Optimal transportation problem. Applications.

Alexander Kolesnikov

Higher School of Economics

Bonn, Hausdorff Research Institute, 2013

Equilibrium theory

We follow **[CE], [CMcN], [Ekeland]**.
Assume we are given

1. Compact sets $X_0 \subset \mathbb{R}^{d_1}$, $Y_0 \subset \mathbb{R}^{d_2}$, $Z_0 \subset \mathbb{R}^{d_3}$
2. Continuous functions $u(x, z)$, $v(y, z)$ defined on $X_0 \times Z_0$, $Y_0 \times Z_0$. $u(v)$ is assumed to be differentiable w.r.t. x (y).
3. Non-negative finite measures μ_0 on X_0 and ν_0 on Y.

Any point $z \in Z_0$ represents **quality** (or contract types) of indivisidle good, $x \in X_0$ represents consumer type, $y \in Y_0$ represent producer type. μ (ν) is a distribution of consumer (producer) types. To allow for the possibility that some agents choose not to participate, we augment the spaces $X = X_0 \cup \emptyset_X$, $Y = Y_0 \cup \emptyset_Y$, $Z = Z_0 \cup \emptyset_Z$ by including an isolated point in each: a partner $\emptyset_X$ for any unmatched sellers, a partner $\emptyset_Y$ for any unmatched buyers, and the null contract $\emptyset_Z$.

Equilibrium theory

We follow **[CE], [CMcN], [Ekeland]**.
Assume we are given

1. Compact sets $X_0 \subset \mathbb{R}^{d_1}$, $Y_0 \subset \mathbb{R}^{d_2}$, $Z_0 \subset \mathbb{R}^{d_3}$
2. Continuous functions $u(x,z)$, $v(y,z)$ defined on $X_0 \times Z_0$, $Y_0 \times Z_0$. $u(v)$ is assumed to be differentiable w.r.t. x (y).
3. Non-negative finite measures μ_0 on X_0 and ν_0 on Y.

Any point $z \in Z_0$ represents **quality** (or contract types) of indivisidle good, $x \in X_0$ represents consumer type, $y \in Y_0$ represent producer type. μ (ν) is a distribution of consumer (producer) types. To allow for the possibility that some agents choose not to participate, we augment the spaces $X = X_0 \cup \emptyset_X$, $Y = Y_0 \cup \emptyset_Y$, $Z = Z_0 \cup \emptyset_Z$ by including an isolated point in each: a partner $\emptyset_X$ for any unmatched sellers, a partner $\emptyset_Y$ for any unmatched buyers, and the null contract $\emptyset_Z$.

Equilibrium theory

We follow **[CE], [CMcN], [Ekeland]**.
Assume we are given

1. Compact sets $X_0 \subset \mathbb{R}^{d_1}$, $Y_0 \subset \mathbb{R}^{d_2}$, $Z_0 \subset \mathbb{R}^{d_3}$
2. Continuous functions $u(x, z)$, $v(y, z)$ defined on $X_0 \times Z_0$, $Y_0 \times Z_0$. $u(v)$ is assumed to be differentiable w.r.t. x (y).
3. Non-negative finite measures μ_0 on X_0 and ν_0 on Y.

Any point $z \in Z_0$ represents **quality** (or contract types) of indivisidle good, $x \in X_0$ represents consumer type, $y \in Y_0$ represent producer type. μ (ν) is a distribution of consumer (producer) types. To allow for the possibility that some agents choose not to participate, we augment the spaces $X = X_0 \cup \emptyset_X$, $Y = Y_0 \cup \emptyset_Y$, $Z = Z_0 \cup \emptyset_Z$ by including an isolated point in each: a partner $\emptyset_X$ for any unmatched sellers, a partner $\emptyset_Y$ for any unmatched buyers, and the null contract $\emptyset_Z$.

Equilibrium theory

We follow **[CE], [CMcN], [Ekeland]**.
Assume we are given

1. Compact sets $X_0 \subset \mathbb{R}^{d_1}$, $Y_0 \subset \mathbb{R}^{d_2}$, $Z_0 \subset \mathbb{R}^{d_3}$
2. Continuous functions $u(x, z)$, $v(y, z)$ defined on $X_0 \times Z_0$, $Y_0 \times Z_0$. $u(v)$ is assumed to be differentiable w.r.t. x (y).
3. Non-negative finite measures μ_0 on X_0 and ν_0 on Y.

Any point $z \in Z_0$ represents **quality** (or contract types) of indivisidle good, $x \in X_0$ represents consumer type, $y \in Y_0$ represent producer type. μ (ν) is a distribution of consumer (producer) types. To allow for the possibility that some agents choose not to participate, we augment the spaces $X = X_0 \cup \emptyset_X$, $Y = Y_0 \cup \emptyset_Y$, $Z = Z_0 \cup \emptyset_Z$ by including an isolated point in each: a partner $\emptyset_X$ for any unmatched sellers, a partner $\emptyset_Y$ for any unmatched buyers, and the null contract $\emptyset_Z$.

Equilibrium theory

We follow **[CE], [CMcN], [Ekeland]**.
Assume we are given

1. Compact sets $X_0 \subset \mathbb{R}^{d_1}$, $Y_0 \subset \mathbb{R}^{d_2}$, $Z_0 \subset \mathbb{R}^{d_3}$
2. Continuous functions $u(x, z)$, $v(y, z)$ defined on $X_0 \times Z_0$, $Y_0 \times Z_0$. $u(v)$ is assumed to be differentiable w.r.t. x (y).
3. Non-negative finite measures μ_0 on X_0 and ν_0 on Y.

Any point $z \in Z_0$ represents **quality** (or contract types) of indivisidle good, $x \in X_0$ represents consumer type, $y \in Y_0$ represent producer type. μ (ν) is a distribution of consumer (producer) types. To allow for the possibility that some agents choose not to participate, we augment the spaces $X = X_0 \cup \emptyset_X$, $Y = Y_0 \cup \emptyset_Y$, $Z = Z_0 \cup \emptyset_Z$ by including an isolated point in each: a partner $\emptyset_X$ for any unmatched sellers, a partner $\emptyset_Y$ for any unmatched buyers, and the null contract $\emptyset_Z$.

We extend the measures

$$\mu = \mu_0 + (1 + \nu(Y_0))\delta_{\emptyset_X}, \quad \nu = \nu_0 + (1 + \mu(X_0))\delta_{\emptyset_Y}$$

and utility functions

$$u(\emptyset_X, z) = \begin{cases} 0, & z = O_Z, \\ -\infty, & \text{else} \end{cases}, \quad v(\emptyset_Y, z) = \begin{cases} 0, & z = O_Z, \\ +\infty, & \text{else} \end{cases}$$

- Assume we are given a **price system** $p : Z_0 \to \mathbb{R}$
- Consumers of type x maximize in z

$$U(x) = \max_{z \in Z}(u(x, z) - p(z)).$$

- Producers of type x maximize in z

$$V(y) = \max_{z \in Z}(p(z) - v(y, z)).$$

We extend the measures

$$\mu = \mu_0 + (1 + \nu(Y_0))\delta_{\emptyset_X}, \quad \nu = \nu_0 + (1 + \mu(X_0))\delta_{\emptyset_Y}$$

and utility functions

$$u(\emptyset_X, z) = \begin{cases} 0, & z = O_Z, \\ -\infty, & \text{else} \end{cases}, \quad v(\emptyset_Y, z) = \begin{cases} 0, & z = O_Z, \\ +\infty, & \text{else} \end{cases}$$

- Assume we are given a **price system** $p : Z_0 \to \mathbb{R}$
- Consumers of type x maximize in z

$$U(x) = \max_{z \in Z}(u(x, z) - p(z)).$$

- Producers of type x maximize in z

$$V(y) = \max_{z \in Z}(p(z) - v(y, z)).$$

We extend the measures

$$\mu = \mu_0 + (1 + \nu(Y_0))\delta_{\emptyset_X}, \quad \nu = \nu_0 + (1 + \mu(X_0))\delta_{\emptyset_Y}$$

and utility functions

$$u(\emptyset_X, z) = \begin{cases} 0, & z = O_Z, \\ -\infty, & \text{else} \end{cases}, \quad v(\emptyset_Y, z) = \begin{cases} 0, & z = O_Z, \\ +\infty, & \text{else} \end{cases}$$

- Assume we are given a **price system** $p : Z_0 \to \mathbb{R}$
- Consumers of type x maximize in z

$$U(x) = \max_{z \in Z}(u(x, z) - p(z)).$$

- Producers of type x maximize in z

$$V(y) = \max_{z \in Z}(p(z) - v(y, z)).$$

We extend the measures

$$\mu = \mu_0 + (1 + \nu(Y_0))\delta_{\emptyset_X}, \quad \nu = \nu_0 + (1 + \mu(X_0))\delta_{\emptyset_Y}$$

and utility functions

$$u(\emptyset_X, z) = \begin{cases} 0, & z = O_Z, \\ -\infty, & \text{else} \end{cases}, \quad v(\emptyset_Y, z) = \begin{cases} 0, & z = O_Z, \\ +\infty, & \text{else} \end{cases}$$

- Assume we are given a **price system** $p : Z_0 \to \mathbb{R}$
- Consumers of type x maximize in z

$$U(x) = \max_{z \in Z}(u(x,z) - p(z)).$$

- Producers of type x maximize in z

$$V(y) = \max_{z \in Z}(p(z) - v(y,z)).$$

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \operatorname{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \operatorname{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \mathrm{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \mathrm{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \operatorname{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \operatorname{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \mathrm{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \mathrm{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \mathrm{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \mathrm{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

If the value $U(V)$ is strictly positive, consumer buys (producer sells) the good item of the maximizing quality z. Otherwise not participates.

We define

1. **Demand set** $D = \mathrm{argmax}\{u(x,z) - p(z) | z \in Z\}$
2. **Supply set** $S = \mathrm{argmin}\{v(y,z) - p(z) | z \in Z\}$
3. **Demand distribution** $\alpha_{X \times Z}$ is a measure on (X, Z) carried on the graph of D with marginal equals to μ
4. **Supply distribution** $\beta_{Y \times Z}$ is a measure on (Y, Z) carried on the graph of S with marginal equals to ν

Definition

Hedonic equilibrium: a pair (α, p) where α is a measure on $X \times Y \times Z$ such that $Pr_{X\times Z}\alpha$ $(Pr_{Y\times Z}\alpha)$ is a demand (supply) distribution.

Define a **cost function** ("**surplus function**")

$$s(x, y) = \sup_z \big(u(x, z) - v(y, z)\big).$$

Assume that the supremum is attainded for every (x, y) that the set

$$Z(x, y) = \mathrm{argmax}\Big((u(x, z) - v(y, z)|z \in Z\Big)$$

is compact, non-empty, and depends hemicontinuously on (x, y). Then one can find a measurable selection $z_0(x, y)$ of the multivalued mapping

$$(x, y) \to \mathrm{argmax}\Big((u(x, z) - v(y, z)|z \in Z\Big)$$

This problem can be viewed as a continuous **matching problem**: every buyer (seller) looking for a best parther.
Ekeland. There exists an (in general non-unique) equilibrium.
Chiappori, McCann, Nesheim. Existence of equilibrium can be established by methods of the transportation theory.

Theorem

Let γ solve the primal and $(q(x), r(y))$ solve the dual transportation problem for the cost $s(x, y)$

$$\int q d\mu + \int r d\nu \to \min, \quad q(x) + r(y) \geq s(x, y).$$

Then there exists $p(z)$ satisfying

$$\inf_y \Big\{ v(y, z) + r(y) \Big\} \geq p(z) \geq \sup_x \Big\{ u(x, z) - q(x) \Big\}. \qquad (1)$$

With $\gamma \circ (x, y, z_0)^{-1}$ any such p forms an equilibrium pair.

This problem can be viewed as a continuous **matching problem**: every buyer (seller) looking for a best parther.

Ekeland. There exists an (in general non-unique) equilibrium.

Chiappori, McCann, Nesheim. Existence of equilibrium can be established by methods of the transportation theory.

Theorem

Let γ solve the primal and $(q(x), r(y))$ solve the dual transportation problem for the cost $s(x, y)$

$$\int q d\mu + \int r d\nu \to \min, \quad q(x) + r(y) \geq s(x, y).$$

Then there exists $p(z)$ satisfying

$$\inf_y \Big\{ v(y, z) + r(y) \Big\} \geq p(z) \geq \sup_x \Big\{ u(x, z) - q(x) \Big\}. \qquad (1)$$

With $\gamma \circ (x, y, z_0)^{-1}$ any such p forms an equilibrium pair.

Remark: note that the statement of the dual problem is slightly different from the standard formulation (but equivalent).

Proof: The solution (q, r) satisfies

$$q(x) + r(y) \geq s(x, y) \geq u(x, z) - v(y, z), \quad on \;\; X \times Y \times Z.$$

Hence

$$v(y, z) + r(y) \geq u(x, z) - q(x)$$

and $\inf_y \big(v(y, z) + r(y)\big) \geq \sup_x \big(u(x, z) - q(x)\big)$. This implies existence of $p(z)$. In particular, (1) implies that for all (x, y, z)

$$q(x) \geq u(x, z) - p(z), \;\; r(y) \geq p(z) - v(y, z). \tag{2}$$

According to the transportational duality

$$q(x) + r(y) = s(x, y) \tag{3}$$

Remark: note that the statement of the dual problem is slightly different from the standard formulation (but equivalent).

Proof: The solution (q, r) satisfies

$$q(x) + r(y) \geq s(x, y) \geq u(x, z) - v(y, z), \quad on \;\; X \times Y \times Z.$$

Hence

$$v(y, z) + r(y) \geq u(x, z) - q(x)$$

and $\inf_y \big(v(y, z) + r(y)\big) \geq \sup_x \big(u(x, z) - q(x)\big)$. This implies existence of $p(z)$. In particular, (1) implies that for all (x, y, z)

$$q(x) \geq u(x, z) - p(z), \;\; r(y) \geq p(z) - v(y, z). \tag{2}$$

According to the transportational duality

$$q(x) + r(y) = s(x, y) \tag{3}$$

In addition, according to the definition of s there exists a contract $\hat{z}$ such that $s(x,y) = u(x,\hat{z}) - v(y,\hat{z})$ and we have

$$s(x,y) = u(x,\hat{z}) - v(y,\hat{z}). \tag{4}$$

for every (x,y) and $\hat{z} \in Z(x,y)$. Combining (4) and (3) we get

$$v(y,\hat{z}) + r(y) = p(\hat{z}) = u(x,\hat{z}) - q(x). \tag{5}$$

for γ-a.e. (x,y) and $\hat{z} \in Z(x,y)$.

Since the measure γ has marginals μ and ν, it remains to show that $\hat{z}$ maximizes $u(x,\hat{z}) - p(\hat{z})$ and $p(\hat{z}) - v(y,\hat{z})$. But this follows immediately from (2), (5).

In addition, according to the definition of s there exists a contract $\hat{z}$ such that $s(x,y) = u(x,\hat{z}) - v(y,\hat{z})$ and we have

$$s(x,y) = u(x,\hat{z}) - v(y,\hat{z}). \tag{4}$$

for every (x,y) and $\hat{z} \in Z(x,y)$. Combining (4) and (3) we get

$$v(y,\hat{z}) + r(y) = p(\hat{z}) = u(x,\hat{z}) - q(x). \tag{5}$$

for γ-a.e. (x,y) and $\hat{z} \in Z(x,y)$.

Since the measure γ has marginals μ and ν, it remains to show that $\hat{z}$ maximizes $u(x,\hat{z}) - p(\hat{z})$ and $p(\hat{z}) - v(y,\hat{z})$. But this follows immediately from (2), (5).

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i} \pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Generalizations

- **Carlier, Ekeland: [CE]** Generalizations of the hedonic equilibruim/matching problem for the case of several types of participants (more than two).
- **Ghoussoub, Pass:** Multimarginal Monge-Kantorovich problem

We are given

- **spaces** $X_1, \cdots, X_N$
- **cost function** $c(x_1, \cdots, x_N)$
- **probability measures** $\mu_1, \cdots, \mu_N$

Find probability measure π such that

- $Pr_{X_i}\pi = \mu_i$
- The functional $\int c(x_1, \cdots, x_N)\, d\pi$ attains its minimum.

Geometry 1. Monge-Ampère equation

Given two measures on $\mathbb{R}^n$ consider the optimal quadratic transportation mapping

$$T = \nabla\varphi.$$

Assume that T is smooth and the measures have densities

$$\mu = \varrho_\mu dx, \ \nu = \varrho_\nu dx.$$

Then the following change of variables formula must hold:

$$\varrho_\nu(\nabla\varphi) \det D^2\varphi = \varrho_\mu. \tag{6}$$

This is a special case of the **Monge–Ampère equation**.
The regularity problem for the Monge–Ampère equation is highly nontrivial. (see works of A.D. Alexandrov, E. Calabi, A.V. Pogorelov, N.V. Krylov, J. Spruck, L. Caffarelli, L. Nirenberg). In geometry there has been a parallel development of a closed direction — the complex Monge–Ampère equation (Sh.-T. Yau, E. Calabi, T. Aubin). The optimal transportation mappings provide **weak** solutions to Monge-Ampère equation (6).

Inequalities

Optimal transportation technique can be applied for proving many classical inequalities with sharp constants (Brunn-Minkowski, Sobolev, isoperimetric inequality ...).

Example: Talagrand's transportation inequality

Consider the standard Gaussian measure

$$\gamma = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} \, dx$$

and another probability measure $\nu = g \cdot \gamma$.

Let $W_2(\gamma, g \cdot \gamma)$ be the Kantorovich quadratic distance:

$$W_2(\gamma, g \cdot \gamma) = \Big(\inf_{\pi \in \Pi(\gamma,\nu)} \int |x-y|^2 d\pi \Big)^{\frac{1}{2}}.$$

Inequalities

Optimal transportation technique can be applied for proving many classical inequalities with sharp constants (Brunn-Minkowski, Sobolev, isoperimetric inequality ...).

Example: Talagrand's transportation inequality

Consider the standard Gaussian measure

$$\gamma = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} \, dx$$

and another probability measure $\nu = g \cdot \gamma$.

Let $W_2(\gamma, g \cdot \gamma)$ be the Kantorovich quadratic distance:

$$W_2(\gamma, g \cdot \gamma) = \Big(\inf_{\pi \in \Pi(\gamma,\nu)} \int |x-y|^2 d\pi \Big)^{\frac{1}{2}}.$$

Inequalities

Optimal transportation technique can be applied for proving many classical inequalities with sharp constants (Brunn-Minkowski, Sobolev, isoperimetric inequality ...).

Example: Talagrand's transportation inequality

Consider the standard Gaussian measure

$$\gamma = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}} \, dx$$

and another probability measure $\nu = g \cdot \gamma$.

Let $W_2(\gamma, g \cdot \gamma)$ be the Kantorovich quadratic distance:

$$W_2(\gamma, g \cdot \gamma) = \Big(\inf_{\pi \in \Pi(\gamma,\nu)} \int |x - y|^2 d\pi \Big)^{\frac{1}{2}}.$$

Theorem

$$\frac{1}{2}W_2^2(\gamma, g\cdot\gamma) \le \mathrm{Ent}_\gamma g, \tag{7}$$

where $\mathrm{Ent}_\gamma g = \int g \log g \; d\gamma$.

More generally: estimates of the type $W_p^p(\nu,\mu) \le \int F\left(\frac{\nu}{\mu}\right) d\mu$ for some $p \ge 1$ and a function F are called transportation inequalities.

Why we are interested in transportation inequalities?

1. Let μ and $\nu = f\cdot\mu$ be probability measures and let $f > 0$. Then **Kullback-Csiszár-Pinsker** inequality holds
$$\frac{1}{2}\|\mu-\nu\|^2 := \frac{1}{2}\left(\int |f-1|\, d\mu\right)^2 \le \mathrm{Ent}_\mu f.$$
Here $\|\mu-\nu\|$ is the variation distance between μ and ν. This inequality can be regarded as the transportation inequality for a suitable cost function (Bolley–Villani).

Theorem

$$\frac{1}{2} W_2^2(\gamma, g \cdot \gamma) \le \mathrm{Ent}_\gamma g, \qquad (7)$$

where $\mathrm{Ent}_\gamma g = \int g \log g \; d\gamma$.

More generally: estimates of the type $W_p^p(\nu, \mu) \le \int F\left(\frac{\nu}{\mu}\right) d\mu$ for some $p \ge 1$ and a function F are called transportation inequalities.

Why we are interested in transportation inequalities?

1. Let μ and $\nu = f \cdot \mu$ be probability measures and let $f > 0$. Then **Kullback-Csiszár-Pinsker** inequality holds

$$\frac{1}{2}\|\mu - \nu\|^2 := \frac{1}{2}\Big(\int |f - 1| \, d\mu\Big)^2 \le \mathrm{Ent}_\mu f.$$

Here $\|\mu - \nu\|$ is the variation distance between μ and ν. This inequality can be regarded as the transportation inequality for a suitable cost function (Bolley–Villani).

Theorem

$$\frac{1}{2} W_2^2(\gamma, g \cdot \gamma) \leq \mathrm{Ent}_\gamma g, \tag{7}$$

where $\mathrm{Ent}_\gamma g = \int g \log g \; d\gamma$.

More generally: estimates of the type $W_p^p(\nu, \mu) \leq \int F\big(\frac{\nu}{\mu}\big) d\mu$ for some $p \geq 1$ and a function F are called transportation inequalities.

Why we are interested in transportation inequalities?

1. Let μ and $\nu = f \cdot \mu$ be probability measures and let $f > 0$. Then **Kullback-Csiszár-Pinsker** inequality holds

$$\frac{1}{2}\|\mu - \nu\|^2 := \frac{1}{2}\Big(\int |f - 1| \, d\mu\Big)^2 \leq \mathrm{Ent}_\mu f.$$

Here $\|\mu - \nu\|$ is the variation distance between μ and ν. This inequality can be regarded as the transportation inequality for a suitable cost function (Bolley–Villani).

Theorem

$$\frac{1}{2}W_2^2(\gamma, g\cdot\gamma) \leq \mathrm{Ent}_\gamma g, \qquad (7)$$

where $\mathrm{Ent}_\gamma g = \int g \log g \; d\gamma$.

More generally: estimates of the type $W_p^p(\nu,\mu) \leq \int F\left(\frac{\nu}{\mu}\right) d\mu$ for some $p \geq 1$ and a function F are called transportation inequalities.

Why we are interested in transportation inequalities?

1. Let μ and $\nu = f\cdot\mu$ be probability measures and let $f > 0$. Then **Kullback-Csiszár-Pinsker** inequality holds

$$\frac{1}{2}\|\mu-\nu\|^2 := \frac{1}{2}\left(\int |f-1|\, d\mu\right)^2 \leq \mathrm{Ent}_\mu f.$$

Here $\|\mu-\nu\|$ is the variation distance between μ and ν. This inequality can be regarded as the transportation inequality for a suitable cost function (Bolley–Villani).

2. Under additional assumptions they are equivalent to Sobolev-type inequalities (e.g. log-Sobolev inequality).
3. They yield the **concentration of measures** phenomenon.
4. This is a natural way to measure distance between solutions to PDE's.
5. Typical application: analysis on Wiener space.

Lemma

For any positive symmetric $d \times d$ matrix A the following inequality holds

$$\mathrm{Tr} A - d - \log \det A \geq 0. \tag{8}$$

2. Under additional assumptions they are equivalent to Sobolev-type inequalities (e.g. log-Sobolev inequality).
3. They yield the **concentration of measures** phenomenon.
4. This is a natural way to measure distance between solutions to PDE's.
5. Typical application: analysis on Wiener space.

Lemma

For any positive symmetric $d \times d$ matrix A the following inequality holds

$$\mathrm{Tr} A - d - \log \det A \geq 0. \qquad (8)$$

2. Under additional assumptions they are equivalent to Sobolev-type inequalities (e.g. log-Sobolev inequality).
3. They yield the **concentration of measures** phenomenon.
4. This is a natural way to measure distance between solutions to PDE's.
5. Typical application: analysis on Wiener space.

Lemma

For any positive symmetric $d \times d$ matrix A the following inequality holds

$$\mathrm{Tr} A - d - \log \det A \geq 0. \tag{8}$$

2. Under additional assumptions they are equivalent to Sobolev-type inequalities (e.g. log-Sobolev inequality).
3. They yield the **concentration of measures** phenomenon.
4. This is a natural way to measure distance between solutions to PDE's.
5. Typical application: analysis on Wiener space.

Lemma

For any positive symmetric $d \times d$ matrix A the following inequality holds

$$\mathrm{Tr} A - d - \log \det A \geq 0. \tag{8}$$

2. Under additional assumptions they are equivalent to Sobolev-type inequalities (e.g. log-Sobolev inequality).
3. They yield the **concentration of measures** phenomenon.
4. This is a natural way to measure distance between solutions to PDE's.
5. Typical application: analysis on Wiener space.

Lemma

For any positive symmetric $d \times d$ matrix A the following inequality holds

$$\mathrm{Tr} A - d - \log\det A \geq 0. \tag{8}$$

Sketch of the proof: Let us consider the optimal transportation $T = \nabla\varphi$ of the measure γ to the measure $g \cdot \gamma$. The desired inequality is equivalent to the estimate

$$\frac{1}{2}\int_{\mathbb{R}^d} |x - \nabla\varphi(x)|^2 \, \gamma(dx) \le \int_{\mathbb{R}^d} g \ln g \, d\gamma. \qquad (9)$$

For the proof of (7) we consider the change of variables formula

$$-\frac{x^2}{2} = \ln g(\nabla\varphi(x)) - \frac{|\nabla\varphi(x)|^2}{2} + \ln \det D^2\varphi(x).$$

Write it as
$|x - \nabla\varphi(x)|^2/2 = \langle x, x - \nabla\varphi(x)\rangle + \ln g(\nabla\varphi(x)) + \ln \det D^2\varphi(x)$
and integrate with respect to the measure γ. Note that

$$\int \Big(\langle x, x - \nabla\varphi(x)\rangle + \ln \det D^2\varphi(x)\Big) \, d\gamma = \int \Big(d - \Delta\varphi + \ln \det D^2\varphi(x)\Big) \, d\gamma.$$

Finally we get (9) from (8).

Log-Sobolev inequality

Let us consider the optimal transportation $T = \nabla\varphi$ of the measure $f^2 \cdot \gamma$ to the measure γ. Apply again the change of variables formula

$$\ln f^2(x) - |x|^2/2 = -|\nabla\varphi(x)|^2/2 + \ln\det D^2\varphi(x).$$

Let us rewrite the obtained equality in the form

$$\ln f^2(x) = -|x - \nabla\varphi(x)|^2/2 - \langle x, \nabla\varphi(x) - x\rangle + \ln\det D^2\varphi(x)$$

and integrate it with respect to the measure $f^2 \cdot \gamma$, which gives the relationship

$$\int_{\mathbb{R}^d} f^2 \ln f^2 \, d\gamma = -\int_{\mathbb{R}^d} \frac{|x - \nabla\varphi(x)|^2}{2} f^2(x)\,\gamma(dx) - \\ - \int_{\mathbb{R}^d} \langle x, \nabla\varphi(x) - x\rangle f^2(x)\,\gamma(dx) + \int_{\mathbb{R}^d} \ln\det D^2\varphi f^2\, d\gamma.$$

Let us apply integration by parts:

$$
\begin{aligned}
&-\int_{\mathbb{R}^d} \langle x, \nabla\varphi(x) - x\rangle f^2(x)\,\gamma(dx) = \\
&= 2\int_{\mathbb{R}^d} \langle \nabla f(x), \nabla\varphi(x) - x\rangle f(x)\,\gamma(dx) - \int_{\mathbb{R}^d} (\Delta\varphi - d) f^2\, d\gamma.
\end{aligned}
$$

Applying (8) we obtain

$$
\begin{aligned}
\int_{\mathbb{R}^d} f^2 \ln f^2\, d\gamma \le \int_{\mathbb{R}^d} \Big[&-\frac{|x - \nabla\varphi(x)|^2}{2} f^2(x) \\
&+ 2\langle \nabla f(x), \nabla\varphi(x) - x\rangle f(x) \Big]\, d\gamma(dx).
\end{aligned}
$$

Finally, we obtain from the estimate

$$
-\frac{|x - \nabla\varphi(x)|^2}{2} f^2(x) + 2\langle \nabla f(x), \nabla\varphi(x) - x\rangle f(x) \le 2|\nabla f(x)|^2.
$$

Theorem

Logarithmic Sobolev inequality. *For every sufficiently regular function f with $\int f^2 \, d\gamma = 1$ one has*

$$\int f^2 \log f^2 \, d\gamma \leq 2 \int |\nabla f(x)|^2 \, d\gamma.$$

A very famous inequality attributed to L. Gross (1975).

PDE's and variational calculus

Philosophy. The space of measures with the distance W_p is an **infinite-dimensional** manifold.

The Benamou–Brenier formula (viewpoint of fluid mechanics).

The Benamou–Brenier formula interprets the metric $W_p(\mathbb{R}^d)$ on the space of $\mathcal{P}_p(\mathbb{R}^d)$ of probability measures with a finite p-th moment as the Riemmannian **length** on the (infinite-dimensional) manifold.

Let $v_t\colon \mathbb{R}^d \to \mathbb{R}^d$, $t \in [0,1]$ be a family of smooth vector fields (a velocity field) and let $T_t\colon \mathbb{R}^d \to \mathbb{R}^d$ be the family of transformations generated by it according to the equation

$$\frac{dT_t}{dt} = v_t(T_t), \;\; T_0(x) = x.$$

Let us consider the family of measures given by

$$\mu_t = \mu_0 \circ T_t^{-1} = \varrho_t \, dx, \; \mu_0 = \varrho_0 \, dx.$$

It is well-known μ_t satisfies the classical continuity equation

$$\frac{\partial \varrho_t}{\partial t} + \operatorname{div}(\varrho_t \cdot v_t) = 0. \tag{10}$$

The following formula (for $p = 2$) was obtained by Benamou and Brenier:

$$W_p^p(\mu_0, \mu_1) = \inf_v \int_0^1 \int |v_t|^p \, d\mu_t \, dt,$$

where $\mu_t = \varrho_t \; dx$ satisfies (10) and inf is taken over vector fields from some suitable class.

Interpretation: W_p is the **geodesic** distance on the space of measures.

Idea of the proof: inequality $W_p^p(\mu_0,\mu_1) \le \int_0^1 \int |v_t|^p \, d\mu_t \, dt$ follows from the relationships

$$W_p^p(\mu_0,\mu_1) \le \int |T_t(x)-x|^p \varrho_0 dx \le \int \int_0^1 \Big|\frac{d}{dt} T_t(x)\Big|^p \, dt \, \varrho_0 \, dx = \\ = \int \int_0^1 |v_t|^p(T_t) \, dt \, d\mu_0 = \int_0^1 \int |v_t|^p \, d\mu_t \, dt.$$

Equality is attained on $T_t(x) = (1-t)x + tT(x)$, where T is the optimal transportation of μ_0 to μ_1 (for $c = |x-y|^p$).

In particular, **geodesics** have a simple description. Any geodesic $\{\mu_t\}$ has the form $\mu_t = \mu_0 \cdot T_t$, where T is the optimal transportation mapping.

Idea of the proof: inequality $W_p^p(\mu_0,\mu_1) \le \int_0^1 \int |v_t|^p \, d\mu_t \, dt$ follows from the relationships

$$W_p^p(\mu_0,\mu_1) \le \int |T_t(x) - x|^p \varrho_0 dx \le \int \int_0^1 \Big| \frac{d}{dt} T_t(x) \Big|^p \, dt \, \varrho_0 \, dx =$$

$$= \int \int_0^1 |v_t|^p (T_t) \, dt \, d\mu_0 = \int_0^1 \int |v_t|^p \, d\mu_t \, dt.$$

Equality is attained on $T_t(x) = (1-t)x + tT(x)$, where T is the optimal transportation of μ_0 to μ_1 (for $c = |x-y|^p$).

In particular, **geodesics** have a simple description. Any geodesic $\{\mu_t\}$ has the form $\mu_t = \mu_0 \cdot T_t$, where T is the optimal transportation mapping.

Gradient flows. Otto calculus.

A gradient flow in $\mathbb{R}^d$ is a solution $X(t)$ of the equation

$$dX(t)/dt = -\nabla E(X(t)),$$

where E is "energy". We observe that $dE(X(t))/dt = -|\nabla E(X(t))|^2$ (the energy is decreasing).

- It is very helpful to interpret an equation as a gradient flow (existence, uniqueness, rate of convergence, numerical solutions ...).
- **Otto calculus:** a heuristic way to interpret PDE's as gradient flows on the space $\mathcal{W}_2(\mathbb{R}^n)$ (which is $\mathbb{P}_2(\mathbb{R}^n)$ endowed with W_2-distance).
- **Difficulties:** $\mathcal{W}_2(\mathbb{R}^n)$ is a non-linear space, but "manifold". There is a natural infinite-dimensional "Riemannian" structure on this space, but it is very difficult to formalize this approach.
- See, however, the book **[AGS]** of Ambrosio, Gigli, and Savarè. Many things can be formalized...

Gradient flows. Otto calculus.

A gradient flow in $\mathbb{R}^d$ is a solution $X(t)$ of the equation

$$dX(t)/dt = -\nabla E(X(t)),$$

where E is "energy". We observe that
$dE(X(t))/dt = -|\nabla E(X(t))|^2$ (the energy is decreasing).

- It is very helpful to interpret an equation as a gradient flow (existence, uniqueness, rate of convergence, numerical solutions ...).
- **Otto calculus:** a heuristic way to interpret PDE's as gradient flows on the space $\mathcal{W}_2(\mathbb{R}^n)$ (which is $\mathbb{P}_2(\mathbb{R}^n)$ endowed with W_2-distance).
- **Difficulties:** $\mathcal{W}_2(\mathbb{R}^n)$ is a non-linear space, but "manifold". There is a natural infinite-dimensional "Riemannian" structure on this space, but it is very difficult to formalize this approach.
- See, however, the book **[AGS]** of Ambrosio, Gigli, and Savarè. Many things can be formalized...

Gradient flows. Otto calculus.

A gradient flow in $\mathbb{R}^d$ is a solution $X(t)$ of the equation

$$dX(t)/dt = -\nabla E(X(t)),$$

where E is "energy". We observe that $dE(X(t))/dt = -|\nabla E(X(t))|^2$ (the energy is decreasing).

- It is very helpful to interpret an equation as a gradient flow (existence, uniqueness, rate of convergence, numerical solutions ...).
- **Otto calculus:** a heuristic way to interpret PDE's as gradient flows on the space $\mathcal{W}_2(\mathbb{R}^n)$ (which is $\mathbb{P}_2(\mathbb{R}^n)$ endowed with W_2-distance).
- **Difficulties:** $\mathcal{W}_2(\mathbb{R}^n)$ is a non-linear space, but "manifold". There is a natural infinite-dimensional "Riemannian" structure on this space, but it is very difficult to formalize this approach.
- See, however, the book **[AGS]** of Ambrosio, Gigli, and Savarè. Many things can be formalized...

Gradient flows. Otto calculus.

A gradient flow in $\mathbb{R}^d$ is a solution $X(t)$ of the equation

$$dX(t)/dt = -\nabla E(X(t)),$$

where E is "energy". We observe that
$dE(X(t))/dt = -|\nabla E(X(t))|^2$ (the energy is decreasing).

- It is very helpful to interpret an equation as a gradient flow (existence, uniqueness, rate of convergence, numerical solutions ...).
- **Otto calculus:** a heuristic way to interpret PDE's as gradient flows on the space $\mathcal{W}_2(\mathbb{R}^n)$ (which is $\mathbb{P}_2(\mathbb{R}^n)$ endowed with W_2-distance).
- **Difficulties:** $\mathcal{W}_2(\mathbb{R}^n)$ is a non-linear space, but "manifold". There is a natural infinite-dimensional "Riemannian" structure on this space, but it is very difficult to formalize this approach.
- See, however, the book **[AGS]** of Ambrosio, Gigli, and Savarè. Many things can be formalized...

Gradient flows. Otto calculus.

A gradient flow in $\mathbb{R}^d$ is a solution $X(t)$ of the equation

$$dX(t)/dt = -\nabla E(X(t)),$$

where E is "energy". We observe that $dE(X(t))/dt = -|\nabla E(X(t))|^2$ (the energy is decreasing).

- It is very helpful to interpret an equation as a gradient flow (existence, uniqueness, rate of convergence, numerical solutions ...).
- **Otto calculus:** a heuristic way to interpret PDE's as gradient flows on the space $\mathcal{W}_2(\mathbb{R}^n)$ (which is $\mathbb{P}_2(\mathbb{R}^n)$ endowed with W_2-distance).
- **Difficulties:** $\mathcal{W}_2(\mathbb{R}^n)$ is a non-linear space, but "manifold". There is a natural infinite-dimensional "Riemannian" structure on this space, but it is very difficult to formalize this approach.
- See, however, the book **[AGS]** of Ambrosio, Gigli, and Savarè. Many things can be formalized...

Energy functionals and their gradient flows

The heat equation

$$E(\mu) = \int \varrho \log \rho \; dx, \quad \frac{\partial \varrho}{\partial t} = \Delta \varrho.$$

The porous media equation:

$$E(\mu) = \frac{1}{m-1} \int \varrho^m \, dx, \quad \frac{\partial \varrho}{\partial t} = \Delta \varrho^m.$$

The McKean–Vlasov equation:

$$E(\mu) = \frac{1}{2} \int W(x-y)\varrho(x)\varrho(y) \, dxdy, \quad \frac{\partial \varrho}{\partial t} = \mathrm{div}(\varrho \nabla(\varrho * \nabla W)).$$

The quantum drift-diffusion equation:

$$E(\mu) = \int \frac{|\nabla \varrho|^2}{\varrho} \, dx, \quad \frac{\partial \varrho}{\partial t} + 4\mathrm{div}\Big(\varrho \nabla \frac{\Delta \sqrt{\varrho}}{\sqrt{\varrho}}\Big) = 0.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\, dx$ consists of all measures with densities of the form $-\mathrm{div}(\varrho \nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\, dx.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\,dx$ consists of all measures with densities of the form $-\mathrm{div}(\varrho\nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\, dx.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\, dx$ consists of all measures with densities of the form $-\mathrm{div}(\varrho \nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\, dx.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\, dx$ consists of all measures with densities of the form $-\text{div}(\varrho \nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\, dx.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\,dx$ consists of all measures with densities of the form $-\mathrm{div}(\varrho\nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\, dx.$$

To compute the gradients of the functionals we need

- Tangent space
- L^2-structure on this space (Riemannian metric)

F. Otto:

- The tangent space $T\mathcal{P}^2(\mathbb{R}^d)$ at a point $\varrho\,dx$ consists of all measures with densities of the form $-\mathrm{div}(\varrho\nabla u)$, where u is a function.
- Define the Riemannian metric on $T\mathcal{P}^2(\mathbb{R}^d)$ as follows:

$$\int_{\mathbb{R}^d} \langle \nabla u_1, \nabla u_2 \rangle \varrho\,dx.$$

Computation of the gradient of the entropy functional

Let us consider the entropy functional

$$\mathrm{Ent}\colon\ \varrho\, dx \mapsto \int_{\mathbb{R}^d} \varrho \ln \varrho\, dx.$$

Differentiate it along a geodesic $t \mapsto \mu_0 \circ T_t^{-1}$. Let us use the change of varibles formula
$\varrho_t((1-t)x + t\nabla\varphi)\det((1-t)I + tD^2\varphi) = \varrho_0$. Using infinitesimal expansion at $t = 0$ we get

$$\langle \nabla\varrho, \nabla\varphi - x\rangle + \dot{\varrho}|_{t=0} + \varrho(\Delta\varphi - d) = 0, \tag{11}$$

$$\frac{\partial \mathrm{Ent}\varrho_t}{\partial t}\Big|_{t=0} = \frac{d}{dt}\int \log \rho_t(T_t)\rho_0\ dx\Big|_{t=0} = -\int_{\mathbb{R}^d} \varrho(\Delta\varphi - d)\, dx$$
$$= \int_{\mathbb{R}^d} \langle \nabla\varrho, \nabla\psi\rangle\, dx,$$

where $\psi(x) = \varphi(x) - |x|^2/2$.

We observe that according to (11)

$$\frac{\partial \varrho_t}{\partial t}|_{t=0} + \mathrm{div}(\nabla\psi \cdot \varrho_0) = 0.$$

Therefore, the gradient of the functional Ent satisfies the equalities

$$\Big\langle \nabla \mathrm{Ent}(\varrho), \frac{\partial \varrho_t}{\partial t} \Big\rangle|_{t=0} := \frac{\partial \mathrm{Ent}\varrho_t}{\partial t}\Big|_{t=0} = \int_{\mathbb{R}^d} \langle \nabla \varrho_0, \nabla\psi \rangle \, dx$$
$$= \int_{\mathbb{R}^d} \Big\langle \frac{\nabla \varrho_0}{\varrho_0}, \nabla\psi \Big\rangle \varrho_0 \, dx,$$

and $\nabla \mathrm{Ent}(\varrho)|_{t=0}$ is identified with the function $-\mathrm{div}\Big(\varrho_0 \frac{\nabla\varrho_0}{\varrho_0}\Big) = -\Delta\varrho_0$ (see the definition of the Riemannian structure!).
Finally we get, that $\nabla \mathrm{Ent}(\varrho) = -\Delta\varrho$, this corresponds to the heat equation $\dot{u} = \Delta u$.

- It can be shown that the entropy functional is **convex**

$$\frac{d^2}{dt^2}\mathrm{Ent}(\mu_t) \geq 0$$

along geodesics. This leads to numerous applications (Sobolev-type inequalities, existence/uniqueness of solutions etc.).

- Consider the Fokker-Planck equation

$$\frac{\partial \mu_t}{\partial t} = \Delta \mu_t + \mathrm{div}(\nabla V \cdot \mu_t),$$

where $\mu_t = \rho_t \, dx$ is probability measure and $D^2 V \geq K \cdot \mathrm{Id}$. Similarly to example with the entropy one can compute that

$$\frac{d}{dt} W_2^2(\mu_t, \nu_t) \leq -2K \cdot W_2^2(\mu_t, \nu_t).$$

for two different solutions. In particular,

$$W_2(\mu_t, \nu_t) \leq W_2(\mu_0, \nu_0) e^{-Kt}.$$

We have exponential convergence to the invariant measure if $K > 0$.

- It can be shown that the entropy functional is **convex**

$$\frac{d^2}{dt^2}\mathrm{Ent}(\mu_t) \geq 0$$

along geodesics. This leads to numerous applications (Sobolev-type inequalities, existence/uniqueness of solutions etc.).

- Consider the Fokker-Planck equation

$$\frac{\partial \mu_t}{\partial t} = \Delta \mu_t + \mathrm{div}(\nabla V \cdot \mu_t),$$

where $\mu_t = \rho_t \, dx$ is probability measure and $D^2 V \geq K \cdot \mathrm{Id}$. Similarly to example with the entropy one can compute that

$$\frac{d}{dt} W_2^2(\mu_t, \nu_t) \leq -2K \cdot W_2^2(\mu_t, \nu_t).$$

for two different solutions. In particular,

$$W_2(\mu_t, \nu_t) \leq W_2(\mu_0, \nu_0) e^{-Kt}.$$

We have exponential convergence to the invariant measure if $K > 0$.

Geometry 2

Positively curved spaces

Consider **Riemannan manifold with density**, i.e. manifold M with metric g and probability measure $e^{-V}dx$, where dx is the Riemanian volume.
Its **Bakry-Emery tensor**

$$BE = \mathrm{Ric} + D^2 V,$$

where Ric is the Ricci tensor of M and $D^2 V$ is the Riemannian Hessian of V, is responsible for analytical properties of manifolds.

Theorem

(Bakry-Emery (1985)). *Let M be a Riemannian manifold with a probability measure $\mu = e^{-V} dx$, where dx is the Riemannian volume. If $BE \geq K \cdot g$, $K > 0$, then there holds the logarithmic Sobolev inequality $\mathrm{Ent}_{\mu} f^2 \leq \frac{2}{K} \int_M |\nabla f|^2 d\mu$.*

Geometry 2

Positively curved spaces

Consider **Riemannan manifold with density**, i.e. manifold M with metric g and probability measure $e^{-V}dx$, where dx is the Riemanian volume.
Its **Bakry-Emery tensor**

$$BE = \mathrm{Ric} + D^2 V,$$

where Ric is the Ricci tensor of M and $D^2 V$ is the Riemannian Hessian of V, is responsible for analytical properties of manifolds.

Theorem

(Bakry-Emery (1985)). *Let M be a Riemannian manifold with a probability measure $\mu = e^{-V} dx$, where dx is the Riemannian volume. If $BE \geq K \cdot g$, $K > 0$, then there holds the logarithmic Sobolev inequality* $\mathrm{Ent}_\mu f^2 \leq \frac{2}{K} \int_M |\nabla f|^2 d\mu$.

Theorem

(Characterization of the property $BE \geq K \cdot g$.) *A smooth manifold M with a measure $\nu = e^{-V}\,dx$, where $V \in C^2(M)$ satisfies $BE \geq K \cdot g$, precisely when the entropy functional*

$$\mathcal{F}(\mu) = \int_M \varrho \ln \varrho \, d\nu, \quad \mu = \rho \cdot \nu,$$

is K-convex on $\mathcal{W}_2(M)$.

K-convexity means that $\frac{d^2}{dt^2}\mathcal{F}(\mu_t) \geq K$, where $\{\mu_t\}$ is a geodesic family of measures.

This characterization can be used to define positively curved **non-smooth** spaces: Gromov-Hausdorff limits of smooth manifolds, infinite-dimensional spaces etc. (**C. Villani, K.-T. Sturm, J. Lott**)

Theorem

(Characterization of the property $BE \geq K \cdot g$.) *A smooth manifold M with a measure $\nu = e^{-V}\,dx$, where $V \in C^2(M)$ satisfies $BE \geq K \cdot g$, precisely when the entropy functional*

$$\mathcal{F}(\mu) = \int_M \varrho \ln \varrho \, d\nu, \quad \mu = \rho \cdot \nu,$$

is K-convex on $\mathcal{W}_2(M)$.

K-convexity means that $\frac{d^2}{dt^2}\mathcal{F}(\mu_t) \geq K$, where $\{\mu_t\}$ is a geodesic family of measures.

This characterization can be used to define positively curved **non-smooth** spaces: Gromov-Hausdorff limits of smooth manifolds, infinite-dimensional spaces etc. (**C. Villani, K.-T. Sturm, J. Lott**)

Theorem

(Characterization of the property $BE \geq K \cdot g$.**)** *A smooth manifold* M *with a measure* $\nu = e^{-V}\,dx$*, where* $V \in C^2(M)$ *satisfies* $BE \geq K \cdot g$*, precisely when the entropy functional*

$$\mathcal{F}(\mu) = \int_M \varrho \ln \varrho \, d\nu, \quad \mu = \rho \cdot \nu,$$

is K*-convex on* $\mathcal{W}_2(M)$*.*

K-convexity means that $\frac{d^2}{dt^2}\mathcal{F}(\mu_t) \geq K$, where $\{\mu_t\}$ is a geodesic family of measures.

This characterization can be used to define positively curved **non-smooth** spaces: Gromov-Hausdorff limits of smooth manifolds, infinite-dimensional spaces etc. (**C. Villani, K.-T. Sturm, J. Lott**)

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Other applications in geometry

- Generalizations of geometrical (comparison theorems) and analytical (isoperimetric inequalities) results for the case of manifolds with densities
- There are many works on optimal transportation on manifolds (in particular, regularity issues).
- Ricci flows from the transportational viewpoint. It is possible to recover some of the crucial Perelman's estimates using optimal transportation (**Topping, McCann (09)**)
- **Oliker (07)** The cost function $-\log\langle x, y\rangle$ leads to transportational solution of a Minkowski-type problem solved earlier by A.D. Alexandrov
- Weak solutions to Gauss curvature flows (**V. Bogachev, K.**(09))

Infinite-dimensional analysis

- **Analysis on Wiener space.** Let γ be a Gaussian measure and $(H, |\cdot|)$ its Cameron-Martin space.
  Problem. Given a measure $\nu = g \cdot \gamma$ find a solution to the following Monge problem

$$\int |T(x) - x|^2 \, d\gamma \to \inf, \tag{12}$$

  where the infimum is taken over mappings T transforming γ to ν.
- Solution to (12) exists under quite general assumptions (**Feyel-Üstünel; K.**)
  The optimal transportation mapping has the form

$$T(x) = x + \nabla\varphi(x),$$

  where φ satisfied $\partial^2_{hh}\varphi + 1 \geq 0$ for every $h \in H, |h| \leq 1$.
  There is no Φ such that $T = \nabla\Phi$!

Infinite-dimensional analysis

- **Analysis on Wiener space.** Let γ be a Gaussian measure and $(H, |\cdot|)$ its Cameron-Martin space.
  Problem. Given a measure $\nu = g \cdot \gamma$ find a solution to the following Monge problem

$$\int |T(x) - x|^2 \, d\gamma \to \inf, \tag{12}$$

  where the infimum is taken over mappings T transforming γ to ν.

- Solution to (12) exists under quite general assumptions (**Feyel-Üstünel; K.**)
  The optimal transportation mapping has the form

$$T(x) = x + \nabla\varphi(x),$$

  where φ satisfied $\partial^2_{hh}\varphi + 1 \geq 0$ for every $h \in H, |h| \leq 1$.
  There is no Φ such that $T = \nabla\Phi$!

- Change of variables formula (**V. Bogachev, K.**)

$$g = \det_2(\mathrm{I} + D^2\varphi)e^{\mathcal{L}\varphi - \frac{1}{2}|\nabla\varphi|^2},$$

where $\det_2$ is the Fredholm–Carleman determinant (for $A : \mathbb{R}^n \to \mathbb{R}^n$ it can be defined by the formula $\det_2 A := e^{\mathrm{Tr}\,(I-A)} \det A$), $L\varphi$ is the Gaussian "Laplacian" $L\varphi = \Delta\varphi - \langle x, \nabla\varphi \rangle$.

- Sufficient conditions for Sobolev regularity of the solution (**[BK]**)

Theorem

Let $T(x) = x + \nabla\varphi$ be optimal transportation mapping of $g \cdot \gamma$ to γ. Then

$$\mathrm{I}_\gamma g = 2\mathrm{Ent}_\gamma g - 2\int \log\det{}_2(\mathrm{Id} + D^2\varphi)\ g d\gamma \tag{13}$$

$$+ \int \|D^2\varphi\|^2_{HS}\ g d\gamma + \sum_{k=1}^{\infty} \int \mathrm{Tr}\big[(I + D^2\varphi)^{-1} D^2\varphi_{x_k}\big]^2\ g d\gamma,$$

where $\mathrm{I}_\gamma g = \int \frac{|\nabla g|^2}{g} d\gamma$ (relative information), $\mathrm{Ent}_\gamma g = \int g \log g\ d\gamma$ (relative entropy). $\|D^2\varphi\|_{HS}$ is the Hilber-Schmidt norm of $D^2\varphi$.

All the terms in (13) are nonnegative. Theorem contains 1) log-Sobolev inequality, 2) a priori estimate of Sobolev norms of φ.

- Sufficient conditions for Sobolev regularity of the solution (**[BK]**)

Theorem

Let $T(x) = x + \nabla\varphi$ be optimal transportation mapping of $g \cdot \gamma$ to γ. Then

$$\mathrm{I}_\gamma g = 2\mathrm{Ent}_\gamma g - 2\int \log \det_2(\mathrm{Id} + D^2\varphi)\, g d\gamma + \int \|D^2\varphi\|_{HS}^2\, g d\gamma + \sum_{k=1}^{\infty} \int \mathrm{Tr}\big[(I + D^2\varphi)^{-1} D^2\varphi_{x_k}\big]^2\, g d\gamma, \tag{13}$$

where $\mathrm{I}_\gamma g = \int \frac{|\nabla g|^2}{g} d\gamma$ *(relative information),* $\mathrm{Ent}_\gamma g = \int g \log g\ d\gamma$ *(relative entropy).* $\|D^2\varphi\|_{HS}$ *is the Hilber-Schmidt norm of* $D^2\varphi$*.*

All the terms in (13) are nonnegative. Theorem contains 1) log-Sobolev inequality, 2) a priori estimate of Sobolev norms of φ.

- Sufficient conditions for Sobolev regularity of the solution (**[BK]**)

Theorem

Let $T(x) = x + \nabla\varphi$ be optimal transportation mapping of $g \cdot \gamma$ to γ. Then

$$\mathrm{I}_\gamma g = 2\mathrm{Ent}_\gamma g - 2\int \log \det_2(\mathrm{Id} + D^2\varphi)\ g d\gamma \tag{13}$$

$$+ \int \|D^2\varphi\|^2_{HS}\ g d\gamma + \sum_{k=1}^{\infty} \int \mathrm{Tr}\big[(I + D^2\varphi)^{-1} D^2\varphi_{x_k}\big]^2\ g d\gamma,$$

where $\mathrm{I}_\gamma g = \int \frac{|\nabla g|^2}{g} d\gamma$ (relative information), $\mathrm{Ent}_\gamma g = \int g \log g\ d\gamma$ (relative entropy). $\|D^2\varphi\|_{HS}$ is the Hilber-Schmidt norm of $D^2\varphi$.

All the terms in (13) are nonnegative. Theorem contains 1) log-Sobolev inequality, 2) a priori estimate of Sobolev norms of φ.

- Optimal transport of martingale measures (very recent) **Beiglböck, Shachermayer, Soner, Touizi, $\cdots$**
- Optimal transportation for measures which are **mutually singular**. (**[KZ]**)

Problem: when there exists "optimal" transportation for mutual singular measures?

Motivating example: (Transportation exists but it is not a minimizer of some Monge functional)

1) Let $\mu = \prod_{i=1}^{\infty} \mu_i(dx_i)$, $\nu = \prod_{i=1}^{\infty} \nu_i(dx_i)$ be product probability measures. Assume that μ_i, ν_i have densities. Set $T(x) = (T_1(x_1), \cdots, T_i(x_i), \cdots)$, where $T_n(x_n)$ is the one-dimensional optimal transportation sending μ_i to ν_i.

2) γ is a Gaussian measure and μ is obtained from γ by a **linear** mapping $T(x) = Ax$ with A symmetric and positive. The measures γ and μ are mutually singular even in the simplest case $A = 2 \cdot \mathrm{Id}$. T is "optimal" because it is linear and given by a positive symmetric operator. Heuristically, $T(x) = \frac{1}{2}\nabla\langle Ax, x\rangle$.

- Optimal transport of martingale measures (very recent) **Beiglböck, Shachermayer, Soner, Touizi,** $\cdots$
- Optimal transportation for measures which are **mutually singular**. (**[KZ]**)

Problem: when there exists "optimal" transportation for mutual singular measures?
Motivating example: (Transportation exists but it is not a minimizer of some Monge functional)
1) Let $\mu = \prod_{i=1}^{\infty} \mu_i(dx_i)$, $\nu = \prod_{i=1}^{\infty} \nu_i(dx_i)$ be product probability measures. Assume that μ_i, ν_i have densities. Set $T(x) = (T_1(x_1), \cdots, T_i(x_i), \cdots)$, where $T_n(x_n)$ is the one-dimensional optimal transportation sending μ_i to ν_i.
2) γ is a Gaussian measure and μ is obtained from γ by a **linear** mapping $T(x) = Ax$ with A symmetric and positive. The measures γ and μ are mutually singular even in the simplest case $A = 2 \cdot \mathrm{Id}$. T is "optimal" because it is linear and given by a positive symmetric operator. Heuristically, $T(x) = \frac{1}{2}\nabla\langle Ax, x\rangle$.

- Optimal transport of martingale measures (very recent) **Beiglböck, Shachermayer, Soner, Touizi,** $\cdots$
- Optimal transportation for measures which are **mutually singular**. (**[KZ]**)

Problem: when there exists "optimal" transportation for mutual singular measures?

Motivating example: (Transportation exists but it is not a minimizer of some Monge functional)

1) Let $\mu = \prod_{i=1}^{\infty} \mu_i(dx_i)$, $\nu = \prod_{i=1}^{\infty} \nu_i(dx_i)$ be product probability measures. Assume that μ_i, ν_i have densities. Set $T(x) = (T_1(x_1), \cdots, T_i(x_i), \cdots)$, where $T_n(x_n)$ is the one-dimensional optimal transportation sending μ_i to ν_i.

2) γ is a Gaussian measure and μ is obtained from γ by a **linear** mapping $T(x) = Ax$ with A symmetric and positive. The measures γ and μ are mutually singular even in the simplest case $A = 2 \cdot \mathrm{Id}$. T is "optimal" because it is linear and given by a positive symmetric operator. Heuristically, $T(x) = \frac{1}{2}\nabla\langle Ax, x\rangle$.

Counterexample. (Reason for non-existence: absence of ergodicity)
Let $\mu = \gamma$ be the standard Gaussian measure on $\mathbb{R}^\infty$ and

$$\nu = \frac{1}{2}(\gamma + \gamma_2)$$

be the average of γ and its homothetic image $\gamma_2 = \gamma \circ S^{-1}$, where $S(x) = 2x$. There is no any mass transportation T of μ to ν which preserves "rotational invariance" or even exchangeability of a set (i.e., if A is invariant with respect to cylindrical rotations, then $T(A)$ is invariant too). Indeed, any mapping of such a type must have the form $T(x) = g(x)(x_1, x_2, \cdots) = g(x) \cdot x$, where g is invariant with respect to any "rotation", in particular, with respect to any coordinate permutation. But any function g of this type is constant γ-a.e. This is a corollary of the Hewitt–Savage $0-1$ law. It is clear that there is no any mass transportation of this type for the given target measure.

Approach

[KZ]. Consider measures which are invariant w.r.t. some sufficiently rich group (permutations of coordinates, Bernoulli shifts). Instead of minimizing the Monge functional

$$\int \sum_{i=1}^{\infty} |T_i(x) - x_i|^2 \, d\mu$$

minimize

$$\int |T_1(x) - x_1|^2 \, d\mu.$$

These formulations are equivalent for in the finite-dimensional case (due to symmetry).

This problem is far from being well understood. Is it true that transportation exists for a large class of ergodic measures?
A positive result: consider a Gibbs measure μ which can be formally written in the form

$$\mu = e^{-H(x)} dx,$$

where the potential H admits the following heuristic representation:

$$H(x) = \sum_{i=-\infty}^{+\infty} V(x_i) + \sum_{i=-\infty}^{+\infty} W(x_i, x_{i+1}).$$

Under reasonable assumptions there exists optimal transportation of μ onto the standard Gaussian product measure on $\mathbb{R}^\infty$

$$\gamma = \prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} \, dx_i.$$

References

AGS Ambrosio L., Gigli N., Savaré G., Gradient flows in metric spaces and in the Wasserstein spaces of probability measures, Birkhäuser, Basel, 2005.

BK Bogachev V., Kolesnikov A. Sobolev regularity for the Monge-Ampere equation in the Wiener space. http://arxiv.org/abs/1110.1822.

CE Carlier G., Ekeland I., Matching for teams. Econ Theory (2010) 42:397–418

CMcN Chiappori P.A, McCann R.J., Nesheim L.P. , Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. Econ Theory (2010) 42:317–354

Ekeland Ekeland I., Existence, uniqueness and efficiency of equilibrium in hedonic markets with multidimensional types. Econ Theory (2010) 42:275–315.

References

KZ KOLESNIKOV A., ZAEV D. Optimal transportation of processes with infinite Kantorovich distance. Independence and symmetry. http://arxiv.org/abs/1303.7255

Vill VILLANI C., Topics in optimal transportation, Amer. Math. Soc. Providence, Rhode Island, 2003.

Vill2 VILLANI C., Optimal transport, old and new, Springer, New York, 2009.