UNIVERSITÀ DI PISA

Corso di Laurea Triennale in Matematica

Tesi di Laurea Triennale

Sinkhorn's algorithm for optimal transport problem

Relatore:
Dott. Dario Trevisan

Candidato:
Vittorio D'Onofrio

Anno Accademico 2017/2018

Contents

Introduction

In this thesis, we will study a method of approximation for the solution of an optimal transport (OT) problem.
Following [2], we will approximate the solution using an entropic constraint and show that the solution of the regularized problem approaches the solution of the optimal transport problem in the generic form.
Some properties of the Kullback-Leibler divergence (KL) allows to prove the existence and unicity of the solution of regularized OT problem. Moreover, the solution is a matrix with given row and column sums, diagonally equivalent to a given matrix.
Sinkhorn's theorem gaurantees that there exists a unique matrix satisfying these properties.
Hence, it's possible to compute the solution of regularized OT problem through Sinkhorn's algorithm.
In the first chapter we will define the entropy and the Kullback-Leibler divergence. An elementary lemma about KL will be proved in order to derive Sinkhorn's theorem.
In the second chapter we will study Sinkhorn's algorithm, that allows computing the solution of regularized OT problem.
Therefore, Hilbert's projective metric allows to define a metric on a suitable space of matrices. Completeness of this space and a bound given by Birkhoff-Hopf theorem are used to prove the convergence of Sinkhorn's algorithm.
In the third chapter we will define an optimal transport problem between two probability vectors. First, we will restrict the problem to a set of matrices with a bound on their KL divergence from a fixed matrix. Then, using Lagrange multipliers, we will turn the problem in a form depending on regularized entropy. Hence, we will prove that the regularized formulation is equivalent to minimize a Kullback-Leibler divergence.
Theorems proved in the previous chapters ensure the existence and unicity of the solution. Moreover, it's possible to compute the solution through Sinkhorn's algorithm. In the last section we will implement Sinkhorn's algorithm in MATLAB and show some results on 1-D marginals and 2-D

marginals that confirm theoretical analysis.

Notation

1. $\mathbb{R}^n_+ = \{x = (x_1, .., x_n) | x_i \geq 0 \ \forall i \leq n\}$.
2. For $x, y \in \mathbb{R}^n_+$ we write $x \leq y$ if $x_i \leq y_i \ \forall i = 1, \ldots, n$.
3. If $x \in \mathbb{R}$, $y \in \mathbb{R}^n$ and $y > 0$, denote $\dfrac{x}{y} = \left(\dfrac{x_1}{y_1}, \ldots, \dfrac{x_n}{y_n}\right)$.
4. $e = (1, .., 1)^T \in \mathbb{R}^n_+$.
5. Let $X \subset \mathbb{R}^n_+$. Let ρ defined on X such that $\forall x, y \in X \ x \rho y \Leftrightarrow x = \lambda y$, $\lambda > 0$. Denote $\mathbb{P}(X) = X/\rho$.
6. If $A, B \in \mathbb{R}^{d \times d}_+$, then $< A, B >= tr(A^T B)$ is the Frobenius dot product.
7. If $x \in \mathbb{R}^n$, $\|x\|$ is the Euclidean norm.
8. If $M \in \mathbb{R}^{n \times n}$, denote by $\exp(-\lambda M)$ the element-wise exponential.

Chapter 1

Sinkhorn's theorem

In this chapter we will define the entropy of a probability distribution and the Kullback-Leibler divergence (KL) of two probability distributions. We can fit these definitions on the set $U(r, c)$ defined in 1.1.
We will prove some properties of the entropy [3] and of the KL. Therefore, we will prove an elementary lemma (1.11) about existence and unicity of the minimum, satisfying a condition on its partial derivatives. Following [5], we will use this lemma to prove a theorem (1.13) about diagonal rescaling of a matrix with nonnegative entries. Theorem 1.13 allows to prove Sinkhorn's theorem.

1.1 Kullback-Leibler divergence

Definition 1.1. Let $\Sigma_d := \{x \in \mathbb{R}_+^d : xe = 1\}$. For two probability vectors r and c in Σ_d we define $U(r, c) := \{P \in \mathbb{R}_+^{d\times d} \,|Pe = r, P^T e = c\}$.

If X and Y are two random variables taking values in $\{1, ..., d\}$, each with distribution r and c respectively, $U(r, c)$ contains all possible joint probability distributions of (X,Y).

Definition 1.2. If Q is a discrete probability distribution on Ω we define the entropy of Q as

$$h(Q) = -\sum_{x\in\Omega} Q(x) \log Q(x)$$

Remark 1.3. If $P \in U(r, c)$ and $r \in \Sigma_d$ then

$$h(r) = -\sum_{i=1}^{d} r_i \log r_i \qquad h(P) = -\sum_{i,j=1}^{d} p_{ij} \log p_{ij}$$

.

Remark 1.4. $p \to h(p)$ is concave.

Definition 1.5. If P, Q are discrete probability distributions on Ω we define the Kullback-Leibler divergence of P, Q as

$$KL(P\|Q) = \sum_{x\in\Omega} P(x) \log \frac{P(x)}{Q(x)}$$

.

Remark 1.6. If $P, Q \in U(r, c)$ then

$$KL(P\|Q) = \sum_{i,j} p_{ij} \log(\frac{p_{ij}}{q_{ij}})$$

Remark 1.7. If $r, c \in \mathbb{R}^n_+$ are two probability vectors, then

$$KL(r\|c) = \sum_{i=1}^{n} r_i \log \frac{r_i}{c_i}$$

Proposition 1.8. *If P, Q are discrete probability distributions then $KL(P\|Q) \geq 0$ with equality iff $P = Q$.*

Proof. Using Jensen's inequality

$$-KL(P\|Q) = \sum_{x\in\Omega} P(x) \log \frac{Q(x)}{P(x)} \leq \log \sum_{x\in\Omega} P(x) \frac{Q(x)}{P(x)} =$$
$$\log \sum_{x\in\Omega} Q(x) = \log 1 = 0$$

with equality iff $P = Q$ by the strict concavity of the logarithm. □

Proposition 1.9. *If $p, q \in \Sigma_d$ and q has uniform density then $KL(p\|q) = -h(p) + \log(d)$. Therefore $\log d \geq h(p)$*

Proof.

$$KL(p\|q) = \sum_i p_i \log(\frac{p_i}{q_i}) = -h(p) - \sum_i p_i \log q_i$$
$$= -h(p) + \log(d)$$

.

Using 1.8, $\log d \geq h(p)$. □

Proposition 1.10. *Let C be an $m \times n$ real matrix. Then $Im(C^T) = Ker(C)^\perp$.*

Proof. For $x, y \in \mathbb{R}^n$ denote $\phi(x, y)$ the standard dot product.
$Im(C^T) \subseteq Ker(C)^\perp$: $\forall z \in Ker(C), \forall x \in \mathbb{R}^n$ we have $\phi(C^T x, z) = \phi(x, Cz) = 0$. Hence $C^T x \in Ker(C)^\perp$.
$Ker(C)^\perp \subseteq Im(C^T)$: it's equivalent to show that $Im(C^T)^\perp \subseteq Ker(C)^{\perp\perp} = Ker(C)$. If $x \in Im(C^T)^\perp$, then $\forall y \in \mathbb{R}^n$ we have $0 = \phi(x, C^T y) = \phi(Cx, y)$. Since this must be true $\forall y \in \mathbb{R}^n$, it's also true for y choosen in the canonic basis of $\mathbb{R}^n$. Then $Cx = 0$ and $x \in Ker(C)$. □

Theorem 1.11. *Let C be an $m \times n$ real matrix. Let b lie in $\mathbb{R}^n$.Assume $Cy = b$ for some $y > 0$. Assume $x \in \mathbb{R}^n_+, x > 0, \sum_{j=1}^{n} x_j = 1$. There exists a unique $u^0 \in \mathbb{R}^n_+$ such that*

$$KL(u^0 \| x) = \min\{KL(u\|x) : u \in \mathbb{R}^n_+, \textstyle\sum u_j = 1, Cu = b\}.$$

Necessarily $u^0 > 0$ and u^0 is the unique point such that $u > 0, Cu = b$ and

$$\frac{\partial}{\partial u_j}[KL(u\|x) - q^T Cu]|_{u=u^0} = 0 \text{ for some } q \in \mathbb{R}^m.$$

The vector q is unique apart from increments ω satysfing $\omega^T C = 0$.

Proof. Existence For $u_j \geq 0$ the function $u_j \log \dfrac{u_j}{x_j}$ is continuous and attains a finite minimum value at $u_j = e^{-1}x_j$; so the sum $KL(u\|x)$ is continuous on $\mathbb{R}^n_+$. The set

$$S = \{u \in \mathbb{R}^n_+, \textstyle\sum_j u_j = 1, Cu = b\}$$

is non-empty because $y \in S$ bounded and closed in $\mathbb{R}^n_+$ so $KL(u\|x)$ attains a finite minimum value u^0 in S.
Positivity. If $0 < \varepsilon < 1$ we define the positive vector $u(\varepsilon) = (1 - \varepsilon)u^0 + \varepsilon y$ and the sets of indices

$$J_0 = \{j : u^0_j = 0\}, \quad J_1 = \{j : u^0_j > 0\}.$$

Then

$$\frac{d}{d\varepsilon} KL(u(\varepsilon)\|x) = \sum_{j=1}^{n} (y_j - u^0_j)(1 + log\frac{u_j(\varepsilon)}{x_j})$$

If J_0 is not empty, as $\varepsilon \to 0$

$$\frac{d}{d\varepsilon} KL(u(\varepsilon)\|x) = (\sum_{J_0} y_j) \log \varepsilon + O(1)$$

which would tend to $-\infty$. So for small ε, $KL \circ u$ is decreasing compared to ε. So this would imply $KL(u(\varepsilon)\|x) < KL(u(0)\|x) = KL(u^0\|x)$ contradicting the minimizing property of u^0.
Uniqueness. Suppose u^1 also minimizes KL under the constraints $u \geq 0$, $Cu = b$. By positivity $u^1 \geq 0$.For $0 \leq \theta \leq 1$ define

$$u(\theta) = (1-\theta)u^0 + \theta u^1$$

If $u^0 \neq u^1$, then

$$\frac{d^2}{d\theta^2} KL(u(\theta)\|x) = \sum_{j=1}^{n} \frac{(u_j^0 - u_j^1)^2}{u_j(\theta)} > 0$$

So $KL \circ u$ is strictly convex with respect to θ and $KL(u^0\|x) = KL(u(0)\|x) > KL(u(\frac{1}{2})\|x)$ contradicting tha fact that u^0 minimizes KL. Therefore, $u^0 = u^1$ and the minimizing u is unique.
Lagrange multipliers. Since $u^0 > 0$, if z is fixed in $\mathbb{R}^n$, then $u^0 + \varepsilon z > 0$ for all sufficiently small ε. If $Cz = 0$ then

$$\frac{d}{d\varepsilon} KL(u^0 + \varepsilon z\|x) = 0 \quad \text{at } \varepsilon = 0,$$

which says that is orthogonal to the gradient of KL at $u = u^0$. Since this must be true for all z in $Ker(C)$, using Proposition 1.10, we have $\nabla KL(u^0\|x) \in Im(C^T)$ and so there exists a vector $q \in \mathbb{R}^m$ such that

$$(\nabla KL(u^0\|x))^T = q^T C$$

This satisfies Lagrange equation $\dfrac{\partial}{\partial u_j}[KL(u\|x) - q^T Cu]|_{u=u^0} = 0$. The vector q is unique apart from increments ω satisfying $\omega^T C = 0$.
Uniqueness. Suppose $u^1 > 0, Cu = b$,$u^1 \neq u^0$ and

$$(\nabla KL(u^1\|x))^T = q^T C$$

By the uniqueness of the minimum, $KL(u^1\|x) > KL(u^0\|x)$. The convexity

of $KL(u\|x)$ now implies

$$\frac{d}{d\varepsilon}KL((1-\varepsilon)u^1+\varepsilon u^0\|x)<0 \quad at \quad \varepsilon=0$$

which says

$$(\nabla KL(u\|x))^T(u^0-u^1)<0 \quad at \quad u=u^1.$$

Now $(\nabla KL(u\|x))^T=q^TC$ implies

$$q^TC(u^0-u^1)<0$$

which is absurd, because $C(u^0-u^1)=b-b=0$. □

1.2 Sinkhorn's theorem

Theorem 1.12. *Let $C=(c_{ij})$ be a real $m\times n$ matrix. Let $b\in\mathbb{R}^n-\{0\}$. Let $K=\{\pi : C\pi=b, \pi\geq 0\}$. Let x and y be two nonnegative vectors with the same zero pattern ($x_i=0\Leftrightarrow y_i=0$). If $y\in K$ then there exists a unique π in K such that*

$$\pi_j=x_j\prod_{i=1}^m z_i^{c_{ij}} \quad j=1,...,n$$

for some $z_i>0$.

Proof. If $x_j=0$ we set $\pi_j=0$, thus without loss of generality we may assume that all components x_j and y_j are positive for $j=1,..,n$.
According to Theorem 1.11 there exists a unique u^0 achieving

$$\min_u\{\sum_{j=1}^n u_j\log\frac{u_j}{x_j} : u\geq 0, \frac{1}{e}Cu=b\}$$

where $u^0>0, \frac{1}{e}Cu^0=b$ and

$$1+\log\frac{u_j^0}{x_j}=\sum_{i=1}^m q_ic_{ij} \quad j=1,..,n.$$

Taking exponentials, we find

$$\pi_j=x_j\prod_{i=1}^m z_i^{c_{ij}} \quad j=1,..,n$$

where $\pi_j=eu_j^0$ and $z_i=\exp(\frac{1}{e}q_i)$. □

Theorem 1.13. *Let $X = (x_{ij})$, $Y = (y_{ij})$ be two $r \times s$ matrices with non-negative entries. Let $x_{ij} = 0 \Leftrightarrow y_{ij} = 0$ for any i,j. Let the row sums and column sums of Y be positive. Then there exist $u_1, ..u_r$, $v_1, ..v_s$ all positive such that $\pi_{ij} = (x_{ij}u_iv_j)$ has the same row sums and column sums as Y.*

Proof. Let $e^s = [1, .., 1] \in \mathbb{R}^s$ and $e^r = [1, .., 1] \in \mathbb{R}^r$.Let s^j be the vector of row sums: $s^j = C(e^s)^T$, so $s^j \in \mathbb{R}^r$ and s^i be the vector of row sums $s^i = C^T(e^r)^T$, $s^i \in \mathbb{R}^s$.
Let C be the matrix in $\mathbb{R}^{m\times n}$ defined as

$$C = \begin{bmatrix} e^s & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & e^s & 0 \\ 0 & \dots & 0 & e^s \\ e^r & 0 & \dots & 0 \\ 0 & e^r & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & 0 & e^r \end{bmatrix}$$

where $m = r + s$ and $n = 2rs$.
The condition on the row and column sums of Y can be expressed in the form $Cy = b$ where $y = \begin{bmatrix} Y_1 & \dots & Y_r & Y^1 & \dots & Y^s \end{bmatrix}^T$ $y \in \mathbb{R}^n$;
$b = \begin{bmatrix} s^j & s^i \end{bmatrix}^T$ and $b \in \mathbb{R}^m$.
If $z \in \mathbb{R}^n$ it's possible to write $z = \begin{bmatrix} z_1 & \dots & z_n \end{bmatrix}^T$ where
$z_1 = \begin{bmatrix} z_{11} & \dots & z_{1s} \end{bmatrix}$,..., $z_r = \begin{bmatrix} z_{r1} & \dots & z_{rs} \end{bmatrix}$
$z_{r+1} = \begin{bmatrix} z_{11} & \dots & z_{r1} \end{bmatrix}$,..., $z_n = \begin{bmatrix} z_{1s} & \dots & z_{rs} \end{bmatrix}$.
For any $q \in \mathbb{R}^m$ we may write

$$q^T = \begin{bmatrix} \alpha_1 & \dots & \alpha_r & \beta_1 & \dots & \beta_s \end{bmatrix}$$

So for all z in $\mathbb{R}^n$ we have

$$q^TCz = \sum_{i=1}^{r}\sum_{j=1}^{s}(\alpha_i + \beta j)z_{ij}$$

Now we consider for all $z \in \mathbb{R}^n$

$$KL(z\|x) = \sum_{v=1}^{n} z_v \log \frac{z_v}{x_v}$$

where x has the same structure as z.
According to Theorem 1.11 there exists a unique z^0 in $\mathbb{R}^n$ such that

$$KL(z^0\|x) = \min\{KL(z\|x) : z \in \mathbb{R}^n, Cz = b\}$$

and satisfying Lagrange equation

$$\frac{\partial}{\partial z_j}[KL(z\|x) - q^T Cz] = 0 \quad j = 1, .., n.$$

It's possible to write Lagrange equation in the form

$$1 + \log \frac{z^0_{ij}}{x_{ij}} - (\alpha_i + \beta_j) = 0$$

Taking exponentials, we find

$$ez^0_{ij} = x_{ij} e^{\alpha_i} e^{\beta_j}$$

Now define

$$\pi_{ij} = ez^0_{ij}, \quad u_i = e^{\alpha_i}, \quad v_j = e^{\beta_j}$$

and we obtain the theorem. □

Theorem 1.14. *Let $A = (a_{ij})$ be a $N \times N$ matrix with $a_{ij} > 0$ $\forall i, j$. Let $p, q \in \mathbb{R}^N$. There exists exactly one matrix $\widehat{B}$ such that the row and column sums are respectively p and q and it can be expressed in the form $\widehat{B} = D_1 A D_2$ where D_1 and D_2 are diagonal matrices with positive diagonals. D_1 and D_2 are unique up to a scalar factor.*

Proof. Let Y be a matrix with positive entries such that $Ye = p$, $Y^T e = q$.By Theorem 1.13, there exist u_i, v_j such that the matrix (π_{ij}) defined as

$$\pi_{ij} = a_{ij} u_i v_j$$

has the same row sums and column sums as Y.
Then, setting $\widehat{B} = (\pi_{ij})$, $D_1 = \text{diag}(u_i)$, $D_2 = \text{diag}(v_j)$, we get the theorem. □

Definition 1.15. A doubly stochastic matrix is a square matrix $A = (a_{ij})$ of nonnegative real numbers each of whose rows and column sums to 1, i.e

$$\sum_j a_{ij} = \sum_i a_{ij} = 1 \qquad \forall i, j$$

Corollary 1.16 (Sinkhorn's theorem)**.** *To a given $N \times N$ matrix $A = (a_{ij})$ with $a_{ij} > 0$ $\forall i, j$, there corresponds exactly one doubly stochastic matrix T_A which can be expressed in the form $T_A = D_1 A D_2$ where D_1 and D_2 are diagonal matrices with positive diagonals. The matrices D_1 and D_2 are unique up to a scalar factor.*

Proof. Using the notation of the previous theorem, we can set $Y = (y_{ij})$ with $y_{ij} = \dfrac{1}{N}$ for all i, j.
By this choice of Y, we have that $T_A = (\pi_{ij})$ is doubly stochastic. □

Chapter 2

Sinkhorn's algorithm

In this chapter we will discuss Sinkhorn's algorithm in order to compute the unique matrix diagonally equivalent to a given matrix with prescribed row and column sums. The existence and unicity of such a matrix was proved in Chapter 1. We will use Hilbert's projective metric d (2.2) to define a metric μ (2.6) on the set $\mathbb{P}(E_A)$ of matrices diagonally equivalent to a given matrix. We will show that the space $\big(\mathbb{P}(\mathbb{R}^n_+), d\big)$ is complete [1].This result allows to show the completeness of $(\mathbb{P}(E_A), \mu)$.
We will prove Birkhoff-Hopf theorem in 2×2 case using Sinkhorn's theorem and then, following [4], we will extend it to the $m \times n$ case. This theorem gives a bound on the contraction ratio defined in 2.10.
We will use this bound and the completeness of $(\mathbb{P}(E_A), \mu)$ to prove the convergence of Sinkhorn's algorithm.

2.1 Hilbert's projective metric

Definition 2.1. If $x, y \in \mathbb{R}^n_+$ we define $M(x/y) = \max\limits_i \dfrac{x_i}{y_i}$ and $m(x/y) = \min\limits_i \dfrac{x_i}{y_i}$

Definition 2.2. Hilbert's projective metric $d(,)$ is defined on $\mathbb{R}^n_+$ by

$$d(x, y) = \log \frac{M(x/y)}{m(x/y)} = \log \max_{i,j} \frac{x_i y_j}{x_j y_i}$$

Proposition 2.3. *$\big(\mathbb{R}^n_+, d\big)$ is a pseudo-metric space and $\big(\mathbb{P}(\mathbb{R}^n_+), d\big)$ is a metric space.*

Proof. Let $x, y, z \in \mathbb{R}^n_+$.It's obvious that $d(x, y) = d(y, x)$. Since $M(x/y) \geq m(x/y)$ we have $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = \alpha y$ where $\alpha > 0$.

We have

$$\max \frac{x_i}{x_j}\frac{y_j}{y_i} \max \frac{x_i}{x_j}\frac{z_j}{z_i}\frac{z_i}{z_j}\frac{y_j}{y_i} \leq \max \frac{x_i z_j}{x_j z_i} \max \frac{z_i y_j}{z_j y_i}.$$

So $d(x,y) \leq d(x,z) + d(z,y)$. If $x, y \in \mathbb{P}(\mathbb{R}^n_+)$ then $d(x,y) = 0 \Leftrightarrow x = y$ so $(\mathbb{P}(\mathbb{R}^n_+), d)$ is a metric space. □

Proposition 2.4. *$E = \big(\mathbb{P}(\mathbb{R}^n_+), d\big)$ is complete.*

Proof. If $x, y \in \mathbb{R}^n_+$ and $x \neq \lambda y \ \forall \lambda > 0$, we will show that

$$\|x - y\| \leq \exp(d(x,y)) - 1.$$

If $x, y \in \mathbb{R}^n_+$ we have

$$m(x/y) \leq 1 \leq M(x/y)$$

Therefore,

$$\begin{aligned}\|x - y\| = \{\sum_i (x_i - y_i)^2\}^{1/2} &\leq \{\sum_i [M(x/y) - m(x/y)]^2 y_i^2\}^{1/2} \\ &\leq M(x/y) - m(x/y) \leq (exp(d(x,y)) - 1) m(x/y)\end{aligned}$$

Moreover,

$$M(x/y) \leq 1 + \frac{\|x - y\|}{m(y/e)}$$

Similarly, if $\|x - y\| \leq m(y/e)$

$$m(x/y) \geq 1 - \frac{\|x - y\|}{m(y/e)}$$

It follows that

$$\|x - y\| \leq m(y/e) \tanh(\frac{1}{2} d(x,y)).$$

Let $p : \mathbb{R}^n_+ \to \mathbb{P}(\mathbb{R}^n_+)$ be the natural projection. So if $\{x_k\}$ is a Cauchy sequence in E, then $\{p^{-1}(x_k)\}$ is a Cauchy sequence in $\{\mathbb{R}^n_+, \| \ \|\}$ and hence converges to a limit $p^{-1}(x)$ with $d(x_k, x) \to 0$ and so x_k converges to x in E. □

Definition 2.5. Let A be a positive $m \times n$ matrix. We denote $A \sim B$ if there exist X and Y diagonal matrices such that $A = XBY$. Denote

$$E_A = \{B : A \sim B, b_{ij} > 0\}$$

the set of matrices diagonally equivalent to A.

Definition 2.6. If $B, B' \in E_A, \quad B = XB'Y, \quad X = \text{diag}(x_i), \quad Y = \text{diag}(y_j)$ we define

$$\mu(B, B') = d(x, e) + d(y, e)$$

Remark 2.7. μ is well defined because by Sinkhorn's theorem X and Y are unique up to a scalar factor.

Proposition 2.8. *(E_A, μ) is a pseudometric space and $(\mathbb{P}(E_A), \mu)$ is a metric space.*

Proof. Let $B, B' \in E_A, B = XB'Y$ with $X = \text{diag}(x)$ and $Y = \text{diag}(y)$. Since $d(,) \geq 0$, $\mu(B, B') = d(x, e) + d(y, e) \geq 0$. $\mu(B, B') = 0$ iff $d(x, e) = -d(y, e)$; by positivity of d we obtain $d(x, e) = d(y, e) = 0$ so $x = \lambda e$ and $y = te$ and $B = \lambda t B'$. We also have $\mu(B', B) = d(\frac{e}{x}, e) + d(\frac{e}{y}, e) = d(x, e) + d(y, e) = \mu(B, B')$.
If $B'' = X''BY''$ with $X'' = \text{diag}(x'')$ and $Y'' = \text{diag}(y'')$ we have $\mu(B, B') = d(x, e) + d(y, e) \leq d(x, x'') + d(x'', e) + d(y, y'') + d(y'', e) = \mu(B, B'') + \mu(B'', B')$. □

Proposition 2.9. *The space $\mathbb{P}(E_A)$ with the metric μ is complete.*

Proof. Let $\{A_k\}_k$ be a Cauchy sequence in $\mathbb{P}(E_A)$. For all $i, j \ A_i \sim A_j$.
So there exist X_i, Y_j such that

$$A_2 = X_1 A_1 Y_1$$

$$A_3 = X_2 A_2 Y_2$$

$$\vdots$$

$$A_{n+1} = X_n A_n Y_n = X_n \dots X_1 A_1 Y_1 \dots Y_n$$

Let $\{x^i\} \subset \mathbb{R}^n_+$ be such that $X_i \dots X_1 = \text{diag}(x^i)$.
Similarly, let $\{y^i\} \subset \mathbb{R}^m_+$ be such that $Y_i \dots Y_1 = \text{diag}(y^i)$.
Since A_k is a Cauchy sequence there exist N such that $\forall \varepsilon > 0 \quad \forall m, n > N$

$$\mu(A_m, A_n) < \varepsilon$$

Suppose $n > m$, so $A_n = X_n \dots X_m A_m Y_m \dots Y_n$ and

$$\mu(A_n, A_m) = d(\frac{x^n}{x^{m-1}}, e) + d(\frac{y^n}{y^{m-1}}, e) = d(x^n, x^{m-1}) + d(y^n, y^{m-1}) < \varepsilon$$

Since $d(,)$ is a distance we have $d(,) \geq 0$ and so

$$d(x^n, x^{m-1}) < \varepsilon \quad d(y^n, y^{m-1}) < \varepsilon$$

Hence, $\{x^i\}$ and $\{y^i\}$ are Cauchy sequences in $(\mathbb{P}(\mathbb{R}^m_+), d)$; since this space is complete

$$x^i \to x \quad y^i \to y$$

Then $\forall m > N$ let

$$B = \operatorname{diag}(\frac{x^m}{x}) A_m \operatorname{diag}(\frac{y^m}{y})$$

We have $B \in \mathbb{P}(E_A)$ and

$$\mu(B, A_m) = d(\frac{x^m}{x}, e) + d(\frac{y^m}{y}, e) = d(x^m, x) + d(y^m, y) < 2\varepsilon$$

So the sequence A_m converges to B in $\mathbb{P}(E_A)$ with the metric μ. □

2.2 Birkhoff-Hopf theorem

Definition 2.10. Given an $m \times n$ matrix A with positive entries, we define

1. $\Delta(A) = \sup\{d(Ay, Ay') | y, y' \in \mathbb{R}^n_+\}$; it measures the diameter of the image.
2. $\kappa(A) = \sup\{\dfrac{d(Ay, Ay')}{d(y, y')} : y, y' \in \mathbb{R}^n_+, y' \neq \alpha y\}$ denote the contraction ratio of A.
3. If $x, y \in \mathbb{R}^n$, $\omega(x/y) = \max_i \dfrac{x_i}{y_i} - \min_i \dfrac{x_i}{y_i}$
4. $N(A) = \sup\{\dfrac{\omega(Ay/Ax)}{\omega(y/x)}, x, y \in \mathbb{R}^n\}$ is the Hopf oscillation ratio

Remark 2.11. If A and B are two $m \times n$ matrices diagonally equivalent with positive entries, i.e., there are two diagonal matrices X and Y with positive diagonal entries such that $B = XAY$, then $\Delta(A) = \Delta(B)$

Remark 2.12. Let A_1 be an $m \times n$ matrix and A_2 be an $n \times m$ matrix. Since $\forall y, y' \in \mathbb{R}^m_+$,

$$d(A_1A_2y, A_1A_2y') \leq \kappa(A_1)d(A_2y, A_2y') \leq \kappa(A_1)\kappa(A_2)d(y, y')$$

then

$$\kappa(A_1A_2) \leq \kappa(A_1)\kappa(A_2)$$

.

Proposition 2.13. *Let A be an $m \times n$ matrix. Then $\Delta(A) = \Delta(A^T)$.*

Proof. Let $y, y' \in \mathbb{R}^n_+$. If e_j are standard basis vector of $\mathbb{R}^n$ then $y = \sum_{s=1}^{n} \lambda_s e_s$; $y' = \sum_{t=1}^{n} \mu_t e_t$.

$$\begin{aligned}\Delta(A) &= \max_{y,y' \in \mathbb{R}^n_+} \log \max_{i,j} \frac{(Ay)_i (Ay')_j}{(Ay)_j (Ay')_i} \\ &= \log \max_{i,j,k,l} \frac{(A\lambda_k e_k)_i (A\mu_l e_l)_j}{(A\lambda_k e_k)_j (A\mu_l e_l)_i} \\ &= \log \max_{i,j,k,l} \frac{(Ae_k)_i (Ae_l)_j}{(Ae_k)_j (Ae_l)_i} \\ &= \log \max_{i,j,k,l} \frac{a_{ik} a_{jl}}{a_{jk} a_{il}}.\end{aligned}$$

Using this equality, $\Delta(A) = \Delta(A^T)$. □

Theorem 2.14. *Given a matrix* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *with positive entries and* $\det(A) \neq 0$, *there exists a matrix* $A' = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$ *where* $\alpha > 1$ *with* $\kappa(A) = \kappa(A')$, $\Delta(A) = \Delta(A')$ *and* $N(A) = N(A')$.

Proof. According to Sinkhorn's theorem there exist diagonal matrices D_1 and D_2 with positive diagonal elements such that $D_1 A D_2$ is doubly stochastic so

$$D_1 A D_2 = \begin{bmatrix} \beta & 1-\beta \\ 1-\beta & \beta \end{bmatrix}$$

Now, if $\det(A) > 0$ let P be the identity matrix and if $det(A) < 0$ let P be the permutation matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

($det(A) \neq 0$) by hypothesis, so

$$P D_1 A D_2 = \begin{bmatrix} \gamma & 1-\gamma \\ 1-\gamma & \gamma \end{bmatrix}$$

where $\gamma > 1/2$. Finally, let D be $1/1-\gamma$ times the identity matrix, so

$$A' = P D_1 A D_2 D = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$$

Using Remark 2.11 and the fact that P and D are bijections in $\mathbb{R}^2$ we have $N(A) = N(A')$, $\Delta(A) = \Delta(A')$ and so $\kappa(A) = \kappa(A')$

□

Proposition 2.15. *Let A be the matrix*

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$$

where $\alpha > 1$. Then $\Delta(A) = d(Ae_1, Ae_2)$ where e_1 and e_2 are the standard basis vectors of $\mathbb{R}^2$.

Proof. Assume $v = e_1 + se_2$, $v' = te_1 + e_2$ where $s, t \geq 0$ and $d(Av, Av') = \Delta(A)$. We want to show $s = t = 0$.
Therefore,

$$Av = \begin{bmatrix} \alpha + s \\ 1 + \alpha s \end{bmatrix} \quad Av' = \begin{bmatrix} \alpha t + 1 \\ t + \alpha \end{bmatrix}$$

Assume without loss of generality

$$\frac{\alpha + s}{\alpha t + 1} > \frac{1 + \alpha s}{t + \alpha}$$

Define

$$f(s, t) = \frac{\alpha + s}{1 + s\alpha} \frac{t + \alpha}{\alpha t + 1}$$

Since $\alpha > 1$,

$$\frac{\alpha + s}{1 + s\alpha} \leq \alpha \qquad \frac{t + \alpha}{\alpha t + 1} \leq \alpha.$$

Hence, $f(s, t) \leq \alpha^2$ and then $s = t = 0$. □

Proposition 2.16. *Let A be the matrix*

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$$

where $\alpha > 1$. Then $N(A) = \kappa(A) = \tanh \frac{1}{4}\Delta(A) = \dfrac{\exp(\Delta(A)/2) - 1}{\exp(\Delta(A)/2) + 1}$

Proof. By the definition

$$N(A) = \sup_{x,y \in \mathbb{R}^2} \frac{w(Ax/Ay)}{w(x/y)}$$

We can consider $x = (1, s)$ and $y = (1, t)$ for non-negative s,t. For $\omega(x/y)$ to exist and be non zero we must have $s, t > 0$ and $s \neq t$.
Using the definition of ω we have

$$N(A) = \sup_{s,t} \frac{\left| \dfrac{\alpha + s}{\alpha + t} - \dfrac{1 + \alpha s}{1 + \alpha t} \right|}{\left| 1 - \dfrac{s}{t} \right|}$$

$$= \sup_{s,t} \frac{(\alpha^2 - 1)t}{(\alpha + t)(\alpha t + 1)}$$
$$= \sup_{t} \phi(t)$$

where

$$\phi(t) = \frac{(\alpha^2 - 1)t}{(\alpha + t)(\alpha t + 1)}$$

ϕ is non-negative, its only stationary point is at 1 and its limits at 0 and ∞ are both 0; it follows that its supremum is attained at 1 and is equal to

$$\phi(1) = \frac{\alpha^2 - 1}{(\alpha + 1)^2} = \frac{\alpha - 1}{\alpha + 1}$$

A similar approach for $\kappa(A)$ gives

$$\kappa(A) = \sup_{s,t} \left| \frac{\log \dfrac{(\alpha + s)(1 + \alpha t)}{(\alpha + t)(1 + \alpha s)}}{\log \dfrac{s}{t}} \right|$$

$$= \sup_{s,t} |\psi(s, t)|$$

where

$$\psi(s, t) = \frac{\log \dfrac{(\alpha + s)(1 + \alpha t)}{(\alpha + t)(1 + \alpha s)}}{\log \dfrac{s}{t}}$$

We can write

$$\psi(s, t) = \frac{f(s) - f(t)}{\log s - \log t}$$

where

$$f(t) = \log \frac{\alpha + t}{1 + \alpha t}$$

Using mean value theorem, we have that for $0 < s < t$ there exists τ with $s \leq \tau \leq t$

$$\psi(s, t) = \frac{f(s) - f(t)}{\log s - \log t}$$
$$= f'(\tau)\tau$$
$$= \frac{(1 - \alpha^2)\tau}{(\alpha + \tau)(1 + \alpha\tau)}$$

$$= -\phi(t).$$

From this we obtain $\sup|\psi| \leq \sup\phi$. To show the opposite inequality, fix $t > 0$ and choose an arbitrary positive ε. Using mean value theorem, $-\psi(t-\varepsilon, t+\varepsilon) = \phi(\tau)$ with $\tau \in [t-\varepsilon, t+\varepsilon]$. By the continuity of ϕ we have that $|\psi|$ attains values arbitrarily close to $\phi(t)$ for any given t.
Hence, we have

$$N(A) = \kappa(A) = \frac{\alpha - 1}{\alpha + 1}$$

Let e_1 and e_2 be the standard basis vectors for $\mathbb{R}^2$. By Proposition 2.15

$$\Delta(A) = d(Ae_1, Ae_2)$$

Since $Ae_1 = (\alpha, 1)$ and $Ae_2 = (1, \alpha)$, we have $d(Ae_1, Ae_2) = \log\alpha^2 = 2\log\alpha$. Now $\alpha = \exp(\Delta(A)/2)$ and

$$N(A) = \kappa(A) = \frac{\alpha - 1}{\alpha + 1} = \frac{\exp(\Delta(A)/2) - 1}{\exp(\Delta(A)/2) + 1} = \tanh\frac{1}{4}\Delta(A)$$

□

Corollary 2.17. *If $A \in \mathbb{R}^{2\times 2}_+$ with positive entries and $\det(A) \neq 0$ then*

$$\kappa(A) = N(A) = \tanh\frac{1}{4}\Delta(A) = \frac{\exp(\Delta(A)/2) - 1}{\exp(\Delta(A)/2) + 1}$$

Proof. Using the previous theorems, there exists a matrix A' such that $\kappa(A) = \kappa(A')$, $N(A) = N(A')$, $\Delta(A) = \Delta(A')$ and

$$\kappa(A') = N(A') = \frac{\exp(\Delta(A')/2) - 1}{\exp(\Delta(A')/2) + 1} = \tanh\frac{1}{4}\Delta(A')$$

□

We now extend the previous result to a generic $m \times n$ matrix with positive entries:

Theorem 2.18 (Birkhoff-Hopf)**.** *Let $A \in \mathbb{R}^{m\times n}_+$ with positive entries. Then*

$$\kappa(A) = N(A) = \frac{\exp(\Delta(A)/2) - 1}{\exp(\Delta(A)/2) + 1} = \tanh\frac{1}{4}\Delta(A)$$

Proof. If $x, y \in \mathbb{R}^n_+$ define $V(x, y) = \{\alpha x + \beta y; \alpha, \beta \in \mathbb{R}\}$
Now define the functions k, N and Δ on $\mathbb{R}^n_+$ by

$$\kappa(x, y) = \sup\{\frac{d(Av, Aw)}{d(v, w)} | v, w \in V(x, y)\}$$

$$N(x,y) = \sup\{\frac{\omega(Av,Aw)}{\omega(v,w)}|v,w \in V(x,y)\}$$
$$\Delta(x,y) = \sup\{d(Av,Aw)|v,w \in V(x,y)\}$$

Since we proved the theorem in dimension 2 we have

$$\kappa(x,y) = N(x,y) = \tanh\frac{1}{4}\Delta(x,y) = \frac{\exp(\Delta(x,y)/2)-1}{\exp(\Delta(x,y)/2)+1}$$

By the definitions it follows that

$$\kappa(A) = \sup\{\kappa(x,y)|x,y \in \mathbb{R}^n_+\}$$
$$N(A) = \sup\{N(x,y)|x,y \in \mathbb{R}^n_+\}$$
$$\Delta(A) = \sup\{\Delta(x,y)|x,y \in \mathbb{R}^n_+\}$$

Now

$$\begin{aligned}\kappa(A) &= \sup\{\kappa(x,y)|x,y \in \mathbb{R}^n_+\}\\ &= \sup\{N(x,y)|x,y \in \mathbb{R}^n_+\}\\ &= N(A)\\ &= \sup\{\tanh\frac{1}{4}\Delta(x,y)|x,y \in \mathbb{R}^n_+\}\\ &= \tanh(\sup\{\frac{1}{4}\Delta(x,y)|x,y \in \mathbb{R}^n_+\})\\ &= \tanh\frac{1}{4}\Delta(A)\end{aligned}$$

□

Remark 2.19. $\kappa(A) < 1$. Using 2.13 $\kappa(A) = \kappa(A^T)$.

2.3 Sinkhorn's Algorithm

The aim of Sinkhorn's iterative algorithm is to find a positive matrix $\widehat{B}$ of the form D_1AD_2 which has prescribed row and column sums. The existence of the matrix $\widehat{B}$ was proved in the first chapter.
Suppose $A = (a_{ij})$ is a positive $m \times n$ matrix and $p \in \mathbb{R}^m$, $q \in \mathbb{R}^n$ with

$$p_1 + ... + p_m = q_1 + ... + q_n.$$

Starting with $A_0 = A$ it's possible to define the sequences A_k, A'_k of column and row normalized matrices by the following algorithm.
If $r^{(k)} = A_k e$ is the vector of row sums, define

$$A'_k = S_k A_k, \qquad S_k = \operatorname{diag}(\frac{p_i}{r_i^{(k)}})$$

If $c^{(k)} = A_k'^T e$ is the vector of column sums, then

$$A_{k+1} = A'_k T_k, \qquad T_k = \operatorname{diag}(\frac{q_j}{c_j^{(k)}})$$

Definition 2.20. We say that A is row normalized by p if $Ae = p$. We say that A is column normalized if by q is $A^T e = q$.

Theorem 2.21. *If $B \in \mathbb{P}(E_A)$ is row normalized, then $f(B) = S_0 B T_0$ is a contraction.*

Proof. Let $B, B' \in \mathbb{P}(E_A)$ be row normalized.
We want to show that

$$\mu(f(B), f(B')) \leq k(\mu(B, B')) \quad k < 1$$

There exist diagonal matrices $X = \operatorname{diag}(x), Y = \operatorname{diag}(y)$ such that

$$B' = XBY$$

and $\mu(B, B') = d(x, e) + d(y, e)$. Now $f(B') = S'_0 XBY T'_0$ and $B = S_0^{-1} f(B) T_0^{-1}$ implies

$$f(B') = S'_0 X S_0^{-1} f(B) T_0^{-1} Y T'_0$$

So

$$\mu(f(B), f(B')) = d(\frac{s'_0}{s_0} x, e) + d(\frac{t'_0}{t_0} y, e)$$

where $S_0 = \operatorname{diag}(s_0)$; $S'_0 = \operatorname{diag}(s'_0)$; $T_0 = \operatorname{diag}(t_0)$ and $T'_0 = \operatorname{diag}(t'_0)$.
Now,

$$B' \frac{e}{y} = \operatorname{diag}(x) B \operatorname{diag}(y) \frac{e}{y} = \operatorname{diag}(x) B e = Bex$$

Hence,

$$\begin{aligned} d(\frac{s'_0}{s_0} x, e) = d(\frac{r'_0}{r_0} x, e) &= d(Bex, B'e) \\ &= d(B' \frac{e}{y}, B'e) \\ &\leq \kappa(B') d(\frac{e}{y}, e) = \kappa(B') d(y, e) \end{aligned}$$

Now we want to show $d(\frac{t_0'}{t_0}y, e) \leq \kappa(B')d(x, e)$.
We have

$$B'^T \frac{e}{x} = \operatorname{diag}(y)B^T \operatorname{diag}(x)\frac{e}{x} = \operatorname{diag}(y)B^T e = B^T ey$$

Therefore,

$$d(\frac{t_0'}{t_0}y, e) = d((S_0 B)^T ey, (S_0' B')^T e) =$$

$$d(B^T S_0 ey, B'^T S_0' e)$$

$$= d(B^T ey, B'^T e)$$

$$= d(B'^T \frac{e}{x}, B'^T e)$$

$$\leq \kappa(B')d(x, e)$$

since $S_0 e = S_0' e = e$ and $\kappa(B'^T) = \kappa(B')$. □

Remark 2.22. It's not known if f is a contraction on $\mathbb{P}(E_A)$.

2.3.1 Convergence

We will use Hilbert projective metric and μ defined in the previous sections to show the convergence.

Theorem 2.23. *Let $A = A_0$ be column normalized. Then*

$$d(r^{(1)}, p) \leq \gamma d(r^{(0)}, p)$$

$$d(c^{(1)}, q) \leq \gamma d(c^{(0)}, q)$$

where $\gamma = \kappa(A)^2$, $r^{(k)} = A_k e$, $c^{(k)} = A_k'^T e$ and A_k, A_k' are the matrices defined by Sinkhorn's algorithm.

Proof. Let $A' = A_0'$. Since $r^{(1)} = A_1 e = A' T_0 e = A' \frac{q}{c^{(0)}}$ and $A'e = p$ we have

$$d(r^{(1)}, p) = d(A' \frac{q}{c^{(0)}}, A'e)$$

$$\leq \kappa(A) d(\frac{q}{c^{(0)}}, e)$$

$$= \kappa(A) d(q, c^{(0)})$$

since $\kappa(A)$ does not change during the iteration.
Now

$$c^{(0)} = A'^T e = A^T S_0 e = A^T \frac{p}{r^{(0)}}$$

and $A^T e = q$ imply

$$d(q, c^{(0)}) = d(A^T e, A^T \frac{p}{r^{(0)}}) \leq \kappa(A^T) d(r^{(0)}, p).$$

Since $\kappa(A) = \kappa(A^T)$ we obtain $d(r^{(1)}, p) \leq \gamma d(r^{(0)}, p)$.
The estimate for the column sums follows similarly. □

Remark 2.24. By construction, all matrices A_k in Sinkhorn's iteration are column normalized. Hence repeated applications of the previous theorem yield

$$d(r^{(k)}, p) \leq \gamma^k d(r^{(0)}, p)$$
$$d(c^{(k)}, q) \leq \gamma^k d(c^{(0)}, q)$$

Proposition 2.25. *The sequence $\{A_k\}$ generated by Sinkhorn's algorithm converges in $(\mathbb{P}(E_A), \mu)$ to the unique matrix $\widehat{B}$ such that $\widehat{B}e = p$, $\widehat{B}^T e = q$ and $\widehat{B} \sim A_0$.*

Proof. We first show that A_k is a Cauchy sequence.
Following the previous theorem,

$$\mu(A_k, A_{k+1}) = d(\frac{p}{r^{(k)}}, e) + d(\frac{q}{c^{(k)}}, e) =$$

$$= d(r^{(k)}, p) + d(c^{(k)}, q)$$
$$\leq \gamma^k \{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

.

Suppose $m, n > N$ and $m > n$, by triangle inequality

$$\mu(A_n, A_m) \leq (\gamma^n + \cdots + \gamma^{m-1})\{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

$$= \gamma^n (1 + \cdots + \gamma^{m-n-1})\{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

$$= \gamma^n \frac{1 - \gamma^{m-n}}{1 - \gamma}\{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

$$\leq \frac{\gamma^n}{1 - \gamma}\{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

Since $\gamma = \kappa(A)^2 < 1$, the last term tends to zero as n approaches ∞.
So there exists the limit $C \in \mathbb{P}(E_A)$ of the sequence A_k; we want to show that $C = \widehat{B}$.
For k large enough, $\forall \varepsilon > 0$

$$\mu(A_k, C) = d(x, e) + d(y, e) < \varepsilon,$$

where $C = \operatorname{diag}(x) A_k \operatorname{diag}(y)$. So C has column sums q.
Recall $A'_k = S_k A_k$ where $S_k = \operatorname{diag}(\frac{p}{r^{(k)}})$.
So

$$\mu(A_k, A'_k) = d(e, e) + d(t^{(k)}, e) = d(q, c^{(k)}) \leq \gamma^k d(c^{(0)}, q)$$

Now, by the triangle inequality

$$\mu(A'_k, C) \leq \mu(A'_k, A_k) + \mu(A_k, C)$$

and the last term tends to zero as k approaches ∞. Then, C has row sums p. So C is diagonally equivalent to $A_0 = A$ and has row sums p and column sums q.By unicity proved in Sinkhorn's theorem, $\widehat{B} = C$. □

Corollary 2.26 (Error bounds)**.** *Using the notation of Proposition 2.25*

$$\mu(A_k, \widehat{B}) \leq \frac{\gamma^k}{1-\gamma} \{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

Proof. If $m > k$ we already proved that

$$\mu(A_k, A_m) \leq \frac{\gamma^k}{1-\gamma} \{d(r^{(0)}, p) + d(c^{(0)}, q)\}.$$

If k is fixed and $m \to \infty$ then $A_m \to \widehat{B}$ and

$$\mu(A_k, \widehat{B}) \leq \frac{\gamma^k}{1-\gamma} \{d(r^{(0)}, p) + d(c^{(0)}, q)\}$$

□

Proposition 2.27 (Error bounds)**.** *If $A \sim B$, $A^T e = B^T e$ and $\mu(A, B) \leq \varepsilon$, then*

$$\exp(-\varepsilon) \leq \frac{b_{ij}}{a_{ij}} \leq \exp(\varepsilon)$$

for all i, j.

Proof. There exist diagonal matrices with positive entries $X = \text{diag}(x), Y = \text{diag}(y)$ such that $B = XAY$ and $d(x, e) \leq \varepsilon$ and $d(y, e) \leq \varepsilon$.
If $x' = \dfrac{1}{x_k}x$ where $x_k = \min\{x_i\}$ then

$$1 \leq x'_i \leq \exp(\varepsilon) \quad i = 1, .., n.$$

Without loss of generality we may assume $x = x'$. From $YA^T = B^T X^{-1}$ and $A^T e = q = B^T e$ it follows that

$$Yq = YA^T e = B^T X^{-1} e.$$

So

$$exp(-\varepsilon) \leq X^{-1} e \leq e$$

and multiplying by B^T,

$$\exp(-\varepsilon) q \leq B^T X^{-1} e \leq q$$

Now

$$\exp(-\varepsilon) \leq y_j \leq 1 \quad j = 1, \ldots, n.$$

Finally

$$\exp(-\varepsilon) \leq x_i y_j = \frac{b_{ij}}{a_{ij}} \leq \exp(\varepsilon)$$

□

Chapter 3

Optimal transport problem

In this chapter we will define an optimal transport problem (3.1) between two probability vectors. It is a known fact that the solution of this problem lies on a vertex of the polyhedral set $U(r, c)$ defined in 1.1; our aim is to approximate the solution.
Following [2], we will introduce a convex subset $U_\alpha(r, c)$ of $U(r, c)$, using a constraint on the Kullback-Leibler divergence. Hence, it's possible to define the optimal transport problem on the set U_α. For α large enough, $U_\alpha(r, c) = U(r, c)$ and the two formulations are equivalent.
Lagrange multipliers' theory allows to turn the problem defined on $U_\alpha(r, c)$ in the form of the definition 3.6. Proposition 3.10 shows that this formulation is equivalent to minimize a Kullback-Leibler divergence on $U(r, c)$.
Therefore, the existence and uniqueness of the solution to the problem defined in 3.6 is ensured by theorem 1.11. According to the condition on the partial derivatives of theorem 1.11, the solution of 3.6 is diagonally equivalent to a given matrix.
Moreover, by Sinkhorn's theorem, there exists a unique matrix of that form, with row sums r and column sums c.
Therefore, it is possible to compute the solution of the problem defined in 3.6 using Sinkhorn's algorithm. In the last section we will show some numerical experiments.

3.1 Entropic constraint

Definition 3.1 (OT problem)**.** Let $U(r, c)$ be the set defined in 1.1. Given a $d \times d$ cost matrix M, the problem

$$d_M(r, c) := \min_{P \in U(r,c)} < P, M >$$

is called an optimal transport problem between r and c with given cost M.

Proposition 3.2. *The set* $U_\alpha(r,c) := \{P \in U(r,c) | KL(P\|rc^T) \leq \alpha\} = \{P \in U(r,c) | h(P) \geq h(r) + h(c) - \alpha\} \subset U(r,c)$ *is convex.*

Proof. The two definitions are equivalent because $KL(P\|rc^T) = h(r)+h(c)-h(P)$. If P, Q lie in $U_\alpha(r,c)$ then $\forall t \in [0,1]$ $h((1-t)P+tQ) \geq (1-t)h(P)+th(Q) \geq (1-t)(h(r)+h(c)-\alpha)+t(h(r)+h(c)-\alpha) = h(r)+h(c)-\alpha$. So $(1-t)P+tQ \in U_\alpha(r,c)$. We used the concavity of the entropy (Remark 1.4) in the first inequality. □

Remark 3.3. rc^T is the joint density in the case r and c are independent.

Definition 3.4 (OT with entropic constraint)**.**

$$d_{M,\alpha}(r,c) := \min_{P \in U_\alpha(r,c)} < P, M >$$

Remark 3.5. Since for any $P \in U(r,c)$ $h(P)$ is lower bounded by $\frac{1}{2}(h(r) + h(c))$, we have that for α large enough $U_\alpha(r,c) = U(r,c)$ and so $d_{M,\alpha}(r,c) = d_M(r,c)$.

Definition 3.6 (dual-Sinkhorn divergence)**.** For $\lambda > 0$,

$$d_M^\lambda(r,c) := \min_{P \in U(r,c)} < P, M > -\frac{1}{\lambda}h(P)$$

Remark 3.7. Equivalence between 3.6 and 3.4 is a consequence of the following proposition.

Proposition 3.8. *Let* $G(P) =< P, M >$*, let* $F : \mathbb{R}^n \to \mathbb{R}$ *be a smooth and strictly convex function. If there exists a unique minimum* $P^\star$ *for* G *such that* $F(P^\star) = \alpha$ *where* $\alpha > 0$*, then there exists a* $\mu < 0$ *such that* $(P^\star, \mu)$ *is a minimum for*

$$\Lambda(P,\mu) = G(P) + \mu F(P)$$

Proof. Using Lagrange multipliers, there exists a $\mu^\star$ such that $(P^\star, \mu^\star)$ is a stationary point for

$$\Lambda(P,\mu) = G(P) + \mu F(P)$$

It's known that ∇G must be proportional to ∇F in the point $P^\star$. So

$$\nabla G(P^\star) = \mu^\star \nabla F(P^\star)$$

Suppose $\mu^\star \geq 0$, so there exists a T minimizing G with $F(T) < \alpha$, this is absurd because by hypothesis the minimum must achieve $F(T) = \alpha$.
Using the convexity of F, $(P^\star, \mu^\star)$ is a minimum for Λ. □

Proposition 3.9. *For every $\alpha \in [0, \infty]$ there exists a $\lambda > 0$ such that $d_{M,\alpha}(r, c) = d_M^\lambda(r, c)$*

Proof. Fix $\alpha > 0$.
If P^α is the matrix such that $< P^\alpha, M >= d_{M,\alpha}(r, c)$, so $P^\alpha \in \partial U_\alpha(r, c)$.
Hence, P^α satisfies $KL(P^\alpha \| rc^T) = \alpha$ which is equivalent to

$$h(P^\alpha) + \alpha - h(r) - h(c) = 0$$

By the previous theorem, there exists a $\mu < 0$ such that (P^α, μ) is a minimum for

$$\Lambda(P, \mu) =< P, M > +\mu KL(P^\alpha \| rc^T)$$

So (P^α, μ) is a minimum for

$$\Lambda'(P, \mu) =< P, M > +\mu h(P)$$

Setting $\mu = -\frac{1}{\lambda}$ where $\lambda > 0$ we get the proposition.

□

Proposition 3.10. $d_M^\lambda(r, c) = \min_{P \in U(r,c)} \frac{KL(P \| \exp(-\lambda M))}{\lambda}$

Proof.

$$\begin{aligned} KL(P \| \exp(-\lambda M)) &= \sum_{i,j} p_{ij} \log \frac{p_{ij}}{\exp(-\lambda m_{ij})} \\ &= \sum_{i,j} p_{ij} \log p_{ij} + \lambda \sum_{i,j} p_{ij} m_{ij} \\ &= -h(P) + \lambda < P, M > \end{aligned}$$

□

Theorem 3.11. *For $\lambda > 0$ the solution $P^\lambda = (p_{ij}^\lambda)$ of the problem defined in 3.6 is unique and has the form $P^\lambda = \operatorname{diag}(u) K \operatorname{diag}(v)$ where $u, v \in \mathbb{R}_+^d$ are uniquely defined up to a multiplicative factor and $K = e^{-\lambda M}$ is the element-wise exponential of $-\lambda M$.*

Proof. The existence and uniqueness of the solution P^λ is ensured by Proposition 3.10 and Theorem 1.11.
By the condition on partial derivatives in theorem 1.11, there exist $\alpha, \beta \in \mathbb{R}_+^d$ such that

$$\frac{\delta}{\delta p_{ij}} \left[KL(P \| e^{-\lambda M}) - \alpha^T P e - \beta^T P^T e \right]_{P = P^\lambda} = 0$$

that is equivalent to

$$\log p_{ij}^{\lambda} + 1 + \lambda m_{ij} - \alpha_j - \beta_i = 0$$

which leads to

$$p_{ij}^{\lambda} = e^{-1/2+\alpha_j} e^{-\lambda m_{ij}} e^{\beta_i - 1/2}.$$

Therefore, P^{λ} is diagonally equivalent to K and has row sums r and column sums c.
Since K has strictly positive entries, Sinkhorn's theorem states that there exists a unique matrix diagonally equivalent to K that belongs to $U(r, c)$. P^{λ} is necessarily that matrix. □

3.2 Numerical experiments

3.2.1 1-D marginals

It is possible to implement Sinkhorn's algorithm in Matlab to find the solution P^{λ} of 3.6. **Algorithm 1** computes Hilbert's projective metric between two given vectors in $\mathbb{R}_+^n$.

Algorithm 1.

```
function d= distance (w,v)
s=v(1)/w(1);
t=v(1)/w(1);
for i=2:length(v)
    if (v(i)/w(i)>s)
        s=v(i)/w(i);
    end
end
for i=2:length(v)
    if(v(i)/w(i)<t)
        t=v(i)/w(i);
    end
end
d=log (s/t);
end
```

Given a matrix cost M and marginals r and c, **Algorithm 2** computes the unique matrix P which has row sums r and column sums c. We will use Hilbert's projective metric computed in **Algorithm 1** as a stopping criterion. We say that Sinkhorn's algorithm converges to P at $k-th$ iteration if the sum of the distance between the row sums and r and column sums and c is less then 10^{-5}.

Algorithm 2.

```
function [P,k,d]=sink (M,r,c,l)
[n,m]=size (M);
e=ones(n,1);
e1=ones(m,1);
P=M;
d=1;
k=1;
while(d>10^-5 && k<l)
      j=P*e1;
      A1=diag(r./j)*P;
      g=A1'*e;
      P=A1*diag(c./g);
      d=distance(j,r)+distance(g,c);
      k=k+1;
end
end
```

Given variance σ^2, mean m and a uniform grid of an interval a, **Algorithm 3** computes a discretized Gaussian distribution on a.

Algorithm 3.

```
function q=gaussd(a,sigma,m)
q=ones(length(a),1);
for j=1:length(a)
q(j,1)=exp(-(a(j)-m)^2/(2*sigma^2));
end
q=(1/sum(q))*q;
end
```

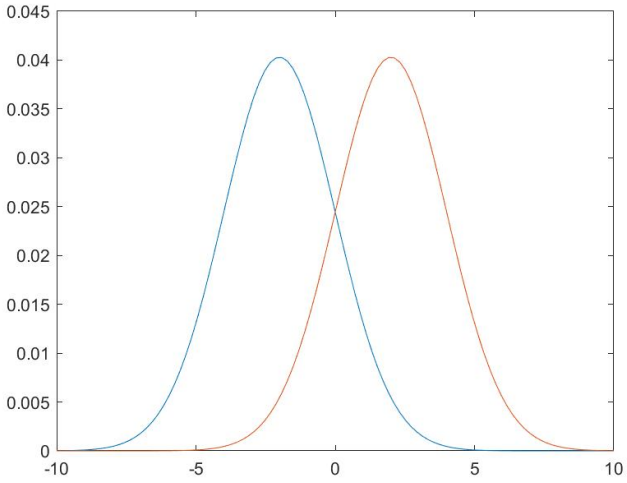
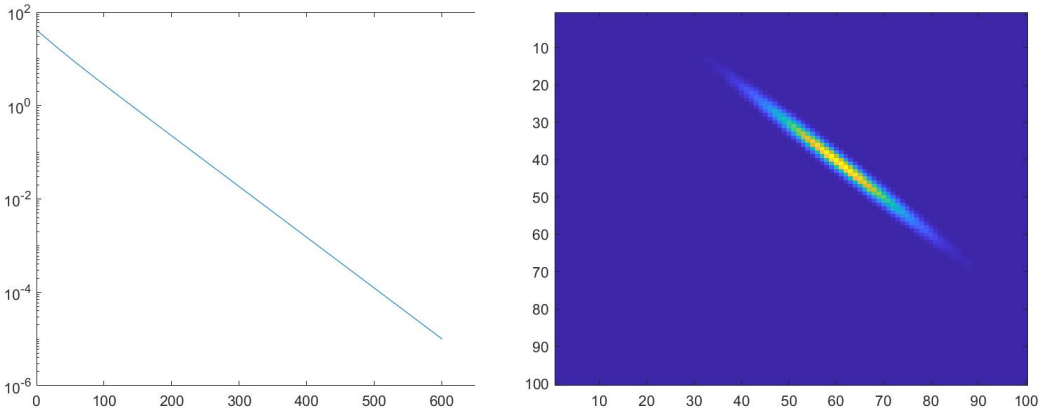


Figure 3.1: Marginals p (blue) and q (red)

Marginals p and q in Figure 3.1 are Gaussian distributions with variance 4, and mean respectively $m_1 = -2$ and $m_2 = 2$; discretized on a uniform grid $(x_i)_{i=1}^{100}$ of 100 points of $[-10, 10]$. Using the cost matrix $M_{i,j} = |x_i - x_j|^2$ and $K = e^{-\lambda M}$ elementwise with $\lambda = 5$, Figure 3.2 (a) shows the convergence of Sinkhorn's algorithm from p to q and Figure 3.2 (b) shows the structure of the transport matrix, that is similiar to a translation.



(a) k=number of iterations d=distance(j,p)+distance(k,q)

(b) Transport matrix P obtained with Sinkhorn's algorithm

Figure 3.2: Convergence of Sinkhorn's algorithm for a fixed λ

Figure 3.3 and Figure 3.4 show the convergence of the algorithm for different values of λ.

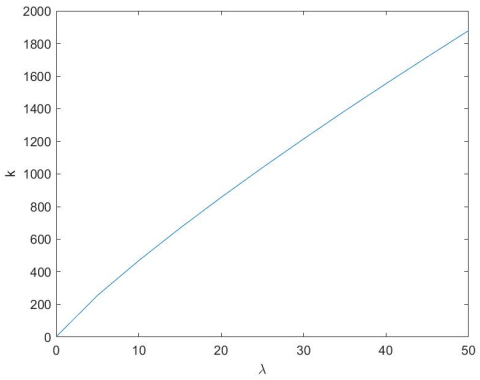


Figure 3.3: k=number of iterations of Sinkhorn's algorithm; λ is the parameter used computing $K = e^{-\lambda M}$

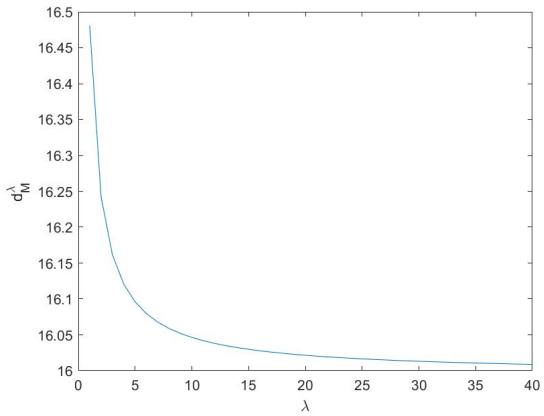


Figure 3.4: For encreasing values of λ, d_M^λ approaches $|m_1 - m_2|^2$

Let $S_k A T_k$ be the k-th iteration of Sinkhorn's algorithm starting from the matrix A. Let p be the normalized row sum and q be the normalized column sum. If $\{q_k\}$ is the sequence of column sums of $S_{k+1}S_k A T_k$, then by Sinkhorn's theorem $S_{k+1}S_k A T_k$ is the unique matrix diagonally equivalent to A with row sum p and column sum q_k. The sequence $\{q_k\}$ converges to q. Therefore, stopping the algorithm at the i-th iteration and normalizing the rows we get a transport matrix between p and q_k. This gives an idea of the convergence from p to q. The computation is performed with cost matrix $M_{i,j} = |x_i - x_j|^2$ and $K = e^{-5M}$ elementwise.

Algorithm 4.

```
function Y=succ(K,p,q,n)
P=sink(K,p,q,n);
[a,b]=size(P);
e=ones(b,1);
r=P*e;
Y=diag(p./r)*P;
end
```

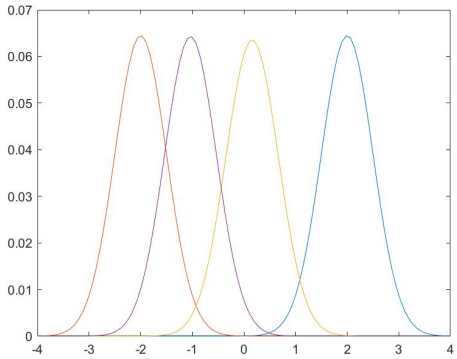


Figure 3.5: Input marginals p (blue) and q (red) are gaussian distributions with the same variance. In yellow the distribution of Ye for $n = 7$. In violet the distribution of Ye for $n = 15$.

Figure 3.5 shows the distribution of Ye where Y is computed using **Algorithm 4** for different values of n. We get similiar results if the input marginals p and q have different variances.

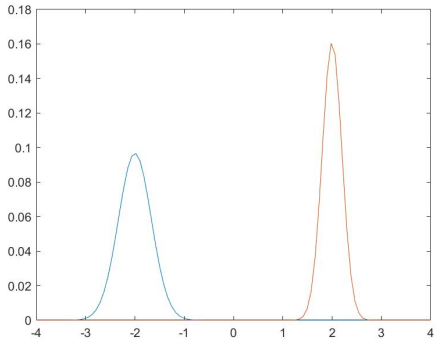


Figure 3.6: Marginals p (blue) and q (red) with different variances

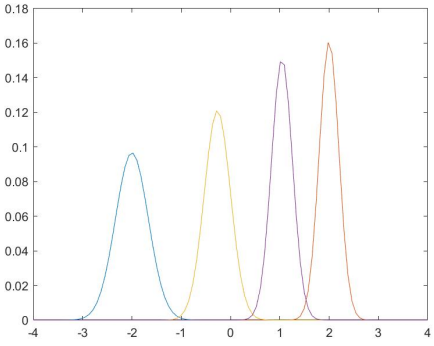


Figure 3.7: In yellow the distribution of Ye for $n = 2$, in violet the distribution of Ye for $n = 4$.

Fix $0 \leq t \leq 1$. Given the transport matrix P obtained with Sinkhorn's algorithm, we can know where the point $(1-t)i+tj$ is sent by the matrix P. If $t = 1$ we get the marginal $q = P'e$, if $t = 0$ we get the marginal $p = Pe$.

Algorithm 5.

```
function r=inter(P,t)
[n,m]=size(P);
r=zeros(n,1);
for i=1:size(P)
    for j=1:size(P)
        k=floor((1-t)*i+t*j);
        r(k,1)=r(k,1)+P(i,j);
    end
end
end
```

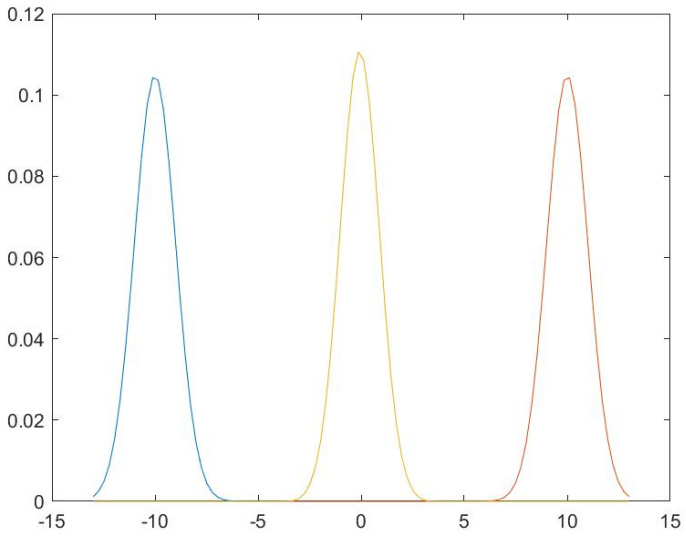
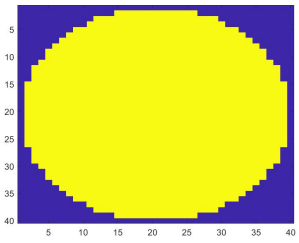


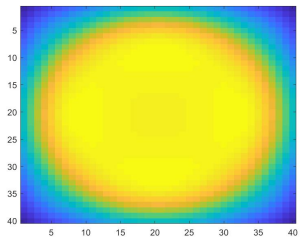
Figure 3.8: Input marignals p (blue) and q (red), in yellow interpolation with $t = 1/2$

3.2.2 2-D marginals

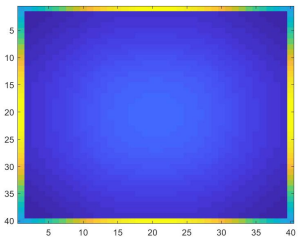
Let $(x_i)_{i=1}^{1600}$ be a uniform grid of 40×40 points in $[-1,1]^2$. If the input marginal s is a uniform density distribution on the discretized ball centered in $(0,0)$ of $[-1,1]^2$ and r is the uniform density on the discretized boundary of the square, Figure 3.9 shows the convergence of Sinkhorn's algorithm with input bidimensional marginals s and r. The computation is performed with cost matrix $M_{ij} = \|x_i - x_j\|^2$ and $K = e^{-10M}$ elementwise. Subfigures from (b) to (e) show the distribution of Ye where Y is computed using **Algorithm 4** for different values of n.



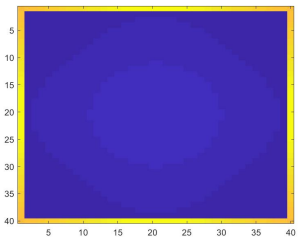
(a) Uniform density on the discretized ball



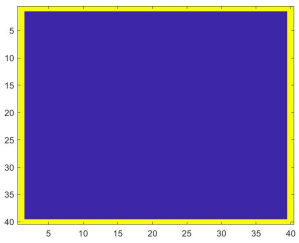
(b) $n = 1$



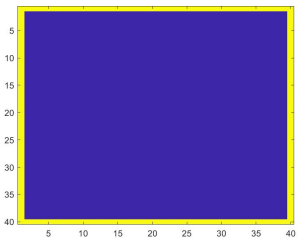
(c) $n = 2$



(d) $n = 5$



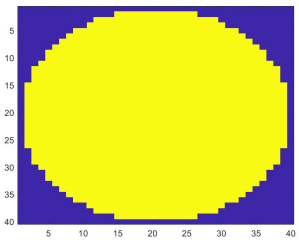
(e) $n = 10$



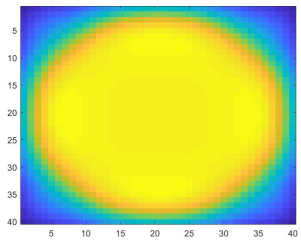
(f) Uniform density on the discretized boundary of the square

Figure 3.9: Sinkhorn's algorithm with input marginals (a) and (f)

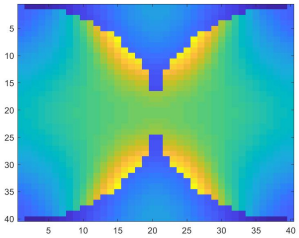
Let $(x_i)_{i=1}^{1600}$ be a grid of 40×40 points in $[-1,1]^2$. Let $M_{ij} = \|x_i - x_j\|^2$ and $K = e^{-10M}$ elementwise. Let s be the uniform density on a ball centered in $(0,0)$. Let r be the uniform density on two balls centered in (-1,0) and in (1,0). The following figure shows the convergence of Sinkhorn's algorithm from s to r. Subfigures from (b) to (e) show the distribution of Ye where Y is computed using **Algorithm 4** with different values of n.



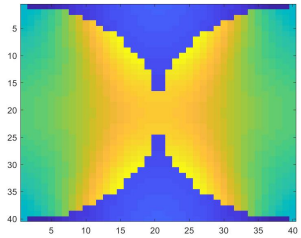
(a) Uniform density on the discretized ball



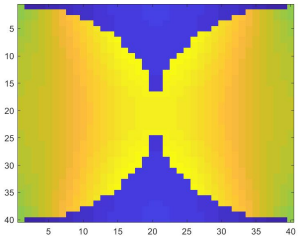
(b) $n = 1$



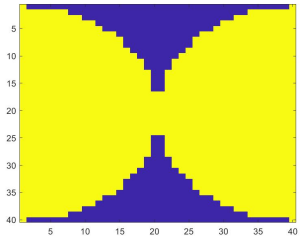
(c) $n = 2$



(d) $n = 4$



(e) $n = 6$



(f) Uniform density on two balls centered in (-1,0) and (1,0)

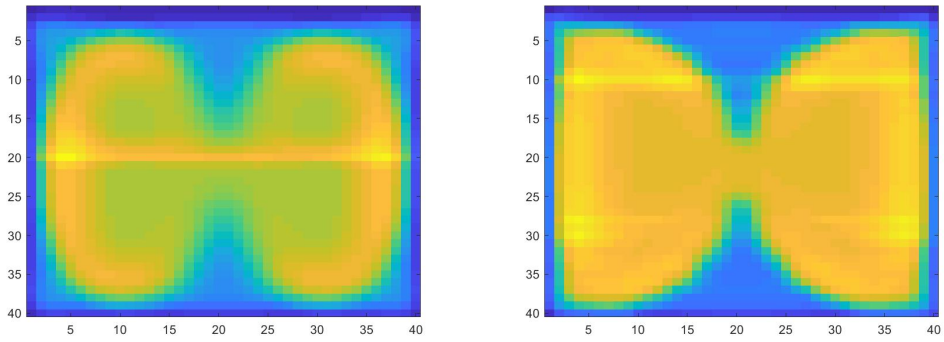
Figure 3.10: Sinkhorn's algorithm with input marginals (a) and (f)

The following algorithm is a variation of **Algorithm 3** for the 2-D case.

Algorithm 6.

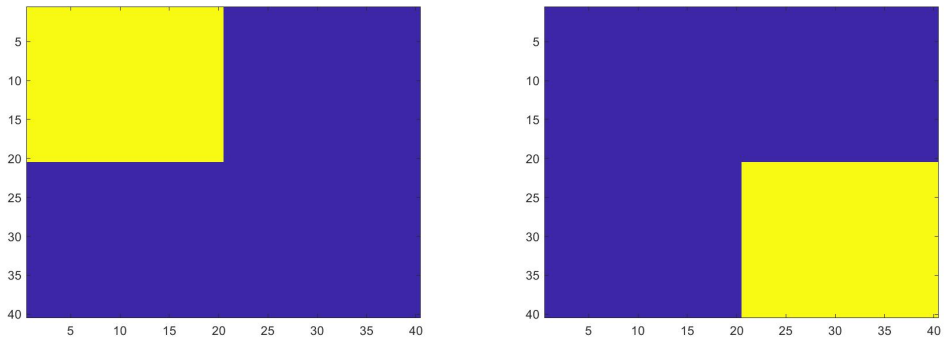
```
function r=inter2 (P,t,n)
r=zeros(n,n);
for i=1:(n^2-1)
    x1=floor(i/n);
    y1=i-n*x1;
    for j=1:(n^2-1)
        x2=floor(j/n);
        y2=j-n*x2;
        w=floor((1-t)*x1+t*x2);
        k=floor((1-t)*y1+t*y2);
        walt=ceil((1-t)*x1+t*x2);
        kalt=ceil((1-t)*y1+t*y2);
        r(k+1,w+1)=r(k+1,w+1)+P(i,j)/2;
        r(kalt+1,walt+1)=r(kalt+1,walt+1)+P(i,j)/2;
    end
end
end
```

Let P be the transport matrix between (a) and (f) in Figure 3.10. Figure 3.11 shows the results of **Algorithm 6** for two values of t and $n = 40$.



(a) $t = \frac{1}{2}$ **(b)** $t = \frac{3}{4}$

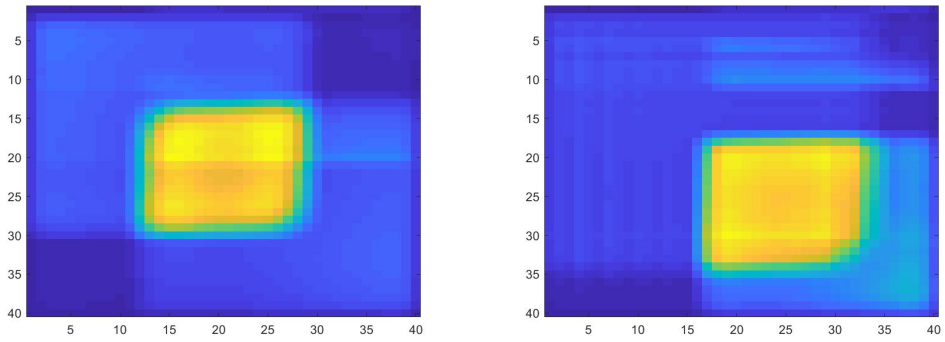
Figure 3.11: Interpolation between uniform density on a ball centered in $(0, 0)$ and uniform density on two balls centered in $(-1, 0)$ and $(1, 0)$.



(a) Marginal s **(b)** Marginal r

Figure 3.12: Marginals s and r are uniform density on a rectangle in the discretized square $[-1, 1]^2$

If P is the matrix obtained with Sinkhorn's algorithm with marginals input s and r in Figure 3.12 , Figure 3.13 shows the results of **Algorithm 6** for two values of t.



(a) $t = \frac{1}{2}$ **(b)** $t = \frac{3}{4}$

Figure 3.13: Interpolation between marginals s and r in Figure 3.12 for two values of t.

Bibliography

[1] P. Bushell. Hilbert's metric and positive contraction mappings in banach space. *IEEE Trans. Inf. Theor.*, 37(6):330–338, September 1973.

[2] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. pages 2292–2300, 2013.

[3] A. Dembo, T. M. Cover, and J. A. Thomas. Information theoretic inequalities. *IEEE Trans. Inf. Theor.*, 37(6):1501–1518, September 2006.

[4] Simon P. Eveson and Roger D. Nussbaum. An elementary proof of the Birkhoff-Hopf theorem. *Math. Proc. Cambridge Philos. Soc.*, 117(1):31–55, 1995.

[5] J. Franklin and J. Lorenz. On the scaling of multi-dimensional matrices. *Linear Algebra and its Applications*, 2(37):717–735, 1989.

Acknowledgements

The author would like to...scherzavo, almeno questi li scrivo in italiano.
Innanzitutto ringrazio il prof. Dario Trevisan che in questi mesi mi ha trasmesso tanto con estrema chiarezza e precisione.
Alla mia famiglia, a cui devo tutto, a chi c'è e a chi mi starà guardando con orgoglio da qualche parte. A mio padre e mia madre, che mi insegnano da sempre a credere nei miei mezzi e hanno piena fiducia in me. A mia sorella, che è stata un esempio da seguire in questi anni. Ai miei zii speciali, che mi hanno insegnato a dare tutto per quello in cui credo e che le difficoltà vanno affrontate con lucidità e freddezza.
Grazie ai miei "frat fidat", abbiamo capito dal primo giorno di scuola che insieme siamo una forza. Grazie ai minuti della TIM con cui chiamo Saverio, che ha sopportato le mie continue lamentele in questi anni. Più passa il tempo e più sono sicuro di poter contare su di te. A Peppe, alla sua risata contagiosa, ai suoi scherzi infami e al suo saper ascoltare quando serve. Orgoglioso di questi anni fianco a fianco.
Alla mia famiglia pisana, a Mattia Gavini con cui ho condiviso tantissimo, a partire da serate a Vettovaglie fino addirittura a una mail (Cari Vittorio e Mattia...) con i risultati di un esame, passata alla storia. Grasie a Mattia Cimorelli e alla tassa quotidiana pagata sui caffè; grazie per tutte le avventure passate negli anni, tra coinquilini che forse ancora ti stanno maledicendo ed esami spollati. Grazie al mio amico filosofo (?) Martino, alla calma olimpica con cui affronta ogni situazione e alle mille serate passate insieme. Grazie per esserci sempre. Grazie a Michele che è una spalla su cui poter fare sempre affidamento. Grazie anche per aver risposto a tutte le mie domande (soprattutto quelle di G2). Ora pensiamo agli obiettivi seri e andiamo a vincere questo fantacalcio. Grazie ad Annalisa che in questi anni ha sopportato le mie continue prese in giro; grazie a Veronica e alle indimenticabili serate passate tutti insieme a cantare a Cavalieri. Grazie a te che mi hai dato tantissimo anche quando potevi evitare e mi hai fatto capire che può fare freddo anche in un giorno di Agosto. Siamo stati capaci di creare un rapporto unico. Grazie per questi anni, Chiara. Grazie a Bargagnati e

alla sua inconfondibile voce tenue. Ad Alfonso, Alessandro, Max, Sciabolata e tutti gli altri della squadra di cui sono orgogliosamente il capitano da tre anni(qualcuno dice anche presidente). Siamo scarsi ma comunque belli veri speciali. Al club di scommettitori anonimi dell'aula 4, all'aula studenti e a tutte le persone con cui ho condiviso qualcosa in questi anni. Ai posti in cui ho passato più tempo che a casa mia: all'aula portatili, al barrique, al sud, a carta gialla, al solito, al cus.
Non ringrazio infine chi non mantiene la parola data: Ceccorivolta.