University of Utah
UNDERGRADUATE RESEARCH JOURNAL

DESIGNED SWARMING BEHAVIOR USING OPTIMAL TRANSPORTATION NETWORKS

Justin Marlin Baker (Elena Cherkaev)
Department of Mathematics

ABSTRACT

Models of swarming behavior aid in disaster planning, direct the actions of warehouse robots, and can map the foraging characteristics of insects. These models use the optimal behavior of individual agents to determine the behavior of the larger population. Optimal transportation is one such model that has been used to successfully describe the behavior of swarming agents[3]. The Monge-Kantorovich formulation of the optimal transportation problem models the best direct route for individual agents in a swarm. The presented work investigates optimal transportation and duality in a Monge-Kantorovich formulation. We reduce the Monge-Kantorovich formulation to the linear programming problem which can be efficiently solved numerically. However this formulation is not applicable in the case where the domain of travel is restricted, so that agents must travel along a particular network of paths. For instance, robots traveling throughout a warehouse must navigate through a network of aisles, and foraging bugs may prefer to travel along premade routes. We extend the Monge-Kantorovich formulation to a network of paths which differs from transfer over a direct route. We show that this extended formulation can also be reduced to the linear programming problem. Finally, we investigate various applications of both the direct and network transfer using numerical simulations.

TABLE OF CONTENTS

1. OPTIMAL TRANSPORTATION PROBLEM

Optimal transportation describes the transfer of mass from one location to another with the least amount of work. The classical problem was formulated by Gaspard Monge in the 1780s as *"How should one transfer, with minimal work, a pile of sand from one location to fill a given hole?"*. Coming from a background of industry and economics, Monge was interested in extending this problem to the transport goods between supply factories and distributed demand. Thus the problem is concerned with how smaller amounts of sand are particularly redistributed, or to which location each individual good travels.

The original problem is rather ill-posed. It is unclear if there is an optimal transportation route, and process of transportation is unconstrained. For this reason, it wasn't until two hundred years later in 1979 that the first general proof was proposed by V.N. Sudakov. However, several years prior Leonid Kantorovich had proved a relaxed version of the problem using the linear programming techniques that he helped develop. The Monge-Kantorovich formulation imposes conditions on the *cost* of work and requires that the masses are transferred along the network in identical amounts.

Due to its history, transportation theory is often motivated by spacial transfer. Physical applications include compression modeling, semi-geostrophic system modeling, kinetic theory, optimal design of lenses, irrigation design and network design. Therefore visualization techniques are highly effective in the study of optimal transportation theory. Section four will discuss these visualization techniques along with the applications of palette swapping, particle swarm optimization, material science, emergency planning, and Weistrass Generative Adversarial Networks (WGAN).

The primary aim of this paper is to extend the Monge-Kantorovich problem to the transfer masses along a directed graph. This modification of the problem requires an understanding of the initial Monge-Kantorovich formulation discussed in sections one and two. Then in section three an adjusted cost function will be introduced along with the extension of the Monge-Kantorovich formulation to directed graphs.

Variational Formulation

The Monge optimal transportation problem is formulated as a problem of calculus of variations. The goal is to formulate the problem mathematically so that the solution is the network which minimizes the cost of transportation.

Consider an *initial domain* $X \in \mathbb{R}^n$ with mass distribution $\mu^+(x)$ for $x \in X$. The *target domain* $Y \in \mathbb{R}^m$ has a distribution which describes the capacity for masses $\mu^-(y)$ for $y \in Y$. The objective is to find the *transportation network* or *mapping* $s(x)$, which maps $\mu^+(x)$ onto $\mu^-(x)$.

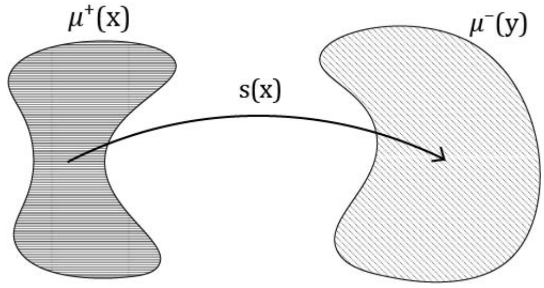


Figure 1: Optimal transportation path between two probability distributions

$$s(x) : \mu^+(x) \to \mu^-(y), \quad x \in X,\, y \in Y$$

This mapping is only feasible if it satisfies a mass balance equation. Thus we constrain the system to having equivalent mass in the initial domain to space in the target domain.

$$\int_X \mu^+(x)dx = \int_Y \mu^-(y)dy$$

Additionally a *cost density* function is used to describe the work of moving masses along the network. In general this cost function is a non-negative real valued function which might take infinite values.

$$c(x, s(x)) : \mathbb{R}^n \times \mathbb{R}^m \to [0, \infty)$$

The mass transfer problem optimizes the cost of redistribution of all particles in the initial domain by selecting an optimal mapping. This process is mathematically formulated as optimization of the following integral, sometimes called the work functional or goal functional.

$$I[s(x)] = \int_{\mathbb{R}^n} c(x, s(x))d\mu^+(x)$$

The variational problem is to select the particular transportation network which minimizes this functional. In addition this network must satisfy the mass balance equation. Let s^* denote the map which minimizes the functional so that $I[s^*] \leq I[s]$ for all possible mappings s.

$$\begin{aligned} I[s^*(x)] &= \min_{s(x)} \int_{\mathbb{R}^n} c(x, s(x))d\mu^+(x) \\ \text{s.t.} \quad & \int_X \mu^+(x)dx = \int_Y \mu^-(y)dy \end{aligned} \tag{1}$$

This formulation is referred to as the *primal problem.* From the primal problem it is not apparent that a mapping s^* which minimizes this functional always exists. And if it exists it is not clear how to find it. This problem can be reformulated using Lagrangian duality.[2] This will allow us to show that an optimal mapping does exist.

Langrangian Duality

The primal problem is formulated as a constrained optimization problem. The constraint is the mass balance equation. To reformulate the problem, we introduce a *Lagrange multiplier* λ accounting for this constraint. This serves as a penalty if the constraint is not satisfied. The Lagrange dual formulation is the reformulation of the primal problem utilizing this penalty term.

The augmented Lagrange functional has the following form

$$L(\lambda, x) = \max_{\lambda} \min_{s(x)} \{I(x, s(x)) + \lambda G(x)\}$$

where

$$G(x) = \int_X \mu^+(x)dx - \int_X \mu^-(s(x))dx = 0.$$

So that this $L(\lambda, x)$ is

$$L(\lambda, x) = \max_{\lambda} \min_{s(x)} \left\{ \int_{\mathbb{R}^n} c(x, s(x))d\mu^+(x) + \lambda \left[\int_X \mu^+(x)dx - \int_X \mu^-(s(x))dx \right] \right\}.$$

The dual formulation identifies both the optimal path s^* and penalty λ. However, the dual variation only provides a lower bound on the solution, which may not be exact. Therefore it must be shown that for the particular problem the solution to the Lagrange dual problem is equivalent to the solution to the primal problem. This is referred to as strong duality and we prove here that strong duality is satisfied.

Theorem 2.1. *Let $\lambda \in \mathbb{R}_+$, $\lambda\mu^+(x) = -u(x)$, and $\lambda\mu^-(y) = v(y)$. The solution to the Lagrangian dual problem*

$$L^*(\lambda) = \max_{u(x),v(y)} \{\int_X u(x)d\mu^+(x) + \int_Y v(y)d\mu^-(y)\} \tag{2}$$
$$s.t. \quad c(x,y)) \geq u(x) + v(y)$$

satisfies the strong duality condition. That is $L^*(\lambda) = I[s^*(x)]$.

Proof. First to see that this is indeed the Lagrangian dual problem let

$$\hat{L}(s(x),\lambda) = \min_{s(x)} \max_{\lambda} \{I(x,s(x) + \lambda G(x)\}.$$

The optimal value of this functional is equal to $I[s^*(x)]$. However, to solve the Lagrangian dual problem, the order of minimization and maximization is switched introducing the duality gap.

$$\max_{\lambda} \min_{s(x)} \{I(x,s(x) + \lambda G(x)\} \leq \min_{s(x)} \max_{\lambda} \{I(x,s(x) + \lambda G(x)\}$$

The solution of the functional $L(s(x),\lambda)$ is given by

$$\begin{aligned} L^*(\lambda) &= \max_{\lambda} \min_{s(x)} \{I(x,s(x) + \lambda G(x)\} \\ &= \max_{\lambda} \{I(x,s^*(x)) + \lambda G(x)\} \end{aligned}$$

First we find the minimum over $s(x)$. Because $u(x)$ and $v(s(x))$ are functions of λ we have that the minimum is not necessarily exactly zero but that $\frac{d}{dx}[I(x,s^*(x) + \lambda G(x)] \geq 0$.

$$\begin{aligned} I(x,s^*(x)) + \lambda G(x) &= \frac{d}{dx}\left[\int_{\mathbb{R}^n} c(x,s(x))d\mu^+(x) - \int_X u(x)dx - \int_X v(s(x))dx\right] \\ &= c(x,s^*(x)) - u(x) - v(s^*(x)) \geq 0 \\ &\Rightarrow c(x,s^*(x)) \geq u(x) + v(s^*(x)) \end{aligned}$$

Then this cost function is substituted into the Lagrangian for $c(x, s(x))$, and this constraint is added. Note the change of integration to integration over the measures $d\mu^+$ and $d\mu^-$ rather than the spatial variable dx. The Lagrangian dual problem is then

$$\begin{aligned} &L^*(\lambda) = \max_{\lambda}\{\int_X u(x)d\mu^+(x) + \int_Y v(y)d\mu^-(y).\} \\ &\text{s.t.} \quad c(x, s^*(x)) \geq u(x) + v(s^*(x)) \end{aligned}$$

Strong duality requires that $\lambda G(x) = 0$ so that

$$L^*(\lambda) = \max_{\lambda}\{I(x, s^*(x)) + \lambda G(x)\} = I(x, s^*(x))$$

Therefore,

$$\lambda \left[\int_X \mu^+(x)dx - \int_X \mu^-(s(x))dx\right] = 0$$

which is true if the mass balance equation is satisfied along the entire transportation network. □

Dual Convex Functions

The final step in showing that there exists a unique transportation route requires the tools of convex analysis. Therefore, the final reformulation of the problem formulates the Lagrange multipliers from the dual problem into dual convex functions. Suppose that the costs functional is uniformly convex and of the form

$$c(x, y) = \frac{1}{2}|x - y|^2.$$

Then the work functional is given by

$$I[s] = \frac{1}{2}\int_{\mathbb{R}^n} |x - s(x)|^2 d\mu^+(x).$$

Furthermore suppose that $s(x)$ is essentially one to one. Then the graph of $\{(x, s^*(x))|x \in X\} \subset \mathbb{R}^n \times \mathbb{R}^n$ is cyclically monotone.[2] Which implies that the optimal mapping s^* lies in the subdifferential of a convex mapping of $\mathbb{R}^n$ to $\mathbb{R}$.[5] Thus for some differentiable convex function ϕ^* and its gradient $D\phi^*(x)$,

$$s^* = D\phi^*$$

The problem is now to find a change of variables from $u(x), v(y)$ to $\phi(x), \psi(y)$ that satisfies this differentiable function. For the given cost function above, this change of variables is

$$\phi(x) = \frac{1}{2}|x|^2 - u(x)$$
$$\psi(y) = \frac{1}{2}|y|^2 - v(y).$$

The constraints in (2) become

$$c(x, y) \geq u(x) + v(y)$$
$$-u(x) - v(y) \geq -\frac{1}{2}|x - y|^2$$
$$\phi(x) + \psi(y) = \frac{1}{2}|x|^2 - u(x) + \frac{1}{2}|y|^2 - v(y)$$
$$\phi(x) + \psi(y) \geq \frac{1}{2}|x|^2 + \frac{1}{2}|y|^2 - \frac{1}{2}|x - y|^2$$
$$\phi(x) + \psi(y) = x \cdot y$$

Thus the convex dual problem is

$$L^*[\phi, \psi] = \min_{\phi,\psi} \quad \int_X \phi(x)dx + \int_Y \psi(y)dy \tag{3}$$
$$\text{s.t.} \quad \phi(x) + \psi(y) \geq x \cdot y$$

The following proof is the crux of the Monge-Kantorovich solution and shown in the work by L.C. Evans[2].

Theorem 2.2. *(i) There exists (ϕ^*, ψ^*) which uniquely solves the minimization problem (3).*
(ii) Furthermore, (ϕ^*, ψ^*) are dual convex functions, in the sense that

$$\phi^*(x) = \max_{y \in Y}(x \cdot y - \phi^*(y)) \quad (x \in X)$$
$$\psi^*(y) = \max_{x \in X}(x \cdot y - \psi^*(x)) \quad (y \in Y)$$

Proof. If ϕ, ψ satisfy the constraint of (3) then

$$\phi(x) \geq \max_{y \in Y}(x \cdot y - \psi(y)) := \hat{\phi}(x)$$

and

$$\hat{\phi}(x) + \psi(y) \geq x \cdot y \quad (x \in X,\, y \in Y).$$

Furthermore,

$$\psi(y) \geq \max_{x \in X}(x \cdot y - \phi(\hat{x})) := \hat{\psi}(y)$$

so that,

$$\hat{\phi} + \hat{\psi} \geq x \cdot y \quad (x \in X,\, y \in Y)$$

Where if $\psi \geq \hat{\psi}$

$$\max_{y \in Y}(x \cdot y - \hat{\psi}(y)) \geq \hat{\phi}(x).$$

$\square$

2. REDUCTION TO THE LINEAR PROGRAMMING PROBLEM

We will now discretize the problem problem and reduce it to the linear programming problem. This allows us to solve the problem computationally. Linear programming is a method which optimizes an objective function subject to linear constraints. It can be generally formulated as

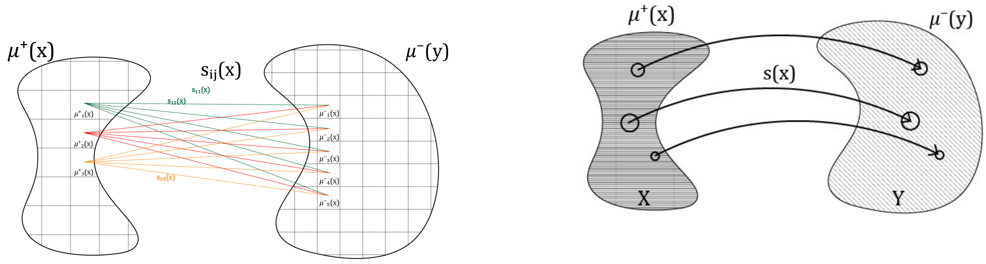
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \\ & h(x) \leq 0 \end{aligned} \tag{4}$$

where $g(x)$ and $h(x)$ are systems of linear equations.

Each of the functionals discussed in the previous section can be formulated as a linear programming problem if considered in a discrete setting. This us to find a computational solution for the variational problems.

Discretized Formulation

To find a numerical solution, we first discretize over the domain space. In the following figure two discretization methods are considered.



(a) Grid spaced discretization

(b) Discretization with equivalent point masses

Figure 2: Comparison of discretization schemes

In both instances the problem is discretized by considering the atomic measure. In other words, the total mass is the sum of several small regions δ_{x_i} of masses.

$$\mu^+(x) = \sum_{i=0}^{n} \mu_i^+(x)\delta_{x_i}$$
$$\mu^-(x) = \sum_{j=0}^{m} \mu_j^-(x)\delta_{y_j}$$

In the grid spaced discretization, each small region is a square on a grid. However, this does not assure that the masses are equivalent in each region. Instead selectively spacing and sizing each δ_{x_i} so that each small region contains equal mass, assures that the transportation map is one to one. Recall that the optimal mapping must be one to one to have a unique minimum. Therefore, the second method is preferred because it transfers masses in equivalent amounts.

Then for each central location x_i and y_j the cost function and the network are discretized as follows.

$$c(x_i, y_j) = c_{ij} = |x_i - y_j|^2$$
$$s(x_i, y_j) = s_{ij}$$

Then the discrete version of (1) is given by

$$\min_{\mu_{ij}} \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} s_{ij} \tag{5}$$
$$\sum_{i=1}^{n} \mu_i^+ = \sum_{j=1}^{m} \mu_j^-$$

Primal Problem

The method of linear programming can be applied to this discretized version of the primal problem. The relaxed transportation problem can be formulated as the following linear programming problem.

$$\begin{aligned} \min \quad & c \cdot s \\ \text{s.t.} \quad & As = b \\ & x \geq 0 \end{aligned}$$

Consider the vector dimensions of each of these, where $\mu_i^+ \in \mathbb{R}^n$, $\mu_j^- \in \mathbb{R}^m$ the cost function is given by $c_{ij} \in \mathbb{R}^{mn}$, and $s_{ij} \in \mathbb{R}^{mn}$ describes the transportation network. The mass balance constraints over the entire network are formulated as

$$\sum_{j=1}^{m} s_{ij} = \mu_i^+, \quad \sum_{i=i}^{n} s_{ij} = \mu_j^-.$$

The objective is to minimize the cost function over the network subject to the constraints above,

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mu_{ij}.$$

To formalize the linear programming problem we define each of these terms. Let $s \in \mathbb{R}^{mn}$, where

$$s = (s_{11}, s_{12}, ..., s_{1m}, s_{21}, ..., s_{nm}).$$

Write b as the right hand side of the constraints, $b \in \mathbb{R}^{m+n}$, where

$$b = (\mu_1^+, \mu_2^+, ..., \mu_n^+, \mu_1^-, ..., \mu_m^-).$$

Then construct A so that the constraints are satisfied. A is an $(n+m) \times nm$ matrix composed of ones vectors $\mathbb{1} \in \mathbb{R}^m$ and the basis vectors $e_m \in \mathbb{R}^m$.

$$A = \begin{pmatrix} \mathbb{1} & 0 & ... & 0 \\ 0 & \mathbb{1} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbb{1} \\ e_1 & e_1 & ... & e_1 \\ e_2 & e_2 & ... & e_2 \\ ... & ... & ... & ... \\ e_m & e_m & ... & e_m \end{pmatrix}$$

The constraints in the problem are then satisfied by the linear function $As = b$. Thus for some c, s^* is the minimizer of $c \cdot s$ and the unique transportation network.

Dual Problem

Developing on the same vectors and matrices A, b and c as in the primal problem, the Lagrangian dual problem may be formulated as.

$$\max \quad b \cdot y$$
$$\text{s.t.} \;\; A^T y \leq c.$$

Here y describes the Lagrange dual pair (u, v). Recall from (2) that the goal is to maximize the value of the (u, v) pair applied to $\int_X d\mu^+$ and $\int_Y d\mu^-$. The discrete formulation is to minimze the inner product of u and v with the distribution μ^+ and μ^-.

To show that the Lagrangian dual problem can be formulated from the discretization of the primal problem consider the constraints given in the discrete primal problem

$$\sum_{i=1}^{n} \mu_{ij} - \mu_i^+ = 0$$
$$\sum_{j=1}^{m} \mu_{ij} - \mu_j^- = 0$$

The discrete Lagrangian dual problem is given by

$$L(u,v) = \min_x \{ f + u(\sum_{i=1}^{n} \mu_{ij} - \mu_i^+) + v(\sum_{j=1}^{m} \mu_{ij} - \mu_j^-)\}$$

and

$$f = c_{ij} \cdot \mu_{ij}.$$

Solving for the minimum over all x gives

$$u_i + v_j \leq c_{ij}.$$

Substituting for c_{ij}

$$\begin{aligned} L(u,v) &= \min_x \{ c_{ij}\mu_{ij} + u(\sum_{i=1}^{n} \mu_{ij} - \mu_i^+) + v(\sum_{j=1}^{m} \mu_{ij} - \mu_j^-)\} \\ &\geq (u_i + v_j)\mu_{ij} + u\sum_{i=1}^{n} \mu_{ij} - u_i\mu_i^+ + v\sum_{j=1}^{m} \mu_{ij} - v_j\mu_j^- \\ &= u\sum_{i=1}^{n} \mu_{ij} + v\sum_{j=1}^{m} \mu_{ij} \end{aligned}$$

This gives us the discrete Lagrangian dual formulation as

$$\max_{u,v} \quad u\sum_{i=1}^{n} \mu_i^+ + v\sum_{j=1}^{m} \mu_j^-$$

$$\text{s.t.} \;\; u_i + v_j \leq c_{ij}$$

Formally, construct y so that $y \in \mathbb{R}^{n+m}$, and $y = (u_1, u_2, .., u_n, v_1, v_2, ..., v_m)$. Then the Lagrangian dual variation is formulated as the linear programming problem described above.

Dual Convex Functions

The same technique may be applied to the dual convex conjugate problem. The linear programming problem is

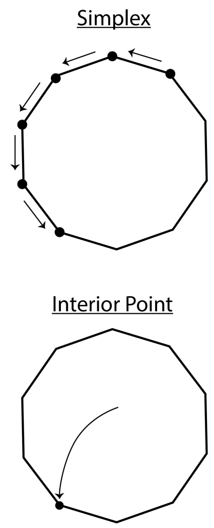
$$\min \quad b \cdot p$$

$$\text{s.t.} \quad A^T p \geq v$$

Where A and b are defined above, $p = (\phi_1, ...\phi_n, \psi_1, ..., \psi_m)$, and $v = (x_1^+, ..., x_n^+, x_1^-, ..., x_m^-)$. This directly follows from (4) as the convex conjugate functions are coordinate transformations of the dual problem.

Solution Methods

There are two approaches to solving a linear programming problem. First note that the solutions to the LP are lie along the boundary points of a convex polytope. The simplex method minimzes over the boundary points. The interior point method minimizes along the interior adding a penalty to the boundary. In addition to direct LP solvers, there are second order conic solvers (SOCP) which can solve for linear programming.[1]

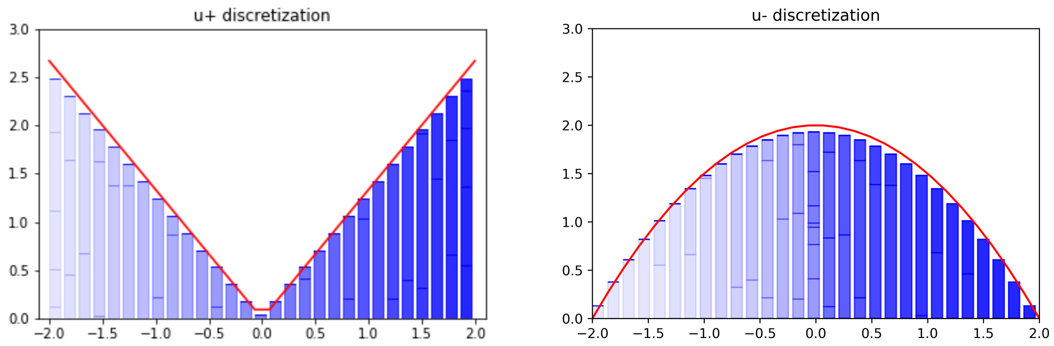


For the purposes of programming the embeded conic solver (ECOS) was used as a solver to the linear programming problem. This solver uses an interior point method to solve for second order conic problems (SOCP). However, SOCP solvers generalize to LP solvers. Thus the method can be used for all of the linear programming problems discussed above.

Linear Transfer Analysis

Consider two mass distributions, one which is bimodal and convex and the other is unimodal and concave. The goal is to push the bimodal distribution onto the unimodal distribution using the discretization and linear programming techniques discussed so far.

The distributions are discretized along the real number line and then shaded by approximate initial location. This local approximation is necessary to view the distribution in this histogram fashion where the sum of the masses within a given region are stacked onto one discrete point. The discretization method itself reduces the mass transfer to equivalent amounts. The colored shading shows the redistribution of each of the cumulative masses between the initial distribution and the final distribution. They are shaded based on initial location, and then this shading is maintained in the final distribution.



(a) Initial distribution with shading (b) Redistribution shaded by initial location

Figure 3: Discretization and transfer of masses using linear programming

3. EXTENSION OF MONGE-KANTOROVICH FORMULATION TO A NETWORK

Building upon the techniques used so far, consider the case where the transportation network $s(x)$ may contain multiple linear or nonlinear paths between two regions. Each path has a starting node and an ending node. All possible transportation networks from x to y are the simple paths along the directed graph between each of the nodes. The cost function is the sum over the entire transportation network of the individual costs between each node. This problem can be formulated and solved similarly to the direct transportation formulations discussed above. This section will describe the necessary modifications to the problem along each of the formulations.

Problem Formulation

First note that several key features between the problems are identical. The primal problem (1) has no constraints on the cost function or network and remains the same.

$$I[s^*(x)] = \min_{s(x)} \int_{\mathbb{R}^n} c(x, s(x)) d\mu^+(x)$$
$$\text{s.t.} \quad \int_X \mu^+(x)dx = \int_Y \mu^-(y)dy$$

This implies that the Lagrangian dual problem (2) is also consistent.

$$L^*(\lambda) = \max_{\lambda}\{\int_X u(x)dx + \int_X v(s(x))dx\}$$
$$\text{s.t.} \quad c(x, s^*(x)) \geq u(x) + v(s^*(x))$$

Utilizing a similar argument for the dual convex problem, formulate the cost function as the individual cost along each of the nodes. Let $\{X_0\}$ be the set of starting nodes, X_N be the set of final nodes and $X_{i=2,\ldots,N-1}$ be the set of nodes along each step in the directed graph. Then for $x \in X$ the cost function is determined by

$$c(x_0, s(x)) = \sum_{i=1}^{N} \frac{1}{2}|x_i - x_{i-1}|^2.$$

This cost function is the sum of uniformly convex functions and is therefore uniformly convex. Therefore by extension of the argument for uniform convexity in §1, the network $s^*(x)$ is cyclically monotone and one to one under the same discretization scheme. This implies that there exists some convex potential function $\phi^*(x)$ so that the optimal transportation scheme lines within the gradient of ϕ^*.

What remains is to verify the coordinate transformation that formulates the dual convex functions $(\phi^*(x), \psi^*(y))$ such that $s^*(x) \subset D\phi^*(x)$. Following the proof by L.C. Evans, we here show that this extends to this formulation as well.

Theorem 4.1. *There exists convex dual functions* $(\phi^*(x), \psi^*(y))$ *given by*

$$\phi^*(x_0) = \sum_{i=1}^{N-1} \frac{1}{2}|x_i - x_{i-1}|^2 - u(x_0)$$
$$\psi^*(x_N) = \sum_{i=2}^{N} \frac{1}{2}|x_i - x_{i-1}|^2 - v(x_N)$$

such that $s^*(x) \subset D\phi^*(x)$ *and*

$$\phi^*(x_0) = \max_{x_N}\{\hat{c} - \psi(x_n)\}$$
$$\psi^*(x_N) = \max_{x_0}\{\hat{c} - \phi(x_0)\}$$

where

$$\hat{c} = \max_{x_0,\ldots,x_N}\{\sum_{i=1}^{N} x_i \cdot x_{i-1}\}.$$

Proof. Note that because $\hat{c}$ is a sum of inner products between each of the components in the graph, thus the solution is invariant to the order in which the points are maximized. If ϕ, ψ satisfy the constraint of (3) then

$$\phi(x) \geq \max_{x_N}(\hat{c} - \psi(y)) := \hat{\phi}(x)$$

and

$$\hat{\phi}(x) + \psi(y) \geq \hat{c} \quad (x \in X,\, y \in Y).$$

Furthermore,

$$\psi(y) \geq \max_{x_0}(\hat{c} - \hat{\phi(x)}) := \hat{\psi}(y)$$

so that,

$$\hat{\phi} + \hat{\psi} \geq \hat{c} \quad (x \in X,\, y \in Y)$$

Where if $\psi \geq \hat{\psi}$

$$\max_{x_n}(\hat{c} - \hat{\psi}(y)) \geq \hat{\phi}(x).$$

□

Finally to see that this is indeed consistent with $\phi^*(x_0)+\psi^*(x_N) \geq \sum_{i=1}^{N} x_i \cdot x_{i-1}$

$$\begin{aligned}
c(x_0, x_N) &\geq u(x_0) + v(x_N) \\
-u(x_0) - v(x_N) &\geq -\sum_{i=1}^{N} \frac{1}{2}|x_i - x_{i-1}|^2 \\
\phi(x_0) + \psi(x_N) &= \sum_{i=1}^{N-1} \frac{1}{2}|x_i - x_{i-1}|^2 - u(x_0) + \sum_{i=2}^{N} \frac{1}{2}|x_i - x_{i-1}|^2 - v(x_N) \\
\phi(x_0) + \psi(x_N) &\geq \sum_{i=1}^{N-1} \frac{1}{2}|x_i - x_{i-1}|^2 + \sum_{i=2}^{N} \frac{1}{2}|x_i - x_{i-1}|^2 - \sum_{i=1}^{N} \frac{1}{2}|x_i - x_{i-1}|^2 \\
\phi(x) + \psi(y) &= \sum_{i=1}^{N} x_i \cdot x_{i-1}
\end{aligned}$$

Thus the convex dual formulation is given by

$$\begin{aligned}
\min_{\phi,\psi} \quad & \int_X \phi(x)dx + \int_Y \psi(y)dy \\
\text{s.t.} \quad & \phi(x) + \psi(y) \geq \sum_{i=1}^{N} x_i \cdot x_{i-1}
\end{aligned} \tag{6}$$

Network Transfer as a Linear Programming Problem

The cost function evaluated as the sum over each network component requires a modification to the discrete network vector s_{ij} and thus the linear programming problem. First note that the dimension of s is increased for each set of points added between x_i and x_j. In our change of notation we let $x_i \in X_0$ and $x_j \in X_N$. Consider the problem when $N = 2$ so that there is one set of k interior nodes X_1. Then the vector s is then given by

$$s_{itj} = \begin{pmatrix} s_{000} & s_{010} & \dots & s_{0km} & s_{100} & \dots & s_{nkm} \end{pmatrix}$$

In general denote the set of all sets of interior points as $X_{k=1,\dots,N-1} = \{K\}$ with ℓ_k nodes in the kth set, and then the dimension of K is given by $\ell = \Pi_{k=1}^{N-1} \ell_k$. Thus the vectors for the network s and the cost c are both in $\mathbb{R}^{m\ell n}$. Thus the matrix A is modified so that the primal problem is consistent

$$\begin{aligned} \min \quad & c \cdot s \\ \text{s.t.} \quad & As = b \\ & s \geq 0 \end{aligned}$$

The constraints are consistent

$$\sum_{j=1}^{m} s_{ikj} = \mu_i^+, \quad \sum_{i=i}^{n} s_{ikj} = \mu_j^- .$$

and so the vector b is invariant. Where $b \in \mathbb{R}^{m+n}$ is given by

$$b = (\mu_1^+, \mu_2^+, ..., \mu_n^+, \mu_1^-, ..., \mu_m^-).$$

Then construct A so that the constraints are satisfied. A is an $(n+m) \times nm\ell$ matrix composed of ones vectors $\mathbb{1} \in \mathbb{R}^{m \cdot \ell}$ and basis vectors e_m where there is a vector of ones of length $\hat{\mathbb{1}} \in \mathbb{R}^{\ell}$. in the $j \cdot \ell$ spot for $j = 1, ..., m$. If ℓ is 2 and m is

3 then $e_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$.

$$A = \begin{pmatrix} \mathbb{1} & 0 & ... & 0 \\ 0 & \mathbb{1} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbb{1} \\ e_1 & e_1 & ... & e_1 \\ e_2 & e_2 & ... & e_2 \\ ... & ... & ... & ... \\ e_m & e_m & ... & e_m \end{pmatrix}$$

Because the Lagrange dual is preserved, so is the dual linear programming problem with A given above.

$$\begin{aligned} \max \quad & b \cdot y \\ \text{s.t.} \quad & A^T y \leq c \end{aligned}$$

where $y \in \mathbb{R}^{n+m}$, and $y = (u_1, u_2, .., u_n, v_1, v_2, ..., v_m)$.

The linear programming problem for the dual convex functions must be modified more substantially. It is easier here to separate ϕ and ψ from the vector p. Then the dual convex linear programming problem is defined as

$$\begin{aligned} \min \quad & \mu^+ \cdot \phi + \mu^- \cdot \psi \\ \text{s.t.} \quad & \phi(x) + \psi(y) \geq \hat{c} \quad (x \in X_0,\ y \in X_N) \end{aligned}$$

where $\hat{c}$ is defined above.

The same techniques from direct transfer can be applied to resolve the optimal transportation problem with a directed graph. It is important to note that with each additional set of points the size of the problem grows exponentially. Thus these systems can become very large and difficult to solve.

Network Transfer Analysis

In the following example four masses are transferred from an initial location to their target solutions by lattice points. From each lattice point the masses may redistribute themselves to any of the eight neighboring lattice points. Thus a directed graph is formed between each of the lattice points. The network is formed from each set of simple paths connecting the initial location to the corresponding final location. However, lattice movement restricted in the first region to only lattice points above the initial location, and along the boundary of a 5×4 grid. With these constraints, the masses attempt to reduce the cost of movement by maximizing transfer along the diagonals.

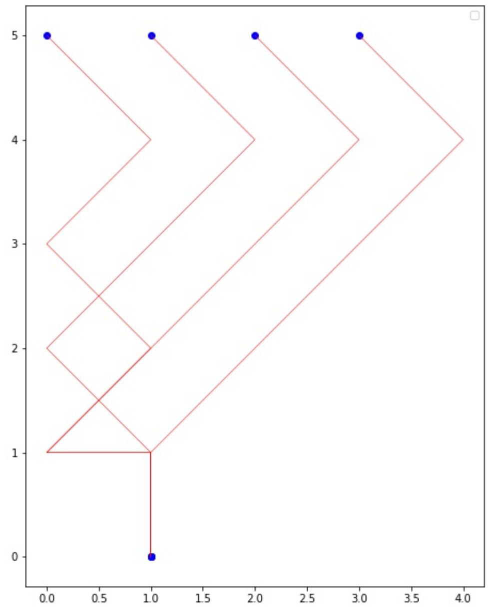


Figure 4: Optimal transportation along a directed graph of lattice points

Network Cost Variation Analysis

Suppose that the cost function is varied over two different regions. Then a set of intermediate nodes may be added along the boundary between the cost variation and the cost function discussed above is modified to reflect the variation between the two regions.

In particular for some α_i acting as a scalar modifier of the cost function in each region, the cost function is defined as

$$c(x_0, s(x)) = \sum_{i=1}^{N} \alpha_i |x_i - x_{i-1}|^2.$$

This cost function is still uniformly convex and thus the analysis holds.

In the following example, the cost is higher in the darker region. While there may be numerous points that are discretized, the visualization displays the primary transfer locations for the masses. In this example the masses are redistributed in the low cost region, and then move linearly to their final location. This is because the cost of moving along a diagonal is higher than moving along a straight line. Thus diagonal transfer occurs in the low cost region .

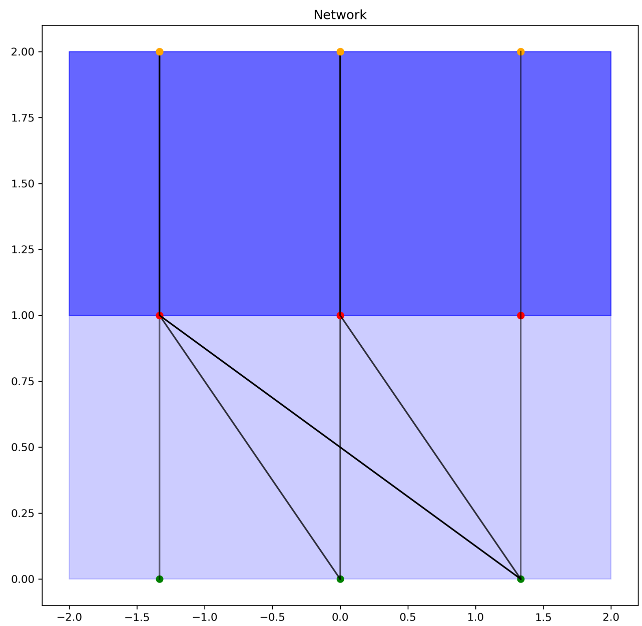


Figure 5: Cost variation over two regions

4. APPLICATIONS

Many applications of the optimal transportation problem model spatial relations between objects. Thus the development of visualization tools can be directly applied to optimal transportation problems. Two primary aspects of visualization are considered here. First is the redistribution network for each individual particle. The second is cost of movement. Consider the following transfer problem.

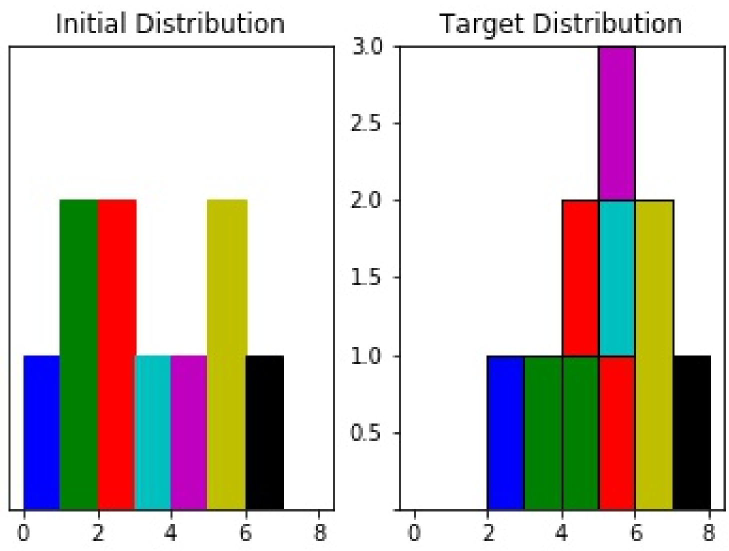


Figure 6: Discrete redistribution by color scheme

One option for visualization is to color the location of masses in the initial location to visualize where they are redistributed to.

This technique may be combined with a heat map of the cost for transfer between two locations. This visualization technique is useful in the study of deep learning networks to visualize the network. This particular technique is used in the study of Weirstrauss Generative Adversarial Networks. Utilizing the Monge-Kantorovich formulation, this process trains a generator and a discriminator on several images. The objective of the generator is to generate images which are close to the images that the discriminator is trying to discriminate between. A further discussion of images as distributions will be introduced in palette swapping, however the generator can be improved by finding optimal redistributions of known image distributions. The implementation of WGAN can reduce the workload on the learning network by reducing the number of updates needed to balance the discriminator and generator. This visualization technique helps remove some of the ambiguity of what happens within the deep learning network.

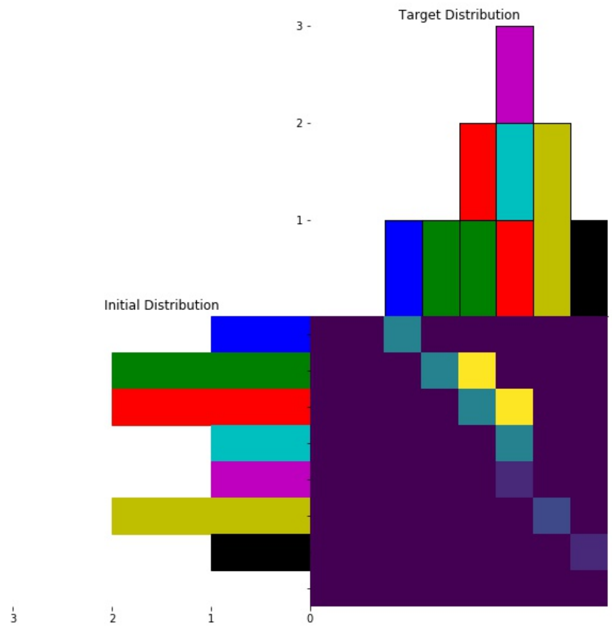
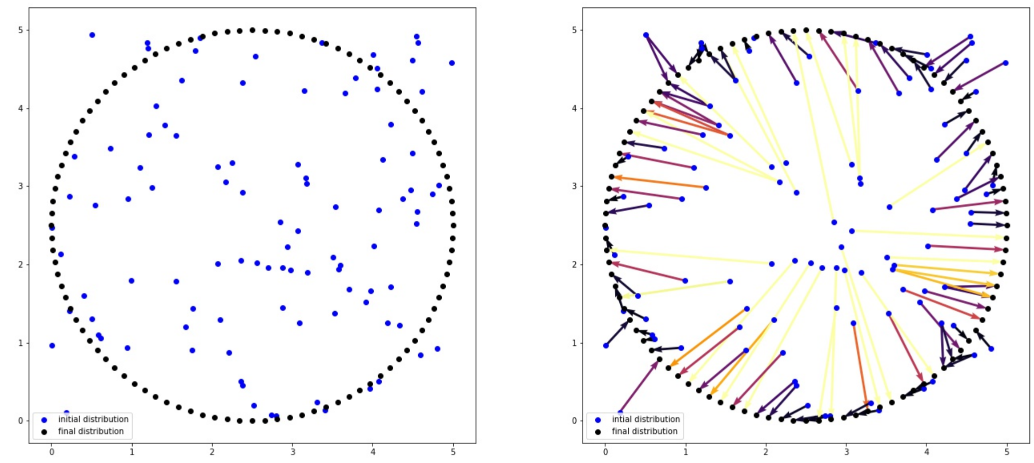


Figure 7: Cost of transportation visualized as a heatmap

Linear Swarm Transfer

These techniques scale up to two and three dimensional representations. Consider the two dimensional transfer problem of redistributing points along a circle. In this particular problem the mass at each point is the same, however the scale of the dots representing masses may be increased to visualize increased mass. The transfer network is represented by the directional arrows. A color gradient denotes the cost of transfer for each of the points, with lighter colors representing higher cost.



(a) Mass distribution visualization (b) Mass transfer network

Figure 8: Two dimensional swarm transfer

This particular example is the application of optimal transportation to particle swarm optimization. Consider this visualization applied to the problem of redistributing particles into an optimal configuration with control over the particle's velocity. The goal is to align all of the particles along the circle at the same time. Then the color of the arrows describing the cost to transport also indicates the necessary velocity to transfer along the ray to align at the appropriate time. These particles may be robots in a factory, swarms of drones or other redistributing particles.

This application may also be scaled to three dimensions. In the following problem particle swarms is redistributed (suppose they are drones) to a constant altitude, beginning from the same starting location on the ground.

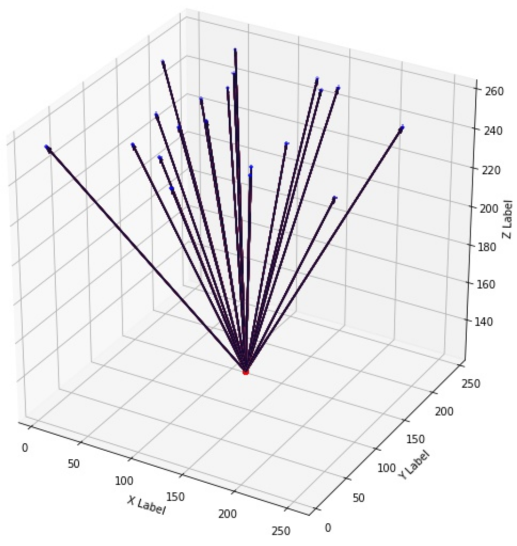


Figure 9: Three dimensional swarm transfer

This type of network is also useful in modern air traffic control. There has been significant developments towards implementing a new generation of air traffic control that reduce delays, holds and separation while maintaining aircraft safety. Aircraft travel along directed graphs between airborne waypoints. Thus the extension of this three dimensional visualization to a network of simple paths along a directed network can be utilized to simulate aircraft transportation.

N-Dimensional Redistribution

It is possible that the problem at hand is larger than three dimensions but visualization is still useful. In this instance the masses may be ordered and then transfer considered by transferring along the ordered masses. In the following example transfer is visualized along lattice points sufficiently spaced so that the network can be sufficiently visualized.

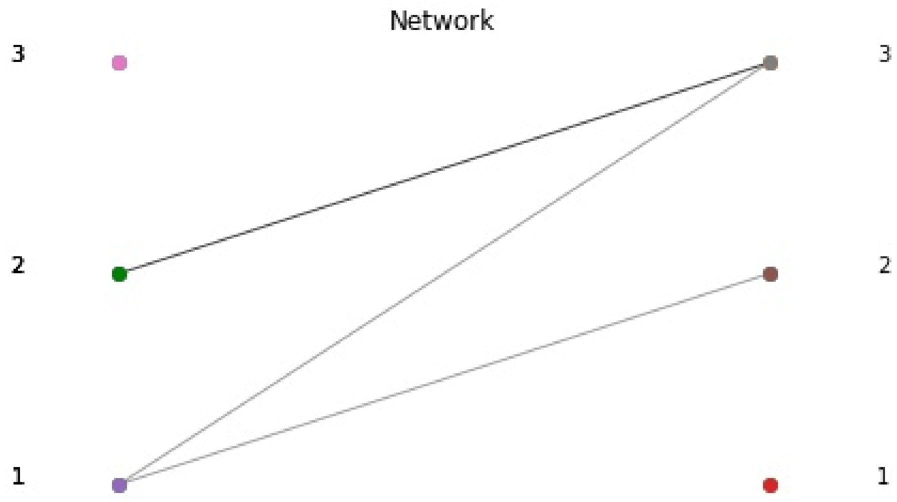


Figure 10: Transfer along a network between ordered nodes

This particular visualization is useful in modeling atmospheric conditions where the transfer is not only a transfer of space but also a transfer of states. In atmospheric studies the transportation of momentum, temperature trace gasses and water or ice crystals are considered along with spatial transfer in three dimensions. Nodes can be ordered into groups by state variables. This technique of visualization can help map out both transportation of states along with physical transportation.

Palette Swapping

Palette swapping is the process of pushing the color palette of one image onto the color palette of a target image. This process has applications in graphic design and photo editing, as well as steganography and object detection and tracking. The problem is formulated by creating a probability distribution of the pixel values in each image, and then pushing one distribution onto the other. Several different distributions may be considered. In the following example the distribution is taken in terms of the pixel values. Other image related swappings include image intensity for brightening or pixel variation for blending.

Consider the following target and palette image.

(a) Palette (b) Target

Figure 11: Initial images for palette swapping

To help reduce run time pixels are grouped into a k-nearest neighbors distribution. The k-means average of each pixel cluster is then the probability distribution pushed between each of the images. Here the clustered images are displayed where each original pixel is assigned to its clustered average.

(a) Palette (b) Target

Figure 12: Clustered images for palette swapping

Each unique pixel color in the reconstructed images are given a mass of one and the optimal transfer between them minimizes the three dimensional cost function $c(x, y) = .5|x - y|^2$ where x and y are pixel values in the palette and target images respectively.

Figure 13: Clustered target image with new palette

This is an example of direct transformation between each of the probability distributions. It may be expected that the butterflies wings remain orange. However, in this case the red of the flower and the orange of the butterfly wings are minimized by transferring orange to red and green to orange. These results can be changed by varying the amount of mass at each location, the number of pixels in the clustering and the cost function.

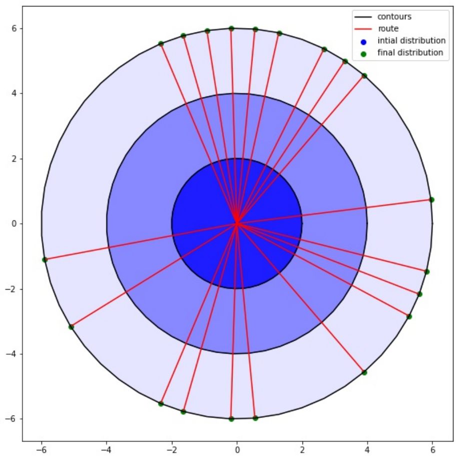
Contour Descent

Contour descent is an example of minimization over an undirected graph. The objective of contour descent is to minimize the distance traveled between two points along a contour map, while avoiding steep descents. The cost function is then given by

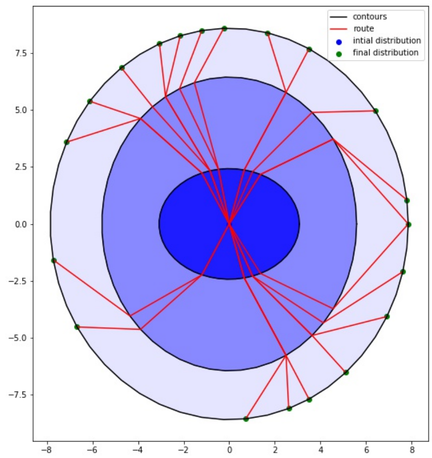
$$c(x_0, x_n) = \sum_{i=0}^{n-1} \alpha_i |x_i - x_{i+1}|^2$$

where α_i represents a variable cost between each of the contours. In the following example α scales with the darkness of the region, so that α is highest in the darkest regions.

Two variations of the problem are considered. In the first case the contours are exact circles. In this case, the minimal distance between the initial location (the center of the contours) and the target points on the boundary of the final contour is the direct line between the center and the boundary. However, if the contour shape is varied to form an elipse, then the masses begin to transfer in various patterns which are dependent on the shape of the contours and the value of α_i in that particular region. These contours can also be extended to any closed polygon



(a) Circular contours



(b) Eliptical contours

Figure 14: Comparison of contour descents by contour shape

Consider the application to descending a mountain slope where a rescue team wants to transfer an individual over as short of a distance as possible while avoiding steep terrain. This problem describes the best path for traversal. The value of α can be varied to correspond to the cost of traversing a particular domain.

Conclusion

The success of the Monge-Kantorovich problems modeling capability encourages further study. This paper extends the Monge-Kantorovich formulation of optimal transportation problem to a network of paths. We reformulate the problem and reduce it to the linear programming problem which allows us to efficiently compute the numerical solution This extends the applications of the Monge-Kantorovich formulation. These applications include but are not limited to, multi-step disaster relief planning, warehouse robotic movement along a restricted set of paths, and improved national air traffic control.

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank my mentor, Elena Cherkaev, for her guidance. I would also like to thank the Department of Mathematics and the Undergraduate Research Opportunities Program at the University of Utah for funding this REU project.

REFERENCES

1. Domahidi, Alexander, Eric Chu, and Stephen Boyd. *ECOS: An SOCP solver for embedded systems.* 2013 European Control Conference (ECC). IEEE, 2013.

2. Evans, L. C. Partial differential equations and Monge–Kantorovich mass transfer. In *Current developments in mathematics*, 1997 (Cambridge, MA). Int. Press, Boston, MA, 1999, pp. 65–126

3. Krishnan, Vishaal and Sonia Martínez *Distributed optimal transport for the deployment of swarms* 2018 (arXiv:1804.01572)

4. Pitié, François, and Anil Kokaram. *The linear monge-kantorovitch linear colour mapping for example-based colour transfer.* 2007 (Trinity Colegge Dublin, Irelnad)

5. Rockafellar, Ralph Tyrell. *Convex analysis.* Princeton university press, 2015.

6. Xia, Qinglan *Motivations, ideas, and applications of ramified optimal transportation.* ESAIM: Mathematical Modelling and Numerical Analysis. Volume 49, Number 6, November-December 2015. Special Issue - Optimal Transport. pp 1791 - 1832