Multi-marginal optimal transport and applications

Brendan Pass

University of Alberta

May 23, 2013

Brendan Pass Multi-marginal optimal transport and applications

• A large construction company wants to build many houses.

- A large construction company wants to build many houses.
- The company employs many carpenters, electricians, plumbers, etc.

- A large construction company wants to build many houses.
- The company employs many carpenters, electricians, plumbers, etc.
- Need to build *teams*.
 - Each team consists of one tradesperson of each type and is responsible for building one house.

- A large construction company wants to build many houses.
- The company employs many carpenters, electricians, plumbers, etc.
- Need to build *teams*.
 - Each team consists of one tradesperson of each type and is responsible for building one house.
- Different combinations of people work more of less well together.
 - Different potential teams have different efficiencies.

- A large construction company wants to build many houses.
- The company employs many carpenters, electricians, plumbers, etc.
- Need to build *teams*.
 - Each team consists of one tradesperson of each type and is responsible for building one house.
- Different combinations of people work more of less well together.
 - Different potential teams have different efficiencies.
- Want to construct teams to make overall process as efficient as possible.

4 3 5 4



P

æ

·≣ ► < ≣ ►



-

э

Multi-marginal problem: Monge formulation

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Multi-marginal problem: Monge formulation

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Monge Problem:

maximize:

$$\int_{M_1} b(x_1, F_2(x_1), F_3(x_1), \dots, F_m(x_1)) d\mu_1(x_1)$$

among (m-1)-tuples of maps $(F_2, F_3, ..., F_m)$ such that $F_i: M_1 \to M_i$ pushes μ_1 to μ_i .

Multi-marginal problem: Monge formulation

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Monge Problem:

maximize:

$$\int_{M_1} b(x_1, F_2(x_1), F_3(x_1), ..., F_m(x_1)) d\mu_1(x_1)$$

among (m-1)-tuples of maps $(F_2, F_3, ..., F_m)$ such that $F_i : M_1 \to M_i$ pushes μ_1 to μ_i .

4 E 6 4 E 6

F pushes μ_1 to μ_2 .



- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Kantorovich Problem:

maximize

$$\int_{M_1\times M_2\times\ldots\times M_m} b(x_1,x_2,\ldots,x_m) d\gamma(x_1,x_2,\ldots,x_m)$$

among measures γ on $M_1 \times M_2 \times ... \times M_m$ that project to the μ_i .

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Kantorovich Problem:

maximize

$$\int_{M_1\times M_2\times\ldots\times M_m} b(x_1,x_2,\ldots,x_m) d\gamma(x_1,x_2,\ldots,x_m)$$

among measures γ on $M_1 \times M_2 \times ... \times M_m$ that project to the μ_i . A Kantorovich solution γ (or $(X_1, X_2, ..., X_m)$) always exists.

- $M_i \subseteq \mathbb{R}^n$, open and bounded, $i = 1, 2, \dots, m$.
- μ_i Borel probability measures on M_i .
- $b: M_1 \times M_2 \times ... \times M_m \rightarrow \mathbb{R}$ smooth *surplus* function.

Kantorovich Problem:

maximize

$$\int_{M_1\times M_2\times\ldots\times M_m} b(x_1,x_2,\ldots,x_m) d\gamma(x_1,x_2,\ldots,x_m)$$

among measures γ on $M_1 \times M_2 \times ... \times M_m$ that project to the μ_i . A Kantorovich solution γ (or $(X_1, X_2, ..., X_m)$) always exists.

γ projects to μ_i .



æ

・聞き ・ ほき・ ・ ほき

 γ projects to μ_i .



Background on (two marginal) optimal transportation

- Optimal transportation with two marginals (m = 2) is an active and well established area of research.
- Many diverse applications, including: fluid mechanics, cosmology, interacting gases, meteorology, image processing, economics, etc.
- Brenier '87, Gangbo '95, Caffarelli '96, Gangbo-McCann '96, Levin '96: Assume μ₁ << dx₁ and that b is twisted, ie:

$$x_2 \mapsto D_{x_1}b(x_1, x_2)$$
 is injective.

Then γ is concentrated on the graph of a function over x_1 and is unique.

• Example: $b(x_1, x_2) = -|x_1 - x_2|^2$.

4 B K 4 B K

Background on multi-marginal problems: good surpluses

- Multi-marginal problems have many emerging applications, in economics, physics, *m*-monotonicity, image processing, financial math, statistics, etc., but are not well understood.
- For certain *special* surplus functions, the optimal γ is unique and is concentrated on a graph over x₁:
 {(x₁, F₂(x₁), ..., F_m(x₁)}.
- Gangbo-Swiech '98: $b(x_1, x_2, ..., x_m) = -\sum_{i \neq j} |x_i x_j|^2$
- Heinich '02: $b(x_1, x_2, ..., x_m) = h(x_1 + x_2 + ... + x_m)$ where $h : \mathbb{R}^n \to \mathbb{R}$ is strictly convex.
- P '11: Strong second order conditions on *b*. For example, when m = 3, we require, for all $x_1, \overline{x_1} \in M_1$, $x_2 \in M_2$, $x_3, \overline{x_3} \in M_3$, we have:

$$\begin{aligned} D_{x_2x_3}^2 b[D_{x_1x_3}^2 b]^{-1} D_{x_1x_2}^2 b(x_1, x_2, x_3) &- D_{x_2x_2}^2 b(x_1, x_2, x_3) \\ &+ D_{x_2x_2}^2 b(\overline{x_1}, x_2, \overline{x_3}) > 0. \end{aligned}$$

• Kim-P '13: $b(x_1, x_2, ..., x_m) = -\inf_{y \in M} \left[\sum_{i=1}^m d^2(x_i, y) \right]$ on a Riemannian manifold M.

Background on multi-marginal problems: bad surpluses

- For other surplus functions, solutions can be non-unique and have high dimensional support. Examples:
- b(x₁, x₂, ..., x_m) = -∑^m_{i≠j} 1/|x_i-x_j|, arises in density functional theory for Coulombic electronic interactions in quantum physics (Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).
- b(x₁, x₂, ..., x_m) = ∑^m_{i≠j} |x_i x_j|², arises when Coulombic interactions are replaced by repulsive, harmonic oscillator interactions.
- $b(x_1, x_2, ..., x_m) = det(x_1, x_2, ..., x_m)$ (when n = m) Carlier-Nazaret '06.
- $b(x_1, x_2, ..., x_m) = h(x_1 + x_2 + ... + x_m)$, h strictly concave .

- 4 E 6 4 E 6

• Model due to Carlier-Ekeland '10 and Chiappori-McCann-Nesheim '10.

医下子 医

- Model due to Carlier-Ekeland '10 and Chiappori-McCann-Nesheim '10.
- Measure μ₁ represents a distribution of buyer types, looking to buy, say, custom built houses; μ_i (i ≥ 2) represents a distribution of a type of worker needed to build houses (ie, carpenters, plumbers, electricians, etc.)

- Model due to Carlier-Ekeland '10 and Chiappori-McCann-Nesheim '10.
- Measure μ₁ represents a distribution of buyer types, looking to buy, say, custom built houses; μ_i (i ≥ 2) represents a distribution of a type of worker needed to build houses (ie, carpenters, plumbers, electricians, etc.)
- Buyer x₁ has a preference f₁(x₁, z) for a house type z ∈ Z ⊆ ℝⁿ; worker x_i (i ≥ 2) has a preference f_i(x_i, z) to build house of type z.

- 4 B b - 4 B b

- Model due to Carlier-Ekeland '10 and Chiappori-McCann-Nesheim '10.
- Measure μ₁ represents a distribution of buyer types, looking to buy, say, custom built houses; μ_i (i ≥ 2) represents a distribution of a type of worker needed to build houses (ie, carpenters, plumbers, electricians, etc.)
- Buyer x₁ has a preference f₁(x₁, z) for a house type z ∈ Z ⊆ ℝⁿ; worker x_i (i ≥ 2) has a preference f_i(x_i, z) to build house of type z.
- Finding an equilibrium in this market is equivalent to solving an optimal transport problem with surplus

$$b(x_1, x_2, ..., x_m) = \sup_{z \in Z} \sum_{i=1}^m f_i(x_i, z)$$

• Model due to Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).

- Model due to Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).
- Measures μ_i represent particle densities of *m* semi-classical electrons.

- Model due to Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).
- Measures μ_i represent particle densities of *m* semi-classical electrons.
- Electrons are indistinguishable $\rightarrow \mu_i = \mu$.

- Model due to Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).
- Measures μ_i represent particle densities of *m* semi-classical electrons.
- Electrons are indistinguishable $\rightarrow \mu_i = \mu$.
- Given μ, the single particle density, want to find the m-particle density (a measure on R^{nm}) minimizing the total interaction energy.

.

- Model due to Cotar-Friesecke-Kluppelberg '11 and Buttazzo-De Pascale-Gori-Giorgi '12).
- Measures μ_i represent particle densities of *m* semi-classical electrons.
- Electrons are indistinguishable $\rightarrow \mu_i = \mu$.
- Given μ, the single particle density, want to find the m-particle density (a measure on R^{nm}) minimizing the total interaction energy.
- Leads to an optimal transport problem with $b(x_1, x_2, ..., x_m) = -\sum_{i \neq j} \frac{1}{|x_i x_j|}$.

通 と イ ヨ と イ ヨ と

• Take n = 1.

э

글 > - < 글 >

• If
$$\frac{\partial^2 b}{\partial x_1 \partial x_2} < 0$$
 the optimal map is decreasing.

글 > - < 글 >

э

글 > - < 글 >

э

Take n = 1.
For m = 2, if ∂²b/∂x₁∂x₂ > 0, optimal map x₂ = F₂(x₁) is increasing.
If ∂²b/∂x₁∂x₂ < 0 the optimal map is decreasing.
For m = 3, if ∂²b/∂x_i∂x_j > 0, for all i ≠ j ∃! optimal maps x₂ = F₂(x₁), x₃ = F₃(x₁), both increasing (Carlier '03)
Coordinate invariant condition: ∂²b/∂x₁∂x₂[∂²b/∂x₃∂x₁] - 1 ∂²b/∂x₃∂x₁ > 0.

• When m = 2, if $det(D^2_{x_1x_2}b) \neq 0$, solution is concentrated on *n*-dimensional Lipschitz submanifold of the product space (McCann-P-Warren '12)

- When m = 2, if $det(D_{x_1x_2}^2 b) \neq 0$, solution is concentrated on *n*-dimensional Lipschitz submanifold of the product space (McCann-P-Warren '12)
- For m = 3, $D^2_{x_i x_i} b > 0$ does not really make sense.

- When m = 2, if $det(D_{x_1x_2}^2 b) \neq 0$, solution is concentrated on *n*-dimensional Lipschitz submanifold of the product space (McCann-P-Warren '12)
- For m = 3, $D^2_{x_i x_j} b > 0$ does not really make sense.
- $D_{x_i x_j}^2 b$ is a bilinear mapping on the product of tangent spaces $T_{x_i} M_i \times T_{x_j} M_j$.

- When m = 2, if $det(D_{x_1x_2}^2 b) \neq 0$, solution is concentrated on *n*-dimensional Lipschitz submanifold of the product space (McCann-P-Warren '12)
- For m = 3, $D^2_{x_i x_j} b > 0$ does not really make sense.
- $D_{x_i x_j}^2 b$ is a bilinear mapping on the product of tangent spaces $T_{x_i} M_i \times T_{x_j} M_j$.
- $(D^2_{x_1x_2}b)[(D^2_{x_3x_2}b)]^{-1}(D^2_{x_3x_1}b)$ is a bilinear mapping on $T_{x_1}M_1 \times T_{x_1}M_1!$

• When m = 2, if $det(D_{x_1x_2}^2 b) \neq 0$, solution is concentrated on *n*-dimensional Lipschitz submanifold of the product space (McCann-P-Warren '12)

• For m = 3, $D^2_{x_i x_j} b > 0$ does not really make sense.

- $D_{x_i x_j}^2 b$ is a bilinear mapping on the product of tangent spaces $T_{x_i} M_i \times T_{x_j} M_j$.
- $(D^2_{x_1x_2}b)[(D^2_{x_3x_2}b)]^{-1}(D^2_{x_3x_1}b)$ is a bilinear mapping on $T_{x_1}M_1 \times T_{x_1}M_1!$
- $(D^2_{x_1x_2}b)[(D^2_{x_3x_2}b)]^{-1}(D^2_{x_3x_1}b) > 0$ makes sense!

$$G = \begin{bmatrix} 0 & D_{x_1x_2}^2 b & D_{x_1x_3}^2 b \\ D_{x_2x_1}^2 b & 0 & D_{x_2x_3}^2 b \\ D_{x_3x_1}^2 b & D_{x_3x_2}^2 b & 0 \end{bmatrix}$$

э

$$G = \begin{bmatrix} 0 & D_{x_1x_2}^2 b & D_{x_1x_3}^2 b \\ D_{x_2x_1}^2 b & 0 & D_{x_2x_3}^2 b \\ D_{x_3x_1}^2 b & D_{x_3x_2}^2 b & 0 \end{bmatrix}$$

• Let the signature of G be $(\lambda_+, \lambda_-, mn - \lambda_+ - \lambda_-)$.

∃ ► < ∃ ►</p>

3

$$G = \begin{bmatrix} 0 & D_{x_1x_2}^2 b & D_{x_1x_3}^2 b \\ D_{x_2x_1}^2 b & 0 & D_{x_2x_3}^2 b \\ D_{x_3x_1}^2 b & D_{x_3x_2}^2 b & 0 \end{bmatrix}$$

- Let the signature of G be $(\lambda_+, \lambda_-, mn \lambda_+ \lambda_-)$.
- $spt(\gamma)$ is spacelike: $V^T \cdot G \cdot V \ge 0$ for all $V \in T(spt(\gamma))$ (P '11).

$$G = \begin{bmatrix} 0 & D_{x_1x_2}^2 b & D_{x_1x_3}^2 b \\ D_{x_2x_1}^2 b & 0 & D_{x_2x_3}^2 b \\ D_{x_3x_1}^2 b & D_{x_3x_2}^2 b & 0 \end{bmatrix}$$

- Let the signature of G be $(\lambda_+, \lambda_-, mn \lambda_+ \lambda_-)$.
- $spt(\gamma)$ is spacelike: $V^T \cdot G \cdot V \ge 0$ for all $V \in T(spt(\gamma))$ (P '11).
- It's dimension is no more than $mn \lambda_-$.

$$G = \begin{bmatrix} 0 & D_{x_1x_2}^2 b & D_{x_1x_3}^2 b \\ D_{x_2x_1}^2 b & 0 & D_{x_2x_3}^2 b \\ D_{x_3x_1}^2 b & D_{x_3x_2}^2 b & 0 \end{bmatrix}$$

- Let the signature of G be $(\lambda_+, \lambda_-, mn \lambda_+ \lambda_-)$.
- $spt(\gamma)$ is spacelike: $V^T \cdot G \cdot V \ge 0$ for all $V \in T(spt(\gamma))$ (P '11).
- It's dimension is no more than $mn \lambda_-$.
- $mn \lambda_{-} = n$ iff $(D^2_{x_1x_2}b)[(D^2_{x_3x_2}b)]^{-1}(D^2_{x_3x_1}b) > 0.$

• Examples: det
$$(x_1x_2...x_m)$$
, $-\sum_{i\neq j} \frac{1}{|x_i - x_j|}$, $\sum_{i\neq j} |x_i - x_j|^2$.

- Examples: det $(x_1x_2...x_m)$, $-\sum_{i\neq j} \frac{1}{|x_i x_j|}$, $\sum_{i\neq j} |x_i x_j|^2$.
- Optimal measure γ is rotationally symmetric. (see, e.g. Carlier-Nazaret '06)

- Examples: det $(x_1x_2...x_m)$, $-\sum_{i\neq j} \frac{1}{|x_i x_j|}$, $\sum_{i\neq j} |x_i x_j|^2$.
- Optimal measure γ is rotationally symmetric. (see, e.g. Carlier-Nazaret '06)
- If $x, y, z \in \operatorname{spt}(\gamma)$, then

- Examples: det $(x_1x_2...x_m)$, $-\sum_{i\neq j} \frac{1}{|x_i x_j|}$, $\sum_{i\neq j} |x_i x_j|^2$.
- Optimal measure γ is rotationally symmetric. (see, e.g. Carlier-Nazaret '06)
- If $x, y, z \in \operatorname{spt}(\gamma)$, then
- $(x, y, z) \in argmax_{|\bar{x}|=r, |\bar{y}|=s, |\bar{z}|=t} b(\bar{x}, \bar{y}, \bar{z})$
- $(Ax, Ay, Az) \in spt(\gamma)$ for any rotation matrix A.

- Examples: det $(x_1x_2...x_m)$, $-\sum_{i\neq j} \frac{1}{|x_i x_j|}$, $\sum_{i\neq j} |x_i x_j|^2$.
- Optimal measure γ is rotationally symmetric. (see, e.g. Carlier-Nazaret '06)
- If $x, y, z \in \operatorname{spt}(\gamma)$, then
- $(x, y, z) \in argmax_{|\bar{x}|=r, |\bar{y}|=s, |\bar{z}|=t} b(\bar{x}, \bar{y}, \bar{z})$
- $(Ax, Ay, Az) \in spt(\gamma)$ for any rotation matrix A.
- Some rotations fix x but not y, assuming x and y are not co-linear (get non Monge solutions).

- Examples: det $(x_1x_2...x_m)$, $-\sum_{i\neq j} \frac{1}{|x_i x_j|}$, $\sum_{i\neq j} |x_i x_j|^2$.
- Optimal measure γ is rotationally symmetric. (see, e.g. Carlier-Nazaret '06)
- If $x, y, z \in \operatorname{spt}(\gamma)$, then
- $(x, y, z) \in argmax_{|\bar{x}|=r, |\bar{y}|=s, |\bar{z}|=t} b(\bar{x}, \bar{y}, \bar{z})$
- $(Ax, Ay, Az) \in spt(\gamma)$ for any rotation matrix A.
- Some rotations fix x but not y, assuming x and y are not co-linear (get non Monge solutions).
- These rotational directions are extra spacelike directions for G.

• For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.
- For m = 3, these type of results hold for

$$b(x_1, x_2, x_3) = \sup_{z \in Z} \sum_{i=1}^3 f_i(x_i, z)$$

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.
- For m = 3, these type of results hold for

$$b(x_1, x_2, x_3) = \sup_{z \in Z} \sum_{i=1}^3 f_i(x_i, z)$$

• This class includes $-\sum_{i=1}^{3} |x_i - x_j|^2$ (Gangbo-Swiech surplus), $h(x_1 + x_2 + x_3)$, for strictly convex h, (Heinich surplus).

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.
- For m = 3, these type of results hold for

$$b(x_1, x_2, x_3) = \sup_{z \in Z} \sum_{i=1}^3 f_i(x_i, z)$$

- This class includes $-\sum_{i=1}^{3} |x_i x_j|^2$ (Gangbo-Swiech surplus), $h(x_1 + x_2 + x_3)$, for strictly convex h, (Heinich surplus).
- Optimal maps factor through a measure on Z (the *generalized barycenter*) Agueh-Carlier '10.

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.
- For m = 3, these type of results hold for

$$b(x_1, x_2, x_3) = \sup_{z \in Z} \sum_{i=1}^3 f_i(x_i, z)$$

- This class includes $-\sum_{i=1}^{3} |x_i x_j|^2$ (Gangbo-Swiech surplus), $h(x_1 + x_2 + x_3)$, for strictly convex h, (Heinich surplus).
- Optimal maps factor through a measure on Z (the *generalized barycenter*) Agueh-Carlier '10.
- Can easily calculate $(D_{x_1x_2}^2 b)[(D_{x_3x_2}^2 b)]^{-1}(D_{x_3x_1}^2 b) > 0$ (under mild conditions on the f_i).

- 4 E 6 4 E 6

- For which surplus functions is the optimizer concentrated on the graph of a function over *x*₁?
- For m = 2, the twist, injectivity of $x_2 \mapsto D_{x_1}b(x_1, x_2)$, suffices.
- For m = 3, these type of results hold for

$$b(x_1, x_2, x_3) = \sup_{z \in Z} \sum_{i=1}^3 f_i(x_i, z)$$

- This class includes $-\sum_{i=1}^{3} |x_i x_j|^2$ (Gangbo-Swiech surplus), $h(x_1 + x_2 + x_3)$, for strictly convex h, (Heinich surplus).
- Optimal maps factor through a measure on Z (the *generalized barycenter*) Agueh-Carlier '10.
- Can easily calculate $(D_{x_1x_2}^2 b)[(D_{x_3x_2}^2 b)]^{-1}(D_{x_3x_1}^2 b) > 0$ (under mild conditions on the f_i).
- One can also prove Monge solutions and uniqueness under strong differential conditions on b (P '11), or under a twist like condition on special sets (Kim-P (in preparation)).

- In the limit as $m \to \infty$, the differences become even more pronounced.
- For the surplus $-\int_0^1 \int_0^1 |x_s x_t|^2 dst dt$, we get unique Monge type solutions (P '13).
- For $-\lim_{m\to\infty} \frac{1}{\binom{m}{2}} \sum_{i\neq j}^{m} \frac{1}{|x_i x_j|}$; the (unique) optimal measure is product measure (Cotar-Friesecke-P (in preparation)).