlasov control problem

Definition. A weak solution to the McKean-Vlasov control problem is a tuple $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}, W, X, \alpha)$ such that:

(i) $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ supports X, BM W, α is $\mathcal{F}$-progress. meas.

(ii) the state equation $dX_t = b_t(X_t, \alpha_t, \mathcal{L}(X_t))\,dt + dW_t$ holds

(iii) if $(\Omega', (\mathcal{F}'_t)_{t\in[0,T]}, \mathbb{P}', W', X', \alpha')$ is another tuple s.t. (i)-(ii) hold,

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^T f_t(X_t, \alpha_t, \mathcal{L}_{\mathbb{P}}(X_t, \alpha_t))\,dt + g(X_T, \mathcal{L}_{\mathbb{P}}(X_T))\right]$$
$$\leq \mathbb{E}^{\mathbb{P}'}\left[\int_0^T f_t\left(X'_t, \alpha'_t, \mathcal{L}_{\mathbb{P}'}(X'_t, \alpha'_t)\right) dt + g\left(X'_T, \mathcal{L}_{\mathbb{P}'}(X'_T)\right)\right]$$

Assumptions

$\rightarrow$ We need some **convexity assumptions**.

Assumptions

→ We need some **convexity assumptions**.

→ In the case of linear drift:

$$dX_t = (c_t^1 X_t + c_t^2 \alpha_t + c_t^3 \mathbb{E}[X_t])dt + dW_t,$$

$c_t^i \in \mathbb{R}, c_t^2 > 0$, our assumptions reduce to: for all x, a, η,

- f_t is bounded from below uniformly in t
- $f_t(x, ., \eta)$ is convex
- $f_t(x, a, .)$ is $\prec_{\text{conv}}$-monotone

Assumptions

→ We need some **convexity assumptions**.

→ In the case of linear drift:

$$dX_t = (c_t^1 X_t + c_t^2 \alpha_t + c_t^3 \mathbb{E}[X_t])dt + dW_t,$$

$c_t^i \in \mathbb{R}, c_t^2 > 0$, our assumptions reduce to: for all x, a, η,

- f_t is bounded from below uniformly in t
- $f_t(x, ., \eta)$ is convex
- $f_t(x, a, .)$ is $\prec_{\text{conv}}$-monotone

E.g. satisfied by the inter-bank borrowing & lending model of Carmona-Fouque-Sun [see slides at the end]

Characterization of weak MKV solutions via COT

Here $\mathcal{X} = \mathcal{Y} = C_0[0,T]$, and for simplicity control = drift, sq. integr.

Theorem

The weak MKV problem is **equivalent** *to the variational problem*

$$\inf_{\nu \ll \gamma} \inf_{\pi \in \Pi_c(\gamma,\nu)} \mathbb{E}^\pi \left[\int_0^T f_t\Big(\overline{\omega}_t, (\dot{\widehat{\overline{\omega} - \omega}})_t, p_t\big((\overline{\omega}, \dot{\widehat{\overline{\omega} - \omega}})_\# \pi\big)\Big) dt + g(\overline{\omega}_T, \nu_T) \right]$$

$$\Big(= \inf_{\pi \in \Pi_c(\gamma,.)} \mathbb{E}^\pi [...] \Big)$$

Notation: $(\dot{\widehat{\overline{\omega} - \omega}})_t = \beta_t$ when $\overline{\omega} - \omega = \int_0^\cdot \beta_s ds$, and $+\infty$ else

[for the general case: see slides at the end]

Characterization of weak MKV solutions via COT

'Equivalence' means that the above variational problem and

$$\inf \; \mathbb{E}^{\mathbb{P}}\left[\int_0^T f_t\left(X_t, \alpha_t, \mathcal{L}(X_t, \alpha_t)\right) dt + g\left(X_T, \mathcal{L}(X_T)\right)\right]$$

have the same value, where the infimum is taken over tuples $(\Omega, (\mathcal{F}_t), \mathbb{P}, W, X, \alpha)$ s.t. $dX_t = b_t\left(X_t, \alpha_t, \mathcal{L}(X_t)\right) dt + dW_t$, and that moreover the optimizers are related via:

- $\nu^* = \mathcal{L}(X^*)$
- $\pi^* \longleftrightarrow \alpha^*$, with $\pi^* = \mathcal{L}(W^*, X^*)$

Characterization of weak MKV solutions via COT

Corollary (Weak closed loop)

1. *The infimum can be taken over tuples s.t. α is $\mathcal{F}^X$-measurable (weak closed loop).*
2. *If the infimum is attained, then the optimal control α is in weak closed loop form.*

Characterization of weak MKV solutions via COT

Corollary (Weak closed loop)

1. *The infimum can be taken over tuples s.t. α is $\mathcal{F}^X$-measurable (weak closed loop).*
2. *If the infimum is attained, then the optimal control α is in weak closed loop form.*

Corollary (Existence)

Minimizations in the variational problem done over compact sets, hence lower-semicontinuity $\Rightarrow$ existence.

Note: in classical the approaches, strong regularity is required

Special case: separable costs [see slides at the end]

Dynamic Cournot-Nash equilibria

Problem formulation

Given: a population of Agents whose **type evolves in time**. Each of them:

- → selects its own actions/strategies
- → faces a cost depending on its own type, action, and on the symmetric interaction with the rest of the population:

$$\text{cost(i)} = \text{fcn}\,\big(\text{type(i), action(i), actions distribution}\big)$$

Problem formulation

Given: a population of Agents whose **type evolves in time**. Each of them:

- → selects its own actions/strategies
- → faces a cost depending on its own type, action, and on the symmetric interaction with the rest of the population:

$$\text{cost(i)} = \text{fcn}\left(\text{type(i), action(i), actions distribution}\right)$$

Crucial: the actions of a player should not anticipate its type!

Problem formulation

Given: a population of Agents whose **type evolves in time**. Each of them:

- → selects its own actions/strategies
- → faces a cost depending on its own type, action, and on the symmetric interaction with the rest of the population:

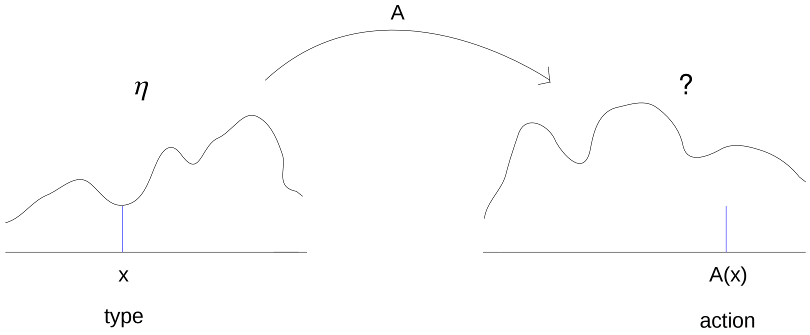
$$\text{cost(i)} = \text{fcn}\,\big(\text{type(i), action(i), actions distribution}\big)$$

Crucial: the actions of a player should not anticipate its type!

Our aim is to:

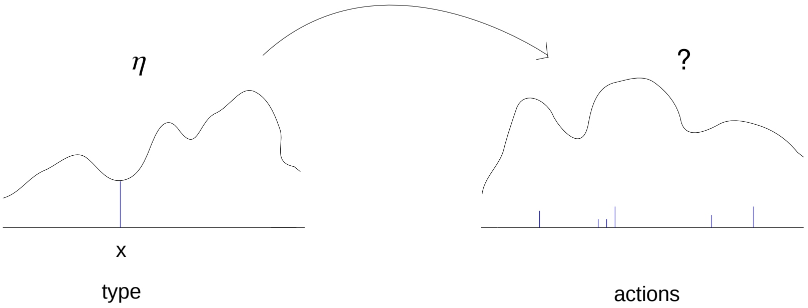
- → find/characterize **equilibria** for games in this setting
- → develop/exploit connection with causal optimal transport

Pure strategies



adapted **pure** strategy = adapted **Monge** transport

Mixed strategies



non-anticipative **mixed** strategy = causal **Kantorovich** transport

Main results

- mixed strategy equilibria are solutions to COT problems
- pure strategy equilibria are solutions to COT problems over Monge maps
- for potential games, we characterize Cournot-Nash equilibria as solutions to a variational problem involving COT problems:
  - ↪ new existence results
  - ↪ new uniqueness results
  - ↪ results on first structure of equilibria

[for precise setting and results see slides at the end]

Value of information in stochastic optimization problems

Value of information

- **Aim:** use causal transport framework to give an estimate of the value of the additional information, for some classical stochastic optimization problems (difference of optimal value of these problems with or without additional information).
- **Idea:** take projection w.r.to causal couplings of the optimizers in the problem with the larger filtration (additional information), so building a feasible element in the problem with the smaller filtration and making a comparison possible.
- Pflug (2009) uses this idea in discrete-time, to gauge the dependence of multistage stochastic programming problems w.r.to different reference probability measures.

$\rightarrow$ Here we see utility maximixation [for optimal stopping: see slides at the end]

Utility maximisation

- B d-dimensional Brownian motion on $(\Omega, \mathbb{P})$.
- Financial market: riskless asset $\equiv 1$, and $m \leq d$ risky assets:

$$dS_t^i = S_t^i \Big(b_t^i \, dt + \sum_{j=1}^{d} \sigma_t^{ij} dB_t^j \Big), \qquad i = 1, \ldots, m.$$

- $|b_t^i(\omega) - b_t^i(\tilde{\omega})| \leq L \sum_{k=1}^{d} \sup_{s \leq t} |\omega_s^k - \tilde{\omega}_s^k|$, same for σ^{ij}, σ bdd

Utility maximisation

- B d-dimensional Brownian motion on $(\Omega, \mathbb{P})$.
- Financial market: riskless asset $\equiv 1$, and $m \leq d$ risky assets:

$$dS_t^i = S_t^i \Big(b_t^i \, dt + \sum_{j=1}^{d} \sigma_t^{ij} dB_t^j \Big), \qquad i = 1, \ldots, m.$$

- $|b_t^i(\omega) - b_t^i(\tilde{\omega})| \leq L \sum_{k=1}^d \sup_{s \leq t} |\omega_s^k - \tilde{\omega}_s^k|$, same for σ^{ij}, σ bdd
- λ_t^i: proportion of an agent's wealth invested in the i^{th} stock at time t: assume $\lambda_t^i \in [0, 1]$ (no short-selling)
- $\mathcal{A}(\mathcal{F}^B)$: set of admissible portfolios for the agent without anticipative information ($\mathcal{F}^B$-progressively measurable λ)
- $\mathcal{A}(\mathcal{G}^B)$: set of admissible portfolios for the agent wit anticipative information ($\mathcal{G}^B$-progressively measurable λ)

Utility maximisation

→ We want to compare the utility maximization problems:

$$v^{\mathcal{F}} = \sup_{\lambda \in \mathcal{A}(\mathcal{F}^B)} \mathbb{E}[U(X_T^\lambda)], \qquad v^{\mathcal{G}} = \sup_{\lambda \in \mathcal{A}(\mathcal{G}^B)} \mathbb{E}[U(X_T^\lambda)].$$

- $(X_t^\lambda)_t$: wealth process corresponding to λ, $X_0^\lambda = 1$.
- utility function $U : \mathbb{R}_+ \to \mathbb{R}$ concave, increasing, and s.t.

$$F := U \circ exp$$ is C-Lipschitz, concave and increasing.

e.g. $U(x) = \frac{x^a}{a}$, $a \leq 0$; $U(x) = ln(x)$; $U(x) = -\frac{1}{a}e^{-ax}$, $a \geq 1$

Utility maximisation

$\rightarrow$ We want to compare the utility maximization problems:

$$v^{\mathcal{F}} = \sup_{\lambda \in \mathcal{A}(\mathcal{F}^B)} \mathbb{E}[U(X_T^\lambda)], \qquad v^{\mathcal{G}} = \sup_{\lambda \in \mathcal{A}(\mathcal{G}^B)} \mathbb{E}[U(X_T^\lambda)].$$

- $(X_t^\lambda)_t$: wealth process corresponding to λ, $X_0^\lambda = 1$.
- utility function $U : \mathbb{R}_+ \rightarrow \mathbb{R}$ concave, increasing, and s.t.

  $F := U \circ exp$ is C-Lipschitz, concave and increasing.

  e.g. $U(x) = \frac{x^a}{a}$, $a \leq 0$; $U(x) = ln(x)$; $U(x) = -\frac{1}{a}e^{-ax}$, $a \geq 1$

Proposition

The following bound holds, for a specific constant K:

$$0 \leq v^{\mathcal{G}} - v^{\mathcal{F}} \leq K \inf_{\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\gamma)} \mathbb{E}^\pi[V_T(\overline{\omega} - \omega)].$$

Causal transports on $C_0[0,T] \times C_0[0,T]$

Utility maximisation

Remark. If complete market, log utility, and initial enlargement, then $v^{\mathcal{G}} - v^{\mathcal{F}}$ is known explicitly (Pikovsky-Karatzas 1996).

Utility maximisation

Remark. If complete market, log utility, and initial enlargement, then $v^{\mathcal{G}} - v^{\mathcal{F}}$ is known explicitly (Pikovsky-Karatzas 1996).

Steps of the proof:

- fix a causal transport $\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma, \gamma)$
- consider $v^{\mathcal{F}}$ to be solved in the ω variable and $v^{\mathcal{G}}$ in $\overline{\omega}$
- take (ϵ-)optimizer $\hat{\lambda} = \hat{\lambda}(\overline{\omega})$ for $v^{\mathcal{G}}$
- $(\pi, \mathcal{F} \otimes \{\emptyset, C\})$-optional projection: $\tilde{\lambda} \in \mathcal{A}(\mathcal{F}^B)$
- in particular $\tilde{\lambda}_t(\omega) = \tilde{\lambda}_t(\omega, \overline{\omega}) = \mathbb{E}^{\pi}[\hat{\lambda}_t | \mathcal{F}_t] = \mathbb{E}^{\pi}[\hat{\lambda}_t | \mathcal{F}_T]$
- substitute in $v^{\mathcal{F}}$

Conclusions

We have exploited causal transport to study several problems:

- semimartingale preservation:

$$\mathcal{G}-\text{BM} \longrightarrow \mathcal{F}-\text{BM}$$

- weak solutions to MKV:

$$\text{noise} \longrightarrow \text{state dynamics}$$

- Cournot-Nash equilibria:

$$\text{types} \longrightarrow \text{actions}$$

- value of information:

$$\mathcal{G}-\text{BM} \longrightarrow \mathcal{F}-\text{BM}$$

THANK YOU FOR YOUR ATTENTION!

Addendum to the Semimartingale preservation section

The absolutely continuous case

- Semimartingale preservation $\Longleftrightarrow$ finite causal transport problem for the total variation cost
- In the A.C. case, the cost functions of interest are

$$c_\rho(\omega,\overline{\omega}) := \int_0^T \rho\big((\dot{\widehat{\overline{\omega}-\omega}})_t\big)dt,$$

where $\rho : \mathbb{R} \to \mathbb{R}_+$ is convex, even, $\rho(0) = 0$, $\rho(+\infty) = +\infty$, and $(\dot{\widehat{\overline{\omega}-\omega}})_t = \beta_t$ when $\overline{\omega} - \omega = \int_0^\cdot \beta_s ds$, and $+\infty$ else. E.g. $\rho(x) = x^2/2 \Rightarrow$ Cameron-Martin cost $c_\rho(\omega,\overline{\omega}) = \frac{1}{2}|\overline{\omega} - \omega|_H^2$.

$\rightarrow$ In the A.C. case we get a characterization of the semim. preservation via COT over π's under which $\overline{\omega} - \omega \ll$ Leb.

The absolutely continuous case

Theorem

For any fixed anticipation $\mathcal{G}$, **TFAE:**

i. *any process B which is a Brownian motion on some $(\Omega, \mathbb{P})$, remains a semimartingale in the enlarged filtration $\mathcal{G}^B$, with absolutely continuous FV part, i.e.*

$$dB_t = d\tilde{B}_t + \alpha_t(B)dt;$$

ii. *for some $\nu \sim \gamma$, and some ρ as above (eqv., for $\rho = |\,.\,|$) the following causal transport problem is finite*

$$\inf_{\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\nu)} \mathbb{E}^\pi[c_\rho(\omega, \overline{\omega})].$$

Optimal transport $\hat{\pi} := (\xi, id)_\# \nu$, where $\xi_t(\overline{\omega}) := \overline{\omega}_t - \int_0^t a_s(\overline{\omega})ds$, a pred.pr. of $\widehat{\overline{\omega} - \omega}$ w.r.t. $(\pi, \{\emptyset, C\} \times \mathcal{G})$, $\forall\, \pi$ with finite cost.

Cameron-Martin cost

Let $c_\rho(\omega,\overline{\omega}) = \frac{1}{2}|\overline{\omega}-\omega|_H^2$. If $\inf_{\pi\in\Pi^{\mathcal{F},\mathcal{G}}(\gamma,\gamma)} \mathbb{E}^\pi[c_\rho] < \infty$, then:

- $dB_t = d\tilde{B}_t + \alpha_t(B)dt$, with α square integrable;
- by Girsanov, B BM w.r.t. $\mathcal{G}^B$ under a new measure $\mathbb{Q}$, and

$$\mathrm{COT} = \frac{1}{2}\mathbb{E}^\gamma\Big[\int_0^T \alpha_t^2 dt\Big] = H(\mathbb{P}|\mathbb{Q});$$

- by martingale representation, $\mathcal{H}'$-hypothesis holds for $\mathcal{F}^B$, $\mathcal{G}^B$, i.e. all $\mathcal{F}^B$-semimartingales are $\mathcal{G}^B$-semimartingales;
- in case of initial enlargement with r.v. $L = L(B)$ with law ℓ:

$$\mathrm{COT} = \int H(\gamma^{L=x}|\gamma)\ell(dx) = \mathrm{I}(B, L(B)) := H(P_{B,L(B)}|P_B \otimes P_{L(B)}),$$

mutual information between B and $L(B)$;
- if L is discrete, COT= entropy of the partition $\{L = x_n\}_n$:

$$\mathrm{COT} = -\sum_n p_n \ln(p_n), \quad p_n = \mathbb{P}(L = x_n).$$

Addendum to the McKean-Vlasov section

Example: Inter-bank borrowing & lending

Carmona-Fouque-Sun (2013)

- Consider a network of N banks, with log-monetary reserve

$$\begin{aligned} dX_t^i &= \left[\frac{k}{N-1}\sum_{j\neq i}\left(X_t^j - X_t^i\right) + \alpha_t^i\right]dt + dW_t^i, \\ &= \left[k\left(\overline{X}_t^{N,-i} - X_t^i\right) + \alpha_t^i\right]dt + dW_t^i, \quad i = 1, ..., N \end{aligned}$$

  - $k \geq 0$ rate of m-r in the interaction from b&l between banks
  - α^i control of bank i, b&l outside of the N bank network

- Bank i tries to minimize the cost

$$\mathbb{E}\left[\int_0^T \left(\frac{1}{2}(\alpha_t^i)^2 - q\alpha_t^i(\overline{X}_t^{N,-i} - X_t^i) + \frac{c}{2}(\overline{X}_t^{N,-i} - X_t^i)^2\right)dt + \frac{d}{2}(\overline{X}_T^{N,-i} - X_T^i)^2\right]$$

  - $q > 0$ incentive to borrowing ($\alpha_t > 0$) or lending ($\alpha_t < 0$)
  - $c, d > 0$ penalize departure from average

Example: Inter-bank borrowing & lending

In the previous example:

- The log-monetary reserve of each bank, asymptotically, is governed by the MKV equation

$$dX_t = [k(\mathbb{E}[X_t] - X_t) + \alpha_t]dt + dW_t$$

(all banks control their rate of b&I with the same policy α)

- Need to minimize the cost

$$\mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\alpha_t^2 - q\alpha_t(\mathbb{E}[X_t]-X_t) + \frac{c}{2}(\mathbb{E}[X_t]-X_t)^2\Big)dt + \frac{d}{2}(\mathbb{E}[X_T]-X_T)^2\Big]$$

Separable cost

Special case: separable running cost = $f_t^1(x,a) + f_t^2(\nu_t, x)$:

$$\inf_{\nu \ll \gamma} \left\{ \mathrm{COT}(\gamma, \nu, c_{f^1}) + F_{f^2,g}(\nu) \right\}$$

↑ standard COT (A. et al. 2016)

↑ penalty

- For COT easy to get existence (& uniqueness) of $\pi^* \in \Pi_c(\gamma, \nu)$
- $\nu \mapsto$COT(γ, ν, c_{f^1}) convex
- Need conditions on F to have existence/uniqueness, e.g.
  - F lsc $\Rightarrow$ exist ν^*
  - F strictly convex $\Rightarrow$ unique ν^*

Separable cost

Example: take $k = q = 0$ in the example above, then

- state dynamics: $dX_t = \alpha_t dt + dW_t$
- cost: $\mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\alpha_t^2 + \frac{c}{2}(\mathbb{E}[X_t] - X_t)^2\Big)dt + \frac{d}{2}(\mathbb{E}[X_T] - X_T)^2\Big]$

$\Rightarrow$ COT w.r.t. Cameron-Martin distance (Lassalle 2015):

$$\frac{1}{2}\inf_{\pi\in\Pi_c(\gamma,\nu)} \mathbb{E}^\pi[|\overline{\omega} - \omega|_H^2] = H(\nu|\gamma), \quad \text{thus}$$

$$\inf_{\nu\ll\gamma}\left\{H(\nu|\gamma) + \frac{c}{2}\int_0^T \mathrm{Var}(\nu_t)dt + \frac{d}{2}\mathrm{Var}(\nu_T)\right\}$$

More generally: for cost $\frac{1}{2}\alpha_t^2 + h_t(X_t, \mathcal{L}(X_t))$, by Sanov theorem:

$$\inf_{\nu\ll\gamma}\left\{H(\nu|\gamma) + F(\nu)\right\} = \lim_{n\to\infty} -\frac{1}{n}\ln\mathbb{E}\, e^{nF(\frac{1}{n}\sum_{i=1}^n \delta_{W^i})}, \ \{W_i\} \text{ ind. BMs.}$$

And this does not seem to be limited to the entropic case $(\frac{1}{2}\alpha_t^2)$.

Characterization: the general case

Assumptions. For all $x, a \in \mathbb{R}, m \in \mathcal{P}(\mathbb{R}), \eta \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$:

(A1) $b_t(x, ., m)$ injective and convex

(A2) f_t bdd below unif. in t, and $f_t(x, b_t^{-1}(x, ., m)(y), \eta)$ convex in y

(A3) $f_t(x, a, .)$ is $\prec_{cm}$-monotone (resp. $\prec_{conv}$-monotone if b is linear)

($\prec_{cm}$ (resp. $\prec_{conv}$) denotes the conv/monotone (resp. conv) order)

Pathwise quadratic variation. For $\omega \in C := C_0[0, T]$, $n \in \mathbb{N}$, let

$$\sigma_0^n(\omega) := 0, \quad \sigma_{k+1}^n(\omega) := \inf\{t > \sigma_k^n(\omega) : |\omega(t) - \omega(\sigma_k^n)| \geq 2^{-n}\}, \; k \in \mathbb{N}$$

We say that ω has quadratic variation if

$$V_n(\omega)(t) := \sum_{k=0}^{\infty} (\omega(\sigma_{k+1}^n \wedge t) - \omega(\sigma_k^n \wedge t))^2 \to_u \; =: \langle \omega \rangle_t \in C$$

Notation. $\tilde{\mathcal{P}} = \{\nu \in \mathcal{P}(C) : \langle \omega \rangle \; \exists \; \nu\text{-a.s., with } \langle \omega \rangle_t = t \text{ for all } t\}$

Characterization: the general case

Under the above assumptions, the following characterization of weak McKean-Vlasov solutions via causal transport holds.

Theorem

The weak MKV problem is **equivalent** *to the following problem*

$$\inf_{\nu \in \tilde{\mathcal{P}}} \inf_{\pi \in \Pi_c(\gamma,\nu)} \mathbb{E}^{\pi}\left[\int_0^T f_t\left(\overline{\omega}_t, u_t^{\nu}(\omega,\overline{\omega}), p_t\big((\overline{\omega}, u^{\nu})_{\#}\pi\big)\right) dt + g(\overline{\omega}_T, \nu_T)\right]$$

where $u_t^{\nu}(\omega,\overline{\omega}) = b_t^{-1}(\overline{\omega}_t, ., \nu_t)\big((\dot{\overbrace{\overline{\omega} - \omega}})_t\big)$ *and* $p_t(\eta) = \eta_t$.

Addendum to the CN-equilibria section

Setting

- Discrete time $\mathbb{T} = \{1, ..., N\}$.
- $\mathcal{X}$ = path-space of types, and $\mathcal{Y}$ = path-space of actions, e.g.
$$\mathcal{X} = \mathcal{Y} = \mathbb{R}^N.$$
- $\eta \in \mathcal{P}(\mathcal{X})$: types' distribution (of public knowledge).

Cost function: $k(x, y, \nu) = c(x, y) + V[\nu](y)$

- x: type $x \in \mathcal{X}$
- y: action $y \in \mathcal{Y}$
- ν: actions distribution $\nu \in \mathcal{P}(\mathcal{Y})$
- $c(x, y)$: idiosyncratic part
- $V[\nu](y)$: mean-field interaction

Usually,

$$c(x, y) = \sum_{t=1}^{N} c_t(x_{1:t}, y_{1:t}), \quad V[\nu](y) = \sum_{t=1}^{N} V_t[\nu_{1:t}](y_{1:t})$$

Setting

The mean-field interaction term may capture repulsive/attractive effects. For example:

→ **Congestion effect**: $V^c[\nu](y) = f\left(y, \frac{d\nu}{dm}(y)\right)$, with $m \in \mathcal{P}(\mathcal{Y})$ reference measure wrt which congestion measured, $f(y,.) \nearrow$

→ **Attractive effect**: $V^a[\nu](y) = \int_{\mathcal{Y}} \phi(y,z)d\nu(z)$, with ϕ cont, symmetric, convex, minimal on the diagonal

Non-cooperative equilibrium with a continuum of agents.

Static case:

- Schmeidler (1973)
- Mas-Colell (1984)
- . . .
- Blanchet and Carlier (2015)

A motivating example: n-player game

- $\mathbb{T} = \{1, 2\}$ meaning *this week* or *next week*.
- $\mathcal{X} = \{(s, s), (s, h), (h, s), (h, h)\}$, with
  s = *sick* and h = *healthy*.
- $\mathcal{Y} = \{(v, v), (v, w), (w, v), (w, w)\}$ with
  v = *vacation* and w = *work*.
- $\eta^{x_1}(x_2) = \mathbb{P}(\text{time-2 type} = x_2 \,|\, \text{time-1 type} = x_1)$.

We denote (x_1^i, x_2^i) and (y_1^i, y_2^i) for Agent i's types and actions.

The Game: Agents must decide "now" how they will distribute *work* and *vacation* for *this* and *next week*, taking into account:

→ current types $x_1^i \in \{s, h\}$, and priors $\eta^{x_1^i} \in \mathcal{P}(\{s, h\})$ (known)

→ the fact that taking holidays gets more expensive if many people are thinking likewise

A motivating example: n-player game

Agent i selects $y_1^i \in \{v,w\}$ for current week, and guesses a time-2 action $y_2^i(x_2^i) \in \{v,w\}$ depending on its unknown time-2 type.

The cost of such arrangement, seen from now, is

$$J^i\left(\{y_1^i, y_2^i(.)\}, \{y_1^k, y_2^k(.)\}_{k\neq i}\right) := c_1(x_1^i, y_1^i) + V_1\left[\frac{1}{n-1}\sum_{k\neq i}\delta_{y_1^k}\right](y_1^i)$$

$$+\int\Big\{c_2(x_1^i, x_2^i, y_1^i, y_2^i(x_2^i)) + V_2\left[\frac{1}{n-1}\sum_{k\neq i}\delta_{(y_1^k, y_2^k(x_2^k))}\right](y_1^i, y_2^i(x_2^i))\Big\}\otimes_k \eta^{x_1^k}(dx_2^k).$$

Definition (Dynamic Nash equilibrium)

$\{y_1^i, y_2^i(\cdot)\}_{i=1}^n$ is a dynamic Nash equilibrium if, for all i,

$$(y_1^i, y_2^i(\cdot)) \in \text{argmin}_{(a,A(\cdot))}\ J^i\left(\{a, A(\cdot)\}, \{y_1^k, y_2^k(.)\}_{k\neq i}\right)$$

A motivating example: n-player game

Problems:

- search for equilibria: very difficult
- even when they exist, difficult to characterize

Idea:

- If size n of population is big, one tries to approximate this difficult equilibrium problem by a hopefully simpler one
- For this we need that, as $n \to \infty$, the empirical distributions of the n-player game equilibria converge to distributions that corresponds to equilibria for infinitely many players
  (dynamic Nash equilibria $\to$ dynamic Cournot-Nash equilibria)

$\rightarrow$ From now on we think of the limiting case (**infinitesimal agents**)

Pure adapted strategies

Pure strategy: all players of type-path $x \in \mathcal{X}$ choose same strategy

$$\mathcal{Y} \ni y = A(x) = (A_t(x))_{t \in \mathbb{T}}$$

Adapted strategy: $A_t(x) = A_t(x_{1:t})$ for all $t \in \mathbb{T}$

Denote by $\mathcal{A}$ the set of pure adapted strategies $A : \mathcal{X} \to \mathcal{Y}$

- types' distribution: $\eta \in \mathcal{P}(\mathcal{X})$ (of public knowledge)
- strategies' distribution: $\nu = A(\eta) \in \mathcal{P}(\mathcal{Y})$ (to be determined in equilibrium)

Pure equilibrium

Definition ((Pure) dynamic Cournot-Nash equilibrium)

$(A^*, \nu^*) \in \mathcal{A} \times \mathcal{P}(\mathcal{Y})$ is called dynamic Cournot-Nash equilibrium if

- A^* attains

$$P(\nu^*) := \inf_{A \in \mathcal{A}} \int_{\mathcal{X}} \{c(x, A(x)) + V[\nu^*](A(x))\} d\eta(x)$$

- and $A^*(\eta) = \nu^*$

→ "minimization of an average cost + fixed point condition"

→ Pure equilibria known to rarely exists, so we shall consider: generalization to mixed-strategy (i.e. randomized) equilibria

Mixed non-anticipative strategy

mixed-strategy: players of same type can choose different actions

non-anticipative: $A_t(x) = \text{fcn}(x_{1:t}) + \text{sth independent of } x$

- Non-anticipative mixed-strategy = causal (Kantorovic) transport
- The causal Monge transports $\pi = (id, A)(\eta) \in \Pi_c(\eta, .)$ are the pure adapted strategies with prior η on types.
- The set of pure adapted strategies with prior η is dense in (and equals the extreme point of) the set $\Pi_c(\eta, .)$.

Mixed-strategy equilibrium

For $\nu \in \mathcal{P}(\mathcal{Y})$, denote $M(\nu) := \inf_{\pi \in \Pi_c(\eta,.)} \mathbb{E}^\pi \left[c(x,y) + V[\nu](y) \right]$.

Definition (Mixed-strategy dynamic Cournot-Nash equilibrium)

$(\pi^*, \nu^*) \in \Pi_c(\eta, .) \times \mathcal{P}(\mathcal{Y})$ is called a mixed-strategy equilibrium if

- π^* attains $M(\nu^*)$,
- with $\nu^* = p_2(\pi^*)$, i.e., $\pi^* \in \Pi_c(\eta, \nu^*)$.

→ Mixed-strategy equilibria are solutions to causal transport problems, i.e. π^* as above does also attain

$$\inf_{\pi \in \Pi_c(\eta,\nu^*)} \mathbb{E}^\pi [c(x,y)].$$

→ Analogously, pure equilibria are solutions to causal transport problems over Monge maps.

Potential games

We have seen that equilibrium $\Longrightarrow$ optimal transport

For **potential games**, we will have "$\Longleftrightarrow$" in a precise sense

Definition (Potential Game)

V is the **first variation** of an energy functional $\mathcal{E} : \mathcal{P}(\mathcal{Y}) \to \mathbb{R}$:

$$\lim_{\epsilon \to 0^+} \frac{\mathcal{E}(\nu + \epsilon(\mu - \nu)) - \mathcal{E}(\nu)}{\epsilon} = \int_{\mathcal{Y}} V[\nu] d(\mu - \nu), \quad \forall\, \nu, \mu \in \mathcal{P}(\mathcal{Y})$$

E.g. for congestion and attractive costs V^c and V^a, we have

$$\mathcal{E}^c(\nu) = \int_{\mathcal{Y}} F\Big(y, \frac{d\nu}{dm}(y)\Big) dm(y), \ \mathcal{E}^a(\nu) = \frac{1}{2} \int_{\mathcal{Y} \times \mathcal{Y}} \phi(y, z) d\nu(z) d\nu(y),$$

where $F(y, u) = \int_0^u f(y, s) ds$.

Potential games

Consider the variational problem

$$(VP) \qquad \inf_{\nu \in \mathcal{P}(\mathcal{Y})} \Big\{ \underbrace{\inf_{\pi \in \Pi_c(\eta,\nu)} \mathbb{E}^\pi[c(x,y)]}_{\mathrm{CT}(\eta,\nu)} + \mathcal{E}[\nu] \Big\}$$

Theorem

Let $\mathcal{E}$ be convex, then the following are **equivalent***:*

(i) (π^*, ν^*) *is a mixed-strategy equilibrium;*

(ii) ν^* *solves (VP), and* π^* *solves* $CT(\eta, \nu^*)$.

Convexity of $\mathcal{E}$ only needed for "$(i) \Rightarrow (ii)$".

Remark. Note:

$$(VP) = \inf_{\pi \in \Pi_c(\eta,\cdot)} \left\{ \mathbb{E}^\pi[c] + \mathcal{E}[p_2(\pi)] \right\}.$$

Thus if $\mathcal{E}$ *concave* and $\exists$ equilibrium $\Rightarrow$ $\exists$ *pure* equilibrium.

Potential games

Corollary (existence)

Let c be l.s.c. and bounded below. Then

- $V = V^c$ *and growth condition on* $f \implies \exists$ *m-s equilibrium;*
- $V = V^a$ *and growth condition on* $c \implies \exists$ *m-s equilibrium.*

Growth conditions ensure existence of a solution ν^* to (VP), and CT(η, ν^*) admits a solution π^* easily. Now apply previous theorem.

Corollary (uniqueness)

If $\mathcal{E}$ strictly convex $\implies$ all m-s equilibria have same second marginal ν^*, i.e., unique optimal distribution of actions.

Indeed, $\nu \mapsto$CT(η, ν) convex, hence $\mathcal{E}$ strictly convex implies unique solution ν^* for (VP). Then apply previous theorem.

Addendum to the Value of information section

Optimal stopping

- With the same method used above, we can estimate the value of information wrt other optimization problems, e.g.

$$v^{\mathcal{F}} := \inf_{\mathcal{F}^W\text{-}st.t.} \mathbb{E}^{\mathbb{P}}\left[\ell(W,\tau)\right], \quad v^{\mathcal{G}} := \inf_{\mathcal{G}^W\text{-}st.t.} \mathbb{E}^{\mathbb{P}}\left[\ell(W,\tau)\right],$$

where $\ell : C[0,T] \times \mathbb{R}_+$ cost function, W BM

Proposition

Let ℓ be $\mathcal{F}$-optional, and K-Lipschitz in its first argument wrt a metric d on $C \times C$, uniformly in time. Then

$$0 \leq v^{\mathcal{F}} - v^{\mathcal{G}} \leq K \inf_{\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\gamma)} \mathbb{E}^{\pi}[d(\omega,\overline{\omega})].$$

E.g. $\ell(x,t) = f(x_t)$ and $\ell(x,t) = f(\sup_{s\leq t} x_s)$ satisfy the above conditions, with $d(\omega,\tilde{\omega}) = \|\omega - \tilde{\omega}\|_\infty$, if f is Lipschitz. In this case

$$0 \leq v^{\mathcal{F}} - v^{\mathcal{G}} \leq K \inf_{\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\gamma)} \mathbb{E}^{\pi}[V_T(\overline{\omega} - \omega)].$$

Optimal stopping

Steps of the proof:

- Fix a causal transport $\pi \in \Pi^{\mathcal{F},\mathcal{G}}(\gamma,\gamma)$.
- Similar idea, but "projecting stopping times" less obvious.
- Randomized stopping time: $\Sigma \in RST(\mathcal{F},\mu)$ is increasing, right-cont., $\mathcal{F}$-adapted, with $\Sigma_0 = 0$ and $\Sigma_T = 1$, μ-a.s.
- For $\Sigma \in RST(\mathcal{G},\gamma)$, let $\widetilde{\Sigma}$ be its $(\pi, \mathcal{F} \otimes \{\emptyset, C\})$-opt. proj.
- By causality: opt. proj. is π-ind. from dual opt. proj.
- Then $\widetilde{\Sigma} \in RST(\mathcal{F},\gamma)$ and

$$\mathbb{E}^{\pi}\left[\int_0^T \ell(\omega,t)d\Sigma_t(\overline{\omega})\right] = \mathbb{E}^{\gamma}\left[\int_0^T \ell(\omega,t)d\tilde{\Sigma}_t(\omega)\right]$$

- Conclude by using $\inf_{\mathcal{F}^W\text{-}st.t.} \mathbb{E}^{\mathbb{P}}[\ell(W,\tau)] = \inf_{\widetilde{\Sigma}\in RST(\mathcal{F},\gamma)} \mathbb{E}^{\gamma}\left[\int \ell_t d\widetilde{\Sigma}_t\right]$ (analogous in the enlarged filtration)