Control of ensembles via structured optimal transport

Axel Ringh

Department of Electronic and Computer Engineering,
The Hong Kong University of Science and Technology, Hong Kong SAR, China.

18th of May, 2021
Complex System Working Lunch - CoSy
Uppsala University

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Acknowledgements

This is based on joint work with Yongxin Chen[1], Isabel Haasler[2], and Johan Karlsson[2].

[1] Isabel Haasler, Johan Karlsson, and Axel Ringh. Control and estimation of ensembles via structured optimal transport: A computational approach based on entropy-regularized multi-marginal optimal transport. *IEEE Control Systems Magazine*, In press, 2021

[2] Isabel Haasler, Axel Ringh, Yongxin Chen, and Johan Karlsson. Multi-marginal optimal transport with tree-structured cost and the Schrödinger bridge problem. Accepted to *SIAM Journal on Control and Optimization*, 2021.

[3] Axel Ringh, Isabel Haasler, Yongxin Chen, and Johan Karlsson. Efficient computations of multi-species mean field games via graph-structured optimal transport. *Submitted*, 2021.

[4] Isabel Haasler, Axel Ringh, Yongxin Chen, and Johan Karlsson. Scalable computation of dynamic flow problems via multi-marginal graph-structured optimal transport. *In preparation*, 2021.

Financial support from

- Knut and Alice Wallenberg Foundation (KAW)
- Swedish Research Council (VR)
- KTH Digital Futures
- National Science Foundation (NSF)

[1] School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

[2] Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden

Outline

- Introduction: a matching problem
- Optimal transport: discrete and continuous formulation
- Matching, minimum energy control, and optimal steering of a swarm
- Numerical solution method: Sinkhorn iterations

Introduction: A matching problem

- Consider two sets of N agents $x^{(0)} := \{x_i^{(0)}\}_{i=1}^N$ and $x^{(1)} := \{x_j^{(1)}\}_{j=1}^N$ in the state space $\mathbb{R}^d$.
- Can define a distance between the two sets as the solution to a matching (assignment) problem

$$d(x^{(0)}, x^{(1)}) := \underset{\phi \in \text{perm}(\{1,2\ldots,N\})}{\text{minimize}} \sum_{i=1}^{N} \|x_i^{(0)} - x_{\phi(i)}^{(1)}\|_2^2.$$

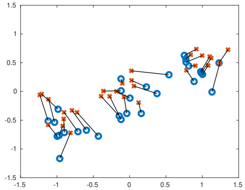


Figure: Example of an optimal association between two distributions of agents.

- Drawback: combinatorial problem! Intractable even for modest N.

Reformulate the problem:

- Parametrize the permutation in terms of the permutation matrix: $M = [M_{ij}]_{i,j=1}^{N}$ with

$$M_{ij} = \begin{cases} 1, & \text{if agent } x_i^{(0)} \text{ and agent } x_j^{(1)} \text{ are associated,} \\ 0, & \text{otherwise.} \end{cases}$$

- Introduce the cost matrix $C_{x^{(0)},x^{(1)}} = [C_{ij}]$, where $C_{ij} = \|x_i^{(0)} - x_j^{(1)}\|_2^2$.
- The optimization problem for $d(x^{(0)}, x^{(1)})$ can be rewritten as

$$\begin{aligned} \underset{M\in\{0,1\}^{N\times N}}{\text{minimize}} \quad & \sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij}M_{ij} \\ \text{subject to} \quad & \sum_{j=1}^{N} M_{ij} = 1,\ i = 1,\dots,N \\ & \sum_{i=1}^{N} M_{ij} = 1\ j = 1,\dots,N \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \underset{M\in\{0,1\}^{N\times N}}{\text{minimize}} \quad & \langle C_{x^{(0)},x^{(1)}}, M\rangle \\ \text{subject to} \quad & M\mathbf{1} = \mathbf{1} \\ & M^T\mathbf{1} = \mathbf{1}. \end{aligned}$$

- **Note:** M is a doubly stochastic matrix.

Introduction: A matching problem

- Relax the variables $M_{ij} \in \{0, 1\}$ to $0 \leq M_{ij} \leq 1$.
- Birkhoff's theorem: permutation matrices are the extreme points of the polytope of doubly stochastic matrices [1].
- All linear programs with an optimal solution has an optimal solution which is an extreme point of the constraint polytope [2].

Theorem

The optimization problems

$$\underset{\phi \in \mathrm{perm}(\{1,2\ldots,N\})}{\text{minimize}} \sum_{i=1}^{N} \|x_i^{(0)} - x_{\phi(i)}^{(1)}\|_2^2 \qquad \textit{and} \qquad \begin{aligned} \underset{M \in [0,1]^{N \times N}}{\text{minimize}} \quad & \langle C_{x^{(0)}, x^{(1)}}, M \rangle \\ \text{subject to} \quad & M\mathbf{1} = \mathbf{1} \\ & M^T\mathbf{1} = \mathbf{1} \end{aligned}$$

attain the same optimal value, and any optimal matrix $\hat{M}$ which is an extreme point of the constraints corresponds to an optimal permutation $\hat{\phi}$.

Note: solving the linear program with the simplex method guarantees that any obtained optimal solution is an optimal permutation matrix.

Note: the linear program is a special instance of a discrete optimal transport problem.

[1] G. Birkhoff, Tres observaciones sobre el algebra lineal, Univ. Nac. Tucuman, Ser. A, vol. 5, pp. 147154, 1946.
[2] D. Luenberger and Y. Ye, Linear and nonlinear programming. Cham: Springer, 2016.

The discrete optimal transport problem is given by

$$\begin{aligned} \underset{M \in \mathbb{R}_+^{N \times N}}{\text{minimize}} \quad & \langle C, M \rangle \\ \text{subject to} \quad & M\mathbf{1} = \mu_0 \\ & M^T\mathbf{1} = \mu_1. \end{aligned}$$

where

- $\mu_0, \mu_1 \in \mathbb{R}_+^N$ are given marginals; must have same total mass: $\mathbf{1}^T \mu_0 = \mathbf{1}^T \mu_1$.
- $C = [C_{ij}]_{i,j=1}^N$ is a cost matrix; C_{ij} = the cost of moving mass from point i to point j.
- $M = [M_{ij}]_{i,j=1}^N$ is the transport plan; M_{ij} = the mass moved from point i to point j.

Interpretation: a classical minimum-cost flow problem on a complete bipartite graph

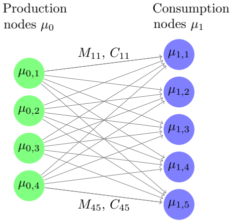


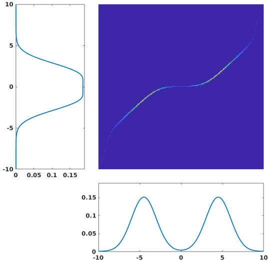
Figure: Small minimum-cost flow problem.

The discrete formulation can be generalized to a continuous setting:

$$\begin{aligned} \underset{m\in\mathcal{M}_+(X\times X)}{\text{minimize}} \quad & \int_{X\times X} c(x_0,x_1)m(x_0,x_1)dx_0dx_1 \\ \text{subject to} \quad & \int_X m(x_0,x_1)dx_1 = \mu_0(x_0), \\ & \int_X m(x_0,x_1)dx_0 = \mu_1(x_1). \end{aligned}$$

where

- $X \subset \mathbb{R}^d$ is the state space.
- $\mathcal{M}_+(\cdot) =$ the set of nonnegative distributions.
- $\mu_0,\ \mu_1 \in \mathcal{M}_+(X)$ are given marginals; must have same total mass: $\int_X \mu_0(x)dx = \int_X \mu_1(x)dx$.
- $c(\cdot,\cdot)$ is a continuous cost function; $c(x_0,x_1) =$ cost of moving mass from x_0 to x_1.
- $m(\cdot,\cdot) \in \mathcal{M}_+(X\times X)$ is the transport plan; $m(x_0,x_1) =$ the mass moved from x_0 to x_1.



This is the so-called Kantorovich formulation of the optimal transport problem.

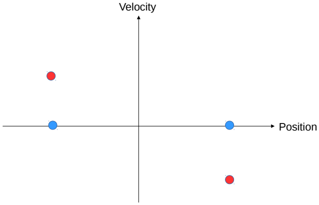
The discrete problem can be seen as a specific instance where, for $k=1,2$,

$$\mu_k(x_k) = \sum_{i=1}^{N} \mu_{k,i}\delta(x-x_i)$$

Matching and minimum energy control

Returning to the matching problem

$$\underset{\phi \in \text{perm}(\{1,2\ldots,N\})}{\text{minimize}} \sum_{i=1}^{N} \|x_i^{(0)} - x_{\phi(i)}^{(1)}\|_2^2.$$



- If X is a state space, distance $\|x_i^{(0)} - x_j^{(1)}\|_2^2$ might not be "relevant".
- Can the dynamics be included in the formulation?

Note: $\|x_i^{(0)} - x_j^{(1)}\|_2^2$ is the geodesic distance in Euclidean space.

Can be formulated as

$$\|x_i^{(0)} - x_j^{(1)}\|_2^2 = \begin{array}{rl} \underset{\gamma \in C([0,1])}{\text{minimize}} & \int_0^1 \|\dot{\gamma}(t)\|_2^2 dt \\ \text{subject to} & \gamma(0) = x_i^{(0)}, \quad \gamma(1) = x_j^{(1)} \end{array} = \begin{array}{rl} \underset{u \in L_2([0,1])}{\text{minimize}} & \int_0^1 \|u(t)\|_2^2 dt \\ \text{subject to} & \dot{x}(t) = u(t), \\ & x(0) = x_i^{(0)}, \quad x(1) = x_j^{(1)}. \end{array}$$

This is a minimum energy control problem for d independent integrators!

Introducing underlying dynamics in the matching problem

This means that the matching problem can be rewritten as

$$\underset{\phi\in\text{perm}(\{1,2\dots,N\})}{\text{minimize}} \sum_{i=1}^{N} \|x_i^{(0)} - x_{\phi(i)}^{(1)}\|_2^2 \quad = \quad \begin{aligned} \underset{\substack{\phi\in\text{perm}(\{1,2\dots,N\})\\ u_i\in L_2([0,1])}}{\text{minimize}} \quad & \sum_{i=1}^{N}\int_0^1 \|u_i(t)\|_2^2 dt \\ \text{subject to} \quad & \dot{x}_i(t) = u_i(t), \\ & x_i(0) = x_i^{(0)}, \quad x_i(1) = x_{\phi(i)}^{(1)}. \end{aligned}$$

Now we can introduce dynamics:

$$\begin{aligned} \underset{\substack{\phi\in\text{perm}(\{1,2\dots,N\})\\ u_i\in L_2([0,1])}}{\text{minimize}} \quad & \sum_{i=1}^{N}\int_0^1 \|u_i(t)\|_2^2 dt \\ \text{subject to} \quad & \dot{x}_i(t) = Ax_i(t) + Bu_i(t), \\ & x_i(0) = x_i^{(0)}, \quad x_i(1) = x_{\phi(i)}^{(1)}. \end{aligned} \quad = \quad \underset{\phi\in\text{perm}(\{1,2\dots,N\})}{\text{minimize}} \sum_{i=1}^{N} c^{A,B}(x_i^{(0)}, x_{\phi(i)}^{(1)})$$

where $c^{A,B}$ is the "control energy distance"

$$\begin{aligned} c^{A,B}(x_0, x_1) \; = \; \underset{u\in L_2([0,1])}{\text{minimize}} \quad & \int_0^1 \|u(t)\|_2^2 dt \\ \text{subject to} \quad & \dot{x}(t) = Ax(t) + Bu(t), \\ & x(0) = x_0, \quad x(1) = x_1. \end{aligned}$$

Dynamics can be included by changing the cost function!

Introducing underlying dynamics in the matching problem

The "control energy distance" has a closed form expression:

$$c^{A,B}(x_0, x_1) = (x_1 - \text{expm}(A)x_0)^T \Sigma_{A,B}^{-1}(x_1 - \text{expm}(A)x_0)$$

where $\Sigma_{A,B}$ is the controllability Gramian $\Sigma_{A,B} = \int_0^1 \text{expm}(A(1-s))BB^T\text{expm}(A^T(1-s))ds$.

The arguments for rewriting the combinatorial problem as a linear program still hold!

- Parametrize ϕ with permutation matrix $M \in \{0,1\}^{N\times N}$.
- Introduce cost $C^{A,B}_{x^{(0)},x^{(1)}} = [C^{A,B}_{ij}]_{i,j=1}^N$, where $C^{A,B}_{ij} = c^{A,B}(x_i^{(0)}, x_j^{(1)})$.
- Write as $\underset{M\in\{0,1\}^{N\times N}}{\text{minimize}} \langle C^{A,B}_{x^{(0)},x^{(1)}}, M\rangle$ subject to $M\mathbf{1} = \mathbf{1}$ and $M^T\mathbf{1} = \mathbf{1}$.
- Relax $M \in \{0,1\}^{N\times N}$ to $M \in [0,1]^{N\times N}$ and use Birkhoff's theorem.

.

$\rightsquigarrow$ Can solve the problem by solving $\underset{M\in[0,1]^{N\times N}}{\text{minimize}} \langle C^{A,B}_{x^{(0)},x^{(1)}}, M\rangle$ subject to $M\mathbf{1} = \mathbf{1}$ and $M^T\mathbf{1} = \mathbf{1}$.

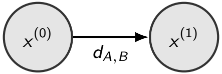
Note: it holds for general nonlinear dynamics, as long as $c(x_0, x_1)$ can be computed effectively.

$$\begin{aligned} c(x_0,x_1) \; = \; & \underset{u\in L_2([0,1])}{\text{minimize}} && \int_0^1 \|u(t)\|_2^2 dt \\ & \text{subject to} && \dot{x}(t) = f(x,u), \\ &&& x(0) = x_0, \quad x(1) = x_1. \end{aligned}$$

Optimal steering of a swarm based on the matching problem

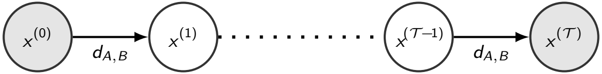
$$\begin{aligned} d_{A,B}(x^{(0)}, x^{(1)}) \; := \; \underset{\substack{\phi \in \mathrm{perm}(\{1,2\dots,N\}) \\ u_i \in L_2([0,1])}}{\text{minimize}} \quad & \sum_{i=1}^{N} \int_0^1 \|u_i(t)\|_2^2 dt & = \underset{M \in [0,1]^{N \times N}}{\text{minimize}} \quad & \langle C^{A,B}_{x^{(0)},x^{(1)}}, M \rangle \\ \text{subject to} \quad & \dot{x}_i(t) = A x_i(t) + B u_i(t), & \text{subject to} \quad & M\mathbf{1} = \mathbf{1} \\ & x_i(0) = x_i^{(0)}, \quad x_i(1) = x^{(1)}_{\phi(i)} & & M^T \mathbf{1} = \mathbf{1}. \end{aligned}$$

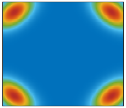
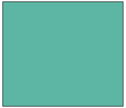
We illustrate such problems graphically as below, where grey circles means known agent distributions.



Next step: can we solve control problems of type $\underset{x^{(i)}}{\text{minimize}} \sum_{i=0}^{\mathcal{T}-1} d_{A,B}(x^{(i)}, x^{(i+1)})$ with given $x^{(0)}$, $x^{(\mathcal{T})}$?

Illustrated with the below graph, where white circles means unknown distributions to be optimized over.



	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	

Optimal steering of a swarm based on the matching problem
Relaxation to convex problem

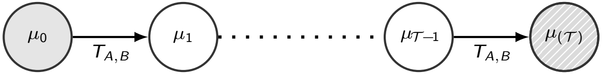
Issue: $d_{A,B}(x^{(0)}, x^{(1)})$ is not convex in the agent positions! (To see this: $C^{A,B}_{ij} = c^{A,B}(x^{(0)}_i, x^{(1)}_j)$).
Remedy: the optimal transport interpretation!

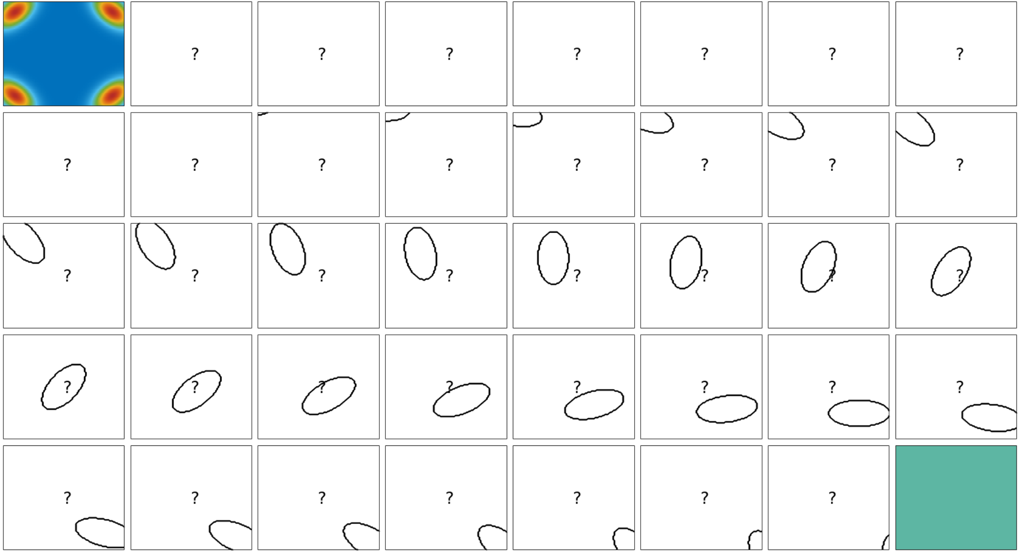
$$d_{A,B}(x^{(0)}, x^{(1)}) = \begin{array}{ll} \underset{M\in[0,1]^{N\times N}}{\text{minimize}} & \langle C^{A,B}_{x^{(0)},x^{(1)}}, M\rangle \\ \text{subject to} & M\mathbf{1} = \mathbf{1} \\ & M^T\mathbf{1} = \mathbf{1} \end{array} \rightsquigarrow \begin{array}{ll} \underset{m\in\mathcal{M}_+(X\times X)}{\text{minimize}} & \displaystyle\int_{X\times X} c^{A,B}(x_0, x_1)m(x_0, x_1)dx_0 dx_1 =: T_{A,B}(\mu_0, \mu_1) \\ \text{subject to} & \displaystyle\int_X m(x_0, x_1)dx_1 = \mu_0(x_0), \\ & \displaystyle\int_X m(x_0, x_1)dx_0 = \mu_1(x_1). \end{array}$$

Relax it to a general density control problem.

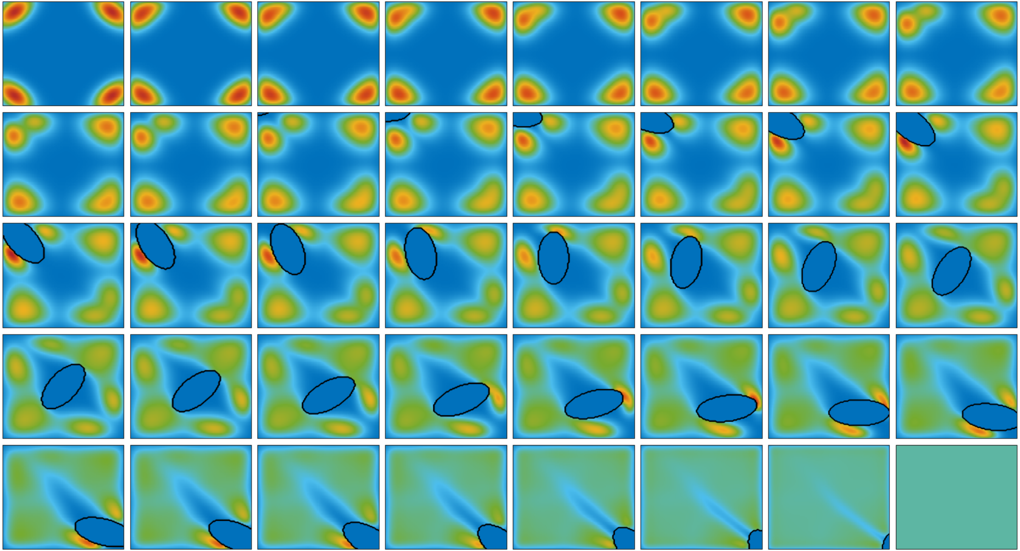
- Corresponds to having two densities of infinitesimal agents, each with dynamics $\dot{x} = Ax + Bu$.
- The optimal transport problem $T_{A,B}(\mu_0, \mu_1)$ is convex in the marginals!

Change $\underset{x^{(i)}}{\text{minimize}} \sum_{i=0}^{\mathcal{T}-1} d_{A,B}(x^{(i)}, x^{(i+1)})$ for the density control problem $\underset{\mu_i}{\text{minimize}} \sum_{i=0}^{\mathcal{T}-1} T_{A,B}(\mu_i, \mu_{i+1})$.
Illustrated with the below graph, where white circles means unknown distributions to be optimized over.
We can also introduce constraints on the unknown marginals.





Optimal steering of a swarm based on the matching problem
Example with constraints



To solve $\underset{\mu_i}{\text{minimize}} \sum_{i=0}^{\mathcal{T}-1} T_{A,B}(\mu_i, \mu_{i+1})$ numerically, we must first understand how to solve $T_{A,B}(\mu_0, \mu_1)$.

Answer: Discretize the state space and solve discrete problem!

State space $\mathbb{R}^d$ with each dimension discretized into $\mathcal{N}$ grid points $\rightsquigarrow$ $\mathcal{N}^d$ grid points $(y_i)_{i=1}^{\mathcal{N}^d}$

- $\rightsquigarrow$ marginals μ_0 and μ_1 are vectors in $\mathcal{N}^d$.
- $\rightsquigarrow$ cost matrix $C^{A,B} \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d} = \mathbb{R}_+^{\mathcal{N}^{2d}}$, with elements $C_{ij}^{A,B} = c^{A,B}(y_i, y_j)$.
- $\rightsquigarrow$ transport plan M with $\mathcal{N}^{2d}$ variables!
  An explosion in the number of variables!
  Even for small numbers like $d = 2$ and $\mathcal{N} = 100$, this gives $100^4 = 10^8 =$ 100 million variables

$$\begin{aligned} \underset{M \in [0,1]^{N \times N}}{\text{minimize}} \quad & \langle C^{A,B}_{x(0),x(1)}, M \rangle \\ \text{subject to} \quad & M\mathbf{1} = \mathbf{1} \\ & M^T\mathbf{1} = \mathbf{1} \end{aligned}$$

- $+$ N^2 varaibles, where N is number of agents; independent of dimension of state space.
- $-$ Not convex in agent position.

$$\begin{aligned} \underset{M \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d}}{\text{minimize}} \quad & \langle C^{A,B}, M \rangle \\ \text{subject to} \quad & M\mathbf{1} = \mu_0 \\ & M^T\mathbf{1} = \mu_1 \end{aligned}$$

- $-$ $\mathcal{N}^{2d} \gg N^2$ variables; huge, and exponential in the dimension of state space.
- $+$ Convex in agent position.

Recently proposed to solve via entropy regularization [1]: $KL(M \mid \mathbf{1}\mathbf{1}^T) = \sum_{i,j=1}^{\mathcal{N}^d} (M_{ij}\log(M_{ij}) - M_{ij} + 1)$,

$$\begin{aligned} \min_{M \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d}} \quad & \langle C, M \rangle + \epsilon KL(M \mid \mathbf{1}\mathbf{1}^T) \\ \text{subject to} \quad & M\mathbf{1} = \mu_0 \\ & M^T\mathbf{1} = \mu_1. \end{aligned}$$

- Let $\exp(\,\cdot\,)$, $\log(\,\cdot\,)$, $./$, $\odot$ denotes the element-wise function.
- For $K = \exp(-C/\epsilon)$, the solution is of the form

$$M = \mathrm{diag}(u)K\mathrm{diag}(v).$$

Theorem (Sinkhorn iterations [2])

For any matrix K with positive elements there are diagonal matrices $\mathrm{diag}(u), \mathrm{diag}(v)$ such that $M = \mathrm{diag}(u)K\mathrm{diag}(v)$ has prescribed row- and column-sums μ_0 and μ_1. The vectors u and v can be obtained by alternating perform:

$$u = \mu_0./(Kv)$$

$$v = \mu_1./(K^T u).$$

[1] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.
[2] R. Sinkhorn. Diagonal equivalence to matrices with prescribed row and column sums. *The American Mathematical Monthly*, 74(4), 402–405, 1967.

Recently proposed to solve via entropy regularization [1]: $KL(M \mid \mathbf{1}\mathbf{1}^T) = \sum_{i,j=1}^{\mathcal{N}^d} (M_{ij}\log(M_{ij}) - M_{ij} + 1)$,

$$\min_{M \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d}} \quad \langle C, M \rangle + \epsilon KL(M \mid \mathbf{1}\mathbf{1}^T)$$

subject to $M\mathbf{1}$

Reduces the number of variables from $\mathcal{N}^{2d}$ to $2\mathcal{N}^d$!

- Let $\exp(\cdot)$
- For $K = \exp(-C/\epsilon)$, the solution is of the form:

$$M = \mathrm{diag}(u) K \mathrm{diag}(v).$$

Theorem (Sinkhorn iterations [2])

For any matrix K with positive elements there are diagonal matrices $diag(u), diag(v)$ such that $M = diag(u)K diag(v)$ has prescribed row- and column-sums μ_0 and μ_1. The vectors u and v can be obtained by alternating perform:

$$u = \mu_0./(Kv)$$

$$v = \mu_1./(K^T u).$$

[1] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.
[2] R. Sinkhorn. Diagonal equivalence to matrices with prescribed row and column sums. *The American Mathematical Monthly*, 74(4), 402–405, 1967.

More in-depth on how to derive the Sinkhorn iterations: consider

$$\begin{aligned} \min_{M \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d}} \quad & \langle C, M \rangle + \epsilon KL(M \mid \mathbf{1}\mathbf{1}^T) \\ \text{subject to} \quad & M\mathbf{1} = \mu_0 \\ & M^T\mathbf{1} = \mu_1. \end{aligned}$$

- Using Lagrangian relaxation gives

$$L(M, \lambda, \zeta) = \langle C, M \rangle + \epsilon KL(M \mid \mathbf{1}\mathbf{1}^T) + \lambda^T(\mu_0 - M\mathbf{1}) + \zeta^T(\mu_1 - M^T\mathbf{1}).$$

- Given dual variables λ, ζ, the minimum M_{ij} is

$$0 = \frac{\partial L(M, \lambda, \zeta)}{\partial M_{ij}} = C_{ij} + \epsilon \log(M_{ij}) - \lambda_i - \zeta_j$$

- Solve for M_{ij} to get

$$M_{ij}^* = e^{\lambda_i/\epsilon} e^{-C_{ij}/\epsilon} e^{\zeta_j/\epsilon}.$$

- Change of variables: $u = \exp(\lambda/\epsilon)$, $v = \exp(\zeta/\epsilon)$. The optimal solution is of the form

$$M^* = \text{diag}(u) K \text{diag}(v)$$

  where $K = \exp(-C/\epsilon)$.

One way to interpreted the Sinkhorn iterations: coordinate ascent in the Lagrangian dual.

- Lagrangian relaxation gave optimal form of the primal variable

$$M^* = \mathrm{diag}(u)K\mathrm{diag}(v)$$

- The Lagrangian dual function:

$$\varphi(u,v) := \min_{M \in \mathbb{R}_+^{\mathcal{N}^d \times \mathcal{N}^d}} L(M,u,v) = L(M^*,u,v) = \ldots = \epsilon \log(u)^T \mu_0 + \epsilon \log(v)^T \mu_1 - \epsilon u^T K v + \epsilon \mathcal{N}^{2d}.$$

- The dual problem is thus

$$\max_{u,v \in \mathbb{R}_+^{\mathcal{N}^d}} \varphi(u,v)$$

- Taking the gradient w.r.t u and putting it equal to zero gives

$$\epsilon \mu_0./u - \epsilon K v = 0 \quad \rightsquigarrow \quad u = \mu_0./(Kv),$$

and w.r.t v gives

$$\epsilon \mu_1./v - \epsilon \left(u^T K\right)^T = 0 \quad \rightsquigarrow \quad v = \mu_1./(K^T u).$$

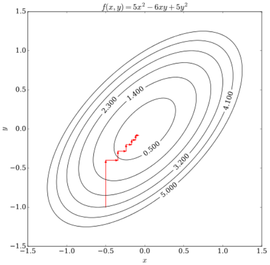


Figure from Wikipedia

These are the Sinkhorn iterations! (cf. [1])

[1] P. Tseng. Dual ascent methods for problems with strictly convex costs and linear constraints: A unified approach. *SIAM Journal on Control and Optimization*, 28(1), 214–242, 1990.

Summary and take-home message:

- Optimal transport can be used to solve matching problems for sets of dynamic agents;
- With this as basis, it can be used to solve optimal control problems for swarm of agents;
- The problems can be solve numerically by developing numerical methods along the lines of the Sinkhorn iterations.

Optimal transport - a viable framework for many problems and applications!

Thank you for your attention!

Questions?