Lectures on optimal transport and applications to economics, statistics and finance

Alfred Galichon (New York University and Sciences Po)

Spring 2018
CFM-Imperial distinguished lectures series, Imperial College, London
Lecture 3, March 15, 2018

Today

- Benamou-Brenier formulation of optimal transport
- Mikami-Thieullen and the Bernstein-Schrodinger problem
- Martingale transport
- Convexification

References

- Villani TOT
- Benamou, Brenier, Num. Math. 2000
- Mikami, Thieullen, SPA 2006
- Carlier, G, COCV 2012
- Tan, Touzi, Ann Proba 2013
- G, Henry-Labordere, Touzi, Ann. Appl. Proba 2014
- Huesmann, Trevisan 2017

Section 1

BENAMOU-BRENIER

- **Fact A.0**: The Benamou-Brenier problem

$$
\begin{aligned}
&\inf \mathbb{E}\left[\int_0^1 l\left(V_t\right) dt\right] \\
&dX_t = V_t dt \\
&X_0 \sim \rho^0,\ X_1 \sim \rho^1
\end{aligned}
$$

has dual

$$
\begin{aligned}
&\sup \int \phi\left(1,.\right) d\rho^1 - \int \phi\left(0,.\right) d\rho^0 \\
&s.t.\ \partial_t \phi + l^*\left(\nabla \phi\right) = 0
\end{aligned}
$$

- **Fact A.1**: once $X_0 \sim \rho^0$ is known, one can recover X_t by its **Eulerian description**
$$dX_t = \nabla l^* \left(\nabla \phi \left(t, X_t \right) \right) dt.$$
or by its **Lagrangian description** $X_t = X_0 + t \nabla l^* \left(\nabla \phi \left(0, X_0 \right) \right)$.
- **Fact A.2**: once $\phi \left(1, . \right)$ is known, one can recover the whole set of dual solutions $\phi \left(t, . \right)$ by the **Hopf-Lax formula**$\phi \left(t, x_t \right) = \sup_{x_1} \left\{ \phi \left(1, x_1 \right) - \left(1 - t \right) l \left(\frac{x_1 - x_t}{1 - t} \right) \right\}$.

- **Fact A.3**: $\phi(0,.)$ and $\phi(1,.)$ are the solutions to the **dual MK problem**

$$\sup \int \phi(1,.)\, d\rho^1 - \int \phi(0,.)\, d\rho^0$$
$$s.t.\ \phi(1,x_1) - \phi(0,x_0) \leq l(x_1 - x_0)$$

while (X_0, X_1) is the solution to the **primal MK problem**

$$\inf \mathbb{E}\left[l(X_1 - X_0)\right]$$
$$s.t.\ X_0 \sim \rho^0,\ X_1 \sim \rho^1$$

- **Fact A.4**: once X_1 is known, then one has

$$dX_t = \frac{X_1 - X_t}{1-t} dt.$$

Section 2

Mikami-Thieullen and the Bernstein-Schrodinger problem

- **Fact B.0**: if we impose an additional diffusion term, the primal Mikami-Thieullen problem

$$\min \mathbb{E}\left[\int_0^1 l\left(V_t\right) dt\right]$$
$$dX_t = V_t dt + \sqrt{T} dW_t$$
$$X_0 \sim \rho^0,\ X_1 \sim \rho^1$$

has dual

$$\sup \int \phi\left(1, .\right) d\rho^1 - \int \phi\left(0, .\right) d\rho^0$$
$$s.t.\ \partial_t \phi + l^*\left(\nabla\phi\right) + \frac{T}{2}\Delta\phi = 0$$

which when $T \to 0$, recovers Benamou-Brenier (fact A.0).

- **Fact B.1**: once $X_0 \sim \rho^0$ is known, one can recover X_t by a **s.d.e. description**

$$dX_t = \nabla l^* \left(\nabla \phi \left(t, X_t \right) \right) dt + \sqrt{T} dW_t,$$

  where W_t is a $\mathbb{P}-$Brownian motion independent from X_0. Again, the Eulerian description is recovered (fact B.1) when $T \to 0$.
- What about extending the other Facts? we will seek to generalize them.

- **Today, we shall take**
$$l(z) = |z|^2/2,$$
so that
$$l^*(p) = |p|^2/2.$$
- Open question (at least for me): does this extend to geodesic distance on a manifold? I should guess so. But I am not sure this extends to more general l.

- In Mikami-Thieullen, ϕ solves the viscous quadratic HJ equation

$$\partial_t \phi + \frac{1}{2} |\nabla \phi|^2 + \frac{T}{2} \Delta \phi = 0$$

- Setting $h = \exp(\phi / T)$, so that $\phi = T \ln h$ (Hopf-Cole transformation) and using $\Delta \ln h = (\Delta h) / h - |\nabla \ln h|^2$, we get

$$T \frac{\partial_t h}{h} + \frac{T^2}{2} |\nabla \ln h|^2 + \frac{T^2}{2} \left(\frac{\Delta h}{h} - |\nabla \ln h|^2 \right) = 0$$

thus the quadratic terms disappear and we are left with the time-reversed heat equation $\partial_t h + \frac{T}{2} \Delta h = 0$, hence $h(t, .) = \exp(\phi(t, .) / T)$ has an explicit representation as a convolution of $h(1, .)$ with the heat kernel:

$$h(t, x_t) = \int_{x1} h(1, x_1) \, \pi^*_{s,1}(x_1 | x_s) \, dx_1.$$

where $\pi^*_{s,1}(x_1 | x_s) = \exp\left(-\frac{(x_1 - x_t)^2}{2T(1-t)} \right) / \left(\sqrt{2\pi T (1-t)} \right)^d$ is the transition kernel of the Brownian motion–or more precisely, scaled by $\sqrt{T}$.

- **Fact B.2**: we get the following representation of $\phi(t,.)$:

$$\phi(t,x_t) = T\log\int_{x_1}\exp\frac{\phi(1,x_1)}{T}\frac{\exp\left(-\frac{(x_1-x_t)^2}{2T(1-t)}\right)}{\left(\sqrt{2\pi T(1-t)}\right)^d}dx_1$$

which tends to Hopf-Lax when $T\to 0$, that is, as we recall:

- *Fact A.2. Once $\phi(1,.)$ is known, one can recover the whole set of dual solutions $\phi(t,.)$ by the Hopf-Lax formula*

$$\phi(t,x_t) = \sup_{x_1}\left\{\phi(1,x_1) - (1-t)\,l\left(\frac{x_1-x_t}{1-t}\right)\right\}.$$

- The heat kernel semigroup replaces the Hamilton-Jacobi semigroup. Probabilistic interpretation: Assume that under $\mathbb{P}^*$, $X_t = X_0 + \sqrt{T}B_t^*$ where B^* is a Brownian motion, then

$$\phi(t,x_t) = T\log\mathbb{E}_{\mathbb{P}^*}\left[\exp\left(\frac{\phi(1,X_1)}{T}\right)|X_t = x_t\right].$$

- **Fact B.3:** As a result, one may obtain $\phi(0,.)$ as a function of $\phi(1,.)$ by

$$\phi(0,x_0) = T \log \int_{x_1} \exp \frac{\phi(1,x_1)}{T} \pi^*_{0,1}(x_1|x_0)\, dx_1$$

and the dual becomes

$$\min_{\phi(1,.)} \left\{ \begin{array}{c} \int T \log \left(\int_{x_1} \exp \frac{\phi(1,x_1)}{T} \pi^*_{0,1}(x_1|x_0)\, dx_1 \right) \rho^0(x_0)\, dx_0 \\ -\int \phi(1,x_1)\, \rho^1(x_1)\, dx_1 \end{array} \right\}$$

- In the $T \to 0$ limit, one recovers Fact A.3: $\phi(1,.)$ is solution to

$$\min_{\phi(1,.)} \left\{ \begin{array}{c} \int \sup_{x_1} \left\{ \phi(1,x_1) - (1-t)\, l\left(\frac{x_1 - x_t}{1-t} \right) \right\} \rho^0(x_0)\, dx_0 \\ -\int \phi(1,x_1)\, \rho^1(x_1)\, dx_1 \end{array} \right\}.$$

- Under $\mathbb{P}^*$, $X_t = X_0 + \sqrt{T} B_t^*$ is a Brownian motion; however the constraint $X_1 \sim \rho^1$ is not met under $\mathbb{P}^*$. To remedy this, consider a change of measure. The new measure $\mathbb{P}$ is such that

$$\frac{d\mathbb{P}}{d\mathbb{P}^*} = \frac{h\left(1, X_1\right)}{h\left(0, X_0\right)} = \xi_1$$

  and look for ξ_1 such that under $\mathbb{P}$, $X_0 \sim \rho^0$, and $X_1 \sim \rho^1$. Let $\xi_t = \mathbb{E}_{\mathbb{P}^0}\left[\xi_1 | t\right] = \frac{h(1,X_t)}{h(0,X_0)}$, so that ξ_t is a $\mathbb{P}^*$-martingale.
- Under $\mathbb{P}$,

$$\left(X_0, X_1\right) \sim \frac{h\left(1, x_1\right)}{h\left(0, x_0\right)} \pi^*\left(x_1 | x_0\right) \rho^0\left(x_0\right)$$

  thus we will look for $h\left(0, x_1\right)$ and $h\left(1, x_1\right)$ such that $\pi\left(x, y\right) = \frac{h(1,x_1)}{h(0,x_0)} \pi^*\left(x_1 | x_0\right) \rho^0\left(x_0\right)$ has margins ρ^0 and ρ^1. More on this later.

- Recall that in MT, the diffusion is $dX_t = T\nabla \ln h\left(t, X_t\right) dt + \sqrt{T} dW_t$. What is $\nabla \ln h\left(t, X_t\right)$?
- Under the new measure $\mathbb{P}$, one has

$$\nabla \ln h\left(t, x_t\right) = \mathbb{E}_{\mathbb{P}}\left[\frac{X_1 - X_t}{T\left(1-t\right)} | X_t = x_t\right]$$

Indeed, $\nabla h\left(t, x_t\right) = \int_{x_1} h\left(1, x_1\right) \left(\frac{x_t - x_1}{T(1-t)}\right) \frac{\exp\left(-\frac{(x_1 - x_t)^2}{2T(1-t)}\right)}{\left(\sqrt{2\pi T(1-t)}\right)^d} dx_1$, thus

$$\begin{aligned}\nabla \log h\left(t, x_t\right) &= \int_{x_1} \frac{h\left(1, x_1\right)}{h\left(t, x_t\right)} \left(\frac{x_t - x_1}{T\left(1-t\right)}\right) \frac{\exp\left(-\frac{\left(x_1 - x_t\right)^2}{2T\left(1-t\right)}\right)}{\left(\sqrt{2\pi T\left(1-t\right)}\right)^d} dx_1 \\ &= \mathbb{E}_{\mathbb{P}}\left[\frac{X_1 - X_t}{T\left(1-t\right)} | X_t = x_t\right].\end{aligned}$$

- **Fact B.4**: once X_1 is known, the one has the *Schrödinger bridge representation*

$$dX_t = \mathbb{E}_{\mathbb{P}}\left[\frac{X_1 - X_t}{1-t}|X_t = x_t\right] dt + \sqrt{T} dW_t$$

which is a generalization of *McCann's interpolant*, recall:

- *Fact A.4: once X_1 is known, then one has*

$$dX_t = \frac{X_1 - X_t}{1-t} dt.$$

- Note that when $X_1 = x_1$ is deterministic, then the Schrödinger bridge boils to a *Brownian bridge*

$$dX_t = \frac{x_1 - X_t}{1-t} dt + \sqrt{T} dW_t$$

- Under $\mathbb{P}^*$, one has $X_0 \sim \rho^0$ and $dX_t = \sqrt{T} dB_t^*$. What is the diffusion of X_t under $\mathbb{P}$? From the Cameron-Martin formula, as

$$\frac{d\mathbb{P}}{d\mathbb{P}^*} = \frac{h(1, X_1)}{h(0, X_0)} = \xi_1$$

we know that $\xi(t, X_t)$ is a $\mathbb{P}^*$-martingale, and $d\xi(t, X_t) = \frac{\nabla h(t,X_t)}{h(0,X_0)} \sqrt{T} dB_t^*$; thus $\frac{d\xi(t,X_t)}{\xi} = \sqrt{T} \nabla \ln h(t, X_t) dB_t^*$, and

$$W_t = B_t^* - \int_0^t \sqrt{T} \nabla \ln h(s, X_s) ds$$

is a $\mathbb{P}$-Brownian motion. As $X_t = X_0 + \sqrt{T} B_t^*$, we get

$$X_t = X_0 + \int_0^t T \nabla \ln h(s, X_s) ds + W_t$$

hence, under $\mathbb{P}$, we recover the s.d.e.

$$dX_t = \nabla \ln h(t, X_t) dt + dW_t.$$

- We are looking for $h(0, x_1)$ and $h(1, x_1)$ so that $\pi(x, y) = \frac{h(1,x_1)}{h(0,x_0)} \pi^*(x_1|x_0) \rho^0(x_0)$ has margins ρ^0 and ρ^1, that is

$$\int \frac{h(1, x_1)}{h(0, x_0)} \pi^*(x_1|x_0) \rho^0(x_0) \, dx_1 = \rho^0(x_1)$$
$$\int \frac{h(1, x_1)}{h(0, x_0)} \pi^*(x_1|x_0) \rho^0(x_0) \, dx_0 = \rho^1(x_1).$$

- This is the Bernstein-Schrödinger problem. Recall from Giulio's presentation that π happens to be the solution of

$$\min_{\pi \in \mathcal{M}(\rho^0, \rho^1)} \mathbb{E}_\pi \left[\ln \frac{\pi(X_0, X_1)}{\pi^0(X_0, X_1)} \right]$$

where $\pi^0(x_0, x_1) = \pi^*(x_1|x_0) \rho^0(x_0)$.

- Let us work out the min-max formulation for this problem. One has

$$\min_{\pi \geq 0} \max_{\phi^0, \phi^1} \left\{ \begin{array}{c} \mathbb{E}_\pi \left[\ln \frac{\pi(X_0, X_1)}{\pi^0(X_0, X_1)} + 1 + \frac{\phi(0, X_0)}{T} - \frac{\phi(1, X_1)}{T} \right] \\ + \int \frac{\phi^1}{T} d\rho^1 - \int \frac{\phi^0}{T} d\rho^0 - 1 \end{array} \right\}$$

thus, by first order conditions

$$\ln \pi \left(x_0, x_1 \right) = \ln \pi^0 \left(x_0, x_1 \right) + \frac{\phi \left(1, x_1 \right) - \phi \left(0, x_0 \right)}{T}$$

$$\pi \left(x_0, x_t \right) = \pi^0 \left(x_0, x_1 \right) \frac{h \left(1, x_1 \right)}{h \left(0, x_0 \right)}$$

- The previous problem is equivalent to the Monge-Kantorovich problem *with an entropic regularization term*

$$\min_{\pi\in\mathcal{M}(\rho^0,\rho^1)} \int \frac{|x_0 - x_1|^2}{2} \pi(x_0, x_1)\, dx_0 dx_1 + T \int \pi(x_0, x_1)\, \pi(x_0, x_1)\, dx_0 dx_1.$$

where $\mathcal{M}(\rho^0, \rho^1)$ is the set of densities of probability measures with marginals P and Q, and $T > 0$ is a temperature parameter.

- Then by first order conditions

$$\pi(x_0, x_1) = \exp\left(\frac{a(x_0) + b(x_1) - |x_1 - x_0|^2 / 2}{T}\right)$$

where the potentials a and b are adjusted so that $\pi \in \mathcal{M}(\rho^0, \rho^1)$.

- Theorem: π is the unique solution of the *Bernstein-Schrödinger system*

$$\left\{ \begin{array}{c} \pi \in \mathcal{M}\left(\rho^0, \rho^1\right) \\ \partial^2_{x_0 x_1} \left\{ \log \pi\left(x_0, x_1\right) + \frac{\left|x_1 - x_0\right|^2}{2T} \right\} = 0 \end{array} \right.$$

  We shall say that π is a σ-optimal coupling between ρ^0 and ρ^1. This allows us to produce a coupling between ρ^0 and ρ^1 which interpolates between optimal transport ($T \to 0$) and independence ($T \to +\infty$).

- Question: is there a stochastic process (X_t) such that:
  - $X_0 \sim \rho^0$ and $X_1 \sim \rho^1$, and
  - for $t > s$, (X_s, X_t) is a $T(t-s)$-optimal coupling between the respective marignals, ρ^s and ρ^t?
- Note that:
  - when $T \to 0$ this becomes a McCann interpolant.
  - when $t \to s$, the dependence between X_s and X_t becomes Gaussian.
- Answer: yes–the Schrödinger bridge!

- Note (cf Pavon) that minimizes $E\left[\int_0^1 \left(\frac{1}{2}\left|V^2\right| + \frac{T^2}{8}\left|\nabla\pi\right|^2\right)\pi dxdt\right]$ s.t. $\partial_t\pi + div\left(\pi V\right) = 0$ and $\rho\left(0,.\right) = ...$, and $\rho\left(1,.\right) = ...$
- Link with Accept/Reject and other MCMC methods
- Graph formulation
- Girsanov on a graph
- Multi-margin case cf. Lions-Rochet's work on multime Hamilton-Jacobi formula.
- Consider another Markov process with kernel π^* and look for π minimizing the KL divergence.

Section 3

MARTINGALE TRANSPORT

Consider a financial market with an asset X_t. Assume:

- X_t is a traded asset, so there is a process (σ_t) and a martingale measure under which $X = X^\sigma$, where

$$\begin{aligned} X_0^\sigma &= x \\ dX_t^\sigma &= \sigma_t dW_t \end{aligned}$$

  where (W_t) is a Brownian motion under the martingale measure.
- There is a complete market of vanilla options at maturity $T = 1$, so the probability distribution Q of X_1 under the martingale measure is given

$$X_1 \sim Q.$$

- The volatility (σ_t) is uncertain, and there is no option market before maturity T that might lead to restrictions on (σ_t). Assume that we need to price an exotic option ξ whose underlying is the wole path $(X_t)_{t\in[0,1]}$. The lower bound on the price of ξ is given by

$$I(Q) = \inf_{(\sigma_t)} \left\{ \mathbb{E}[\xi] : X_0^\sigma = x,\ X_1^\sigma \sim Q \right\}.$$

Assume we want to price an option of maturity $T = 1$ on two underlyings X_1 and Y_1. The payoff of the option at date $T = 1$ is

$$\Phi\left(X_1, Y_1\right)$$

e.g. spread options $\Phi\left(X, Y\right) = \left(X - Y - k\right)^+$; cheapest to deliver $\Phi\left(X, Y\right) = \min\left(X, Y\right)$; etc.

Assume there is a perfectly liquid and complete market of single-name vanilla options on X_1 and Y_1, so that the risk neutral marginal probabilities P of X_1 and Q of Y_1 are known. Let $\mathcal{M}\left(P, Q\right)$ be the set of probabilities with these marginals.

The arbitrage bounds on the option price V are

$$\begin{gathered}\min_{\pi \in \mathcal{M}(P,Q)} \mathbb{E}_\pi\left[\Phi\left(X_1, Y_1\right)\right] \\ \leq V \leq \\ -\min_{\pi \in \mathcal{M}(P,Q)} \mathbb{E}_\pi\left[-\Phi\left(X_1, Y_1\right)\right].\end{gathered}$$

which is a (classical) optimal transport problem.

Consider a European option of payoff $\varphi\left(X_1\right)$ at maturity 1, and assume there is a complete market for call and puts on X_1 at the same maturity. Let $P_0\left(k\right)$ and $C_0\left(k\right)$ the price at time 0 of these options, respectively. Then (Breeden-Litzenberger 1978)

$$\begin{aligned} \mathbb{E}\left[\varphi\left(X\right)\right] &= \varphi\left(x\right) + \varphi'\left(x\right)\left(X_0 - x\right) \\ &+ \int_0^x P_0\left(k\right)\varphi''\left(k\right)dk + \int_k^{+\infty} C_0\left(k\right)\varphi''\left(k\right)dk \end{aligned}$$

hence:

- the vanilla market perfectly determines the risk-neutral probability of X at time 1
- replicating portfolios can be formed directly from vanilla options.

Consider

$$\min_{(X,Y)\sim\pi} \mathbb{E}_\pi \left[\Phi\left(X, Y\right)\right] \tag{P}$$
$$s.t. X \sim P,\ \ Y \sim Q.$$

Dual of this problem is

$$\max_{\varphi_0,\varphi_1} \mathbb{E}\left[\varphi_1\left(Y\right) - \varphi_0\left(X\right)\right] \tag{D}$$
$$s.t. \varphi_1\left(y\right) - \varphi_0\left(x\right) \leq \Phi\left(x, y\right).$$

Weak duality: easy. For $\left(X, Y\right) \sim \pi \in \mathcal{M}\left(P, Q\right)$, and $\left(\varphi_0, \varphi_1\right)$ such that $\varphi_1\left(y\right) - \varphi_0\left(x\right) \leq \Phi\left(x, y\right)$, have

$$\varphi_1\left(Y\right) - \varphi_0\left(X\right) \leq \Phi\left(X, Y\right)$$

and taking expectations,

$$\mathbb{E}\varphi_1\left(Y\right) - \mathbb{E}\varphi_0\left(X\right) \leq \mathbb{E}\Phi\left(X, Y\right)$$

thus value of (D) weakly greater than value of (P). Converse relies on a separation theorem.

- The payoff $\varphi_1(y) - \varphi_0(x)$ is obtained from a subreplicating portfolio composed of vanilla single-name puts and calls. $\mathbb{E}\varphi_1(Y) - \mathbb{E}\varphi_0(X)$ its price.
- Monge-Kantorovich: Price of most expensive superreplicating portfolio=min price of option. Then φ_1 and φ_2 can be taken such that

$$\varphi_1(y) = \inf_x(\varphi_0(x) + \Phi(x,y))$$
$$\varphi_0(x) = \sup_y(\varphi_1(y) - \Phi(x,y))$$

(generalized convex duality).

- Now back to the original problem. Instead of pricing an option on underlying X and Y at time 1, the underlying is now X_1 and X_2, the price of the same asset at two dates forward.
- We still assume that the risk-neutral distributions can be implied from the option prices. The only difference with the previous setting is that we now have the restrictions implies by the fundamental law of asset pricing: under the risk-neutral distribution,

$$E\left[X_2|X_1\right] = X_1$$

- The upper bound on the option price is now

$$\begin{aligned} &\min_{(X_1,X_2)\sim\pi} \mathbb{E}_\pi\left[\Phi\left(X_1,X_2\right)\right] \qquad (1) \\ &s.t. X_1 \sim P,\ X_2 \sim Q,\ E\left[X_2|X_1\right] = X_1 \end{aligned}$$

- The dual problem is

$$\max_{\varphi_0,\varphi_1} \mathbb{E}\left[\varphi_2\left(X_2\right)-\varphi_1\left(X_1\right)\right] \tag{2}$$
$$s.t. \varphi_2\left(x_2\right)-\varphi_1\left(x_1\right)-a\left(x_1\right)\left(x_2-x_1\right) \leq \Phi\left(x_1, x_2\right).$$

- Interpretation: this is the optimal subreplicating portfolio made of vanilla calls and puts on X at maturities 1 and 2, plus rebalancing the quantity of the asset at period 1.

- Consider the minimization of

$$A=\mathbb{E}\left[\int_0^1 L\left(t,X_t,\mu\left(t,X_t\right),\sigma\left(t,X_t\right)\right)dt\right]$$

over the processes X_t and drifts $\mu\left(t,x\right)$ as well as diffusion parameter σ such that

$$dX_t=\mu\left(t,X_t\right)dt+\sigma\left(t,X_t\right)dW_t$$

and $X_0\sim P$, $X_1\sim Q$. For notational convenience, introduce $\Sigma=\sigma\sigma^*/2$, and assume $L\left(t,x,\mu,\Sigma\right)$.

- The value of the problem is given by

$$A=\min_{\mu,\Sigma,p}\int\int_0^1 L\left(t,x,\mu,\Sigma\right)p_t\left(x\right)dtdx \tag{3}$$

subject to

$$p_0=P,\ p_1=Q$$

$$\partial_t p+\nabla.\left(p\mu\right)-\partial^2_{ij}\left(p\Sigma^{ij}\right)=0$$

- The dual problem is

$$\max_{(\varphi_t)} \int \varphi_1 dQ - \int \varphi_0 dP \tag{4}$$
$$\partial_t \varphi + H\left(t, x, \nabla\varphi, D^2\varphi\right) = 0$$

where

$$H(t, x, p, M) = \max_{\mu,\Sigma} \left\{\mu.p + Tr(\Sigma M) - L(t, x, \mu, \Sigma)\right\}.$$

- Further

$$\varphi_t(y) = \inf \left\{ \begin{array}{c} \varphi_0(x) \\ +\mathbb{E}\left[\int_0^t L(t, X_t, \mu(t, X_t), \sigma(t, X_t))\, dt\right] \end{array} \right\}$$
$$s.t. X_1 = y$$
$$dX_t = \mu(t, X_t)\, dt + \sigma(t, X_t)\, dW_t.$$

The saddlefunction for this problem is

$$
\begin{aligned}
&\int \int_0^1 L\left(t,x,\mu,\Sigma\right) p_t\left(x\right) dtdx \\
&+ \int u dp_0 - \int u dP - \int v dp_1 + \int v dQ \\
&+ \int_0^1 \int \varphi_t \left(\partial_t p + \nabla . \left(p\mu\right) - \partial^2_{ij}\left(p\Sigma^{ij}\right)\right) dxdt
\end{aligned}
$$

which is equal to

$$
\begin{array}{c}
\int \varphi_1 dQ - \int \varphi_0 dP \\
- \int \int_0^1 \left(\begin{array}{c} \partial_t \varphi_t + \mu . \nabla \varphi + Tr\left(\Sigma D^2 \varphi\right) \\ -L\left(t,x,\mu,\Sigma\right) \end{array} \right) p_t\left(x\right) dtdx
\end{array}
$$

The minimax formulation of the problem is

$$\max_{\varphi} \min_{p} \left\{ \begin{array}{c} \int \varphi_1 dQ - \int \varphi_0 dP \\ - \int \int_0^1 \left(\partial_t \varphi_t + H\left(t, x, \nabla \varphi, D^2 \varphi\right)\right) p_t\left(x\right) dtdx \end{array} \right\}$$

where

$$H\left(t, x, p, M\right) = \sup_{\mu, \Sigma} \left\{\mu.p + Tr\left(\Sigma M\right) - L\left(t, x, \mu, \Sigma\right)\right\}$$

and one has the following expression for the dual problem

$$A = \max_{(\varphi_t)} \int \varphi_1 dQ - \int \varphi_0 dP$$

$$s.t.\ \partial_t \varphi + H\left(t, x, \nabla \varphi, D^2 \varphi\right) = 0.$$

See Tan and Touzi (2012).

Now assume that one wishes to constrain X_t to be a Markov martingale, i.e. $dX_t = \sigma(t, X_t)\, dW_t$. Then the value of the problem

$$A = \min_{\Sigma, p} \int \int_0^1 L(t, x, \Sigma)\, p_t(x)\, dtdx \tag{5}$$

subject to

$$p_0 = P,\ \ p_1 = Q$$

$$\partial_t p - \partial^2_{ij}\left(p\Sigma^{ij}\right) = 0$$

coincides with its dual formulation, that is

$$\max_{(\varphi_t)} \int \varphi_1 dQ - \int \varphi_0 dP \tag{6}$$

$$\partial_t \varphi + H\left(t, x, D^2\varphi\right) = 0$$

where

$$H(t, x, M) = \max_{\Sigma}\left\{ Tr(\Sigma.M) - L(t, x, \Sigma)\right\}.$$

Section 4

CONVEXIFICATION

Take

$$L(t,x,\Sigma) = 0 \text{ if } Tr(\Sigma) \leq 1$$
$$= +\infty \text{ otherwise.}$$

Then

$$H(t,x,M) = \max_{Tr(\Sigma)\leq 1} \{Tr(\Sigma.M)\}$$
$$= \max(Sp(M), 0)$$

the equation in (6) is

$$\partial_t \varphi + \max\left(Sp\left(D^2\varphi\right), 0\right) = 0$$

thus letting $\psi(t,x) = -\varphi(-t,x)$, the equation becomes

$$\partial_t \psi = \min\left(Sp\left(D^2\psi\right), 0\right) = 0$$

which is the convexification equation of L. Vese (1999), further studied in Carlier and Galichon (2012).