RAIRO Operations Research
RAIRO Oper. Res. **40** (2006) 1–17
DOI: 10.1051/ro:2006011

NUMERICAL SOLUTIONS OF THE MASS TRANSFER PROBLEM *

Serge Dubuc[1] and Issa Kagabo[2]

Abstract. Let μ and ν be two probability measures on the real line and let c be a lower semicontinuous function on the plane. The mass transfer problem consists in determining a measure ξ whose marginals coincide with μ and ν, and whose total cost $\iint c(x,y)\,\mathrm{d}\xi(x,y)$ is minimum. In this paper we present three algorithms to solve numerically this Monge-Kantorovitch problem when the commodity being shipped is one-dimensional and not necessarily confined to a *bounded interval*. We illustrate these numerical methods and determine the convergence rate.

Keywords. Continuous programming, transportation, mass transfer, optimization.

1. Introduction

The mass transfer problem, also known as the Monge-Kantorovitch problem, involves leveling a piece of land. It is natural to remove soil from areas whose level is above the average, and put it in the hollows whose level is below it. To minimize the work done, we have to find a model that minimizes the total displacement of earth.

The Monge-Kantorovitch problem was first studied by the French mathematician Monge [14] in 1781. His work has been extended by several mathematicians,

Received December 10, 2003. Accepted November 7, 2005.

* *The first author is supported by a research grant of NSERC (Federal Government of Canada).*

[1] Département de mathématiques et de statistique, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada; dubucs@dms.umontreal.ca

[2] Department of Mathematics, Columbia Union College, 7600 Flower Avenue, Takoma Park, MD, 20912, USA; ikagabo@cuc.edu

among them Appell [4, 5] and Hitchcock [9], who formulated its discrete version, and Kantorovitch [11], who formulated it as a mathematical program in a function space. This latter approach has been adopted in the modern literature.

Let μ and ν be regular measures defined on the topological spaces X and Y respectively with $\mu(X) = \nu(Y)$. Let c denote a lower semicontinuous function on the product space $X \times Y$. The mass transfer problem consists in determining a measure ξ on $X \times Y$ with marginals μ and ν and such that the total transportation cost (primal objective)

$$\iint_{X\times Y} c(x, y)\,\mathrm{d}\xi(x, y)$$

be minimal.

In this paper we present three algorithms to solve the Monge-Kantorovitch problem numerically when the commodity to be shipped is one-dimensional. We obtain solutions in the extended case where the material is not necessarily located on a bounded interval. We then provide the theoretical solution of four particular problems. Later we illustrate these numerical methods in three examples. In each example, we know the exact solution for the Monge-Kantorovitch problem and the numerical approximation, which allows us to obtain the convergence rate.

2. Preliminary results

The mass transfer problem, or primal problem, is defined as

$$\gamma = \inf_{\xi\geq 0} \iint_{X\times Y} c(x, y)\,\mathrm{d}\xi(x, y) \quad \text{subject to} \quad \begin{array}{l} P_X\xi = \mu \\ P_Y\xi = \nu. \end{array} \tag{1}$$

$P_X\xi$ and $P_Y\xi$ are the projections of ξ onto X and Y. We say that a measure ξ on the product space $X \times Y$ is feasible if, for any choice of compact subsets $K \subseteq X$ and $L \subseteq Y$ we have that

$$\xi(K \times Y) = \mu(K) \;\text{ and }\; \xi(X \times L) = \nu(L).$$

This condition is satisfied if and only if, for each pair of continuous functions ϕ and ψ defined on X and Y respectively, we have:

$$\iint_{X\times Y} \phi(x)\,\mathrm{d}\xi(x, y) = \int_X \phi(x)\,\mathrm{d}\mu(x)$$

and

$$\iint_{X\times Y} \psi(y)\,\mathrm{d}\xi(x, y) = \int_Y \psi(y)\,\mathrm{d}\nu(y).$$

One says that ξ is a measure on the product space $X \times Y$ whose marginals are μ and ν respectively. Throughout the paper we denote $\Gamma(\mu, \nu)$ the space of measures on the product space $X \times Y$ with marginals μ and ν.

Kantorovitch [11] has shown that the set of optimal measures is nonempty. However he didn't construct a solution. Under some conditions on the cost function, some authors have determined an optimal measure explicitly. For instance if the cost function c is superadditive, the lower Fréchet bound is an optimal solution. Before stating this result, proved by Tchen [16], let us recall the definition of a superadditive function and introduce Fréchet bounds.

We say that a two-variable function c is superadditive if

$$c(x,y) + c(x',y') - c(x',y) - c(x,y') \geq 0 \tag{2}$$

whenever $x \leq x'$ and $y \leq y'$. The functions xy, $(x+y)^2$, $\min\{x,y\}$, $-\max\{x,y\}$, $-|x-y|^p$ $(p \geq 1)$, $f(x-y)$ (f concave) are examples of superadditive functions.

The lower and upper Fréchet bounds are the joint distributions respectively defined as

$$\xi_*(x,y) = \max\{\mu(x) + \nu(y) - 1, 0\}, \quad \xi^*(x,y) = \min\{\mu(x), \nu(y)\}. \tag{3}$$

Both ξ_* and ξ^* have marginals μ and ν.

Based on Fréchet bounds, Rachev and Rüschendorf [15] state the following result that characterizes the set of feasible measures of the primal problem (1).

Theorem 2.1. *The measure ξ is in $\Gamma(\mu,\nu)$ if and only if*

$$\xi_*(x,y) \leq \xi(x,y) \leq \xi^*(x,y) \quad \text{for all} \quad (x,y) \in X \times Y.$$

In an instance of problem (1) with $c(x,y) = \pm xy$, Hoeffding [10] and Fréchet [7] have proved that

$$\int_{\mathbb{R}^2} xy \, \mathrm{d}\xi_* \leq \int_{\mathbb{R}^2} xy \, \mathrm{d}\xi \leq \int_{\mathbb{R}^2} xy \, \mathrm{d}\xi^*.$$

Tchen [16] extended this result to superadditive functions costs. He proved the following key result.

Theorem 2.2. *If $c : \mathbb{R}^2 \to \mathbb{R}$ is continuous and superadditive then every probability measure ξ in $\Gamma(\mu,\nu)$ satisfies:*

$$\int_{X \times Y} c \, \mathrm{d}\xi_* \leq \int_{X \times Y} c \, \mathrm{d}\xi \leq \int_{X \times Y} c \, \mathrm{d}\xi^*. \tag{4}$$

Throughout the paper we say that a function c fulfills the condition (5) if two functions $p(x)$ and $q(y)$ can be found such that

$$|c(x,y)| \leq p(x) + q(y) \tag{5}$$

for every (x,y) in $X \times Y$, with $p \in L^1(\mu)$, $q \in L^1(\nu)$ and for every $x \in X$, $p(x) < \infty$, for every $y \in Y$, $q(y) < \infty$. Kellerer [12] was the first to state this condition. When this condition is fulfilled the primal problem (1) has a finite optimal solution.

We illustrate how to verify this condition in a given example. We consider the mass transfer problem (1) where $X = Y = \mathbb{R}$, $\mu = \nu$ is the probability measure whose distribution function is a normal probability distribution and $c(x, y) = \exp(|x| + |y|)$. For this mass transfer problem, condition (5) is satisfied, we may choose $p(x) = \exp(2|x|)$, $q(y) = \exp(|2|y|)$.

3. Algorithms for the continuous transportation problem

We suppose that μ and ν are two probability measures on the real line. Before stating our algorithms, let us recall that the mass transfer problem constitutes a continuous extension of the Hitchcock transportation problem also known as the classical transportation problem. The latter is one of the most studied problems in the field of network flows.

Let a_1, a_2, ..., a_m denote the supply at origins 1, 2, ..., m and b_1, b_2, ..., b_n the demand at destinations 1, 2, ..., n. Let c_{ij} represents the unit transportation cost from an origin i to a destination j. The classical transportation problem consists in determining the amount of the commodity x_{ij} to be shipped between an origin i and a destination j at a minimum transportation cost, while satisfying the demand at each destination.

Mathematically, it takes the form of the linear program:

$$\begin{aligned} \min \quad \rho &= \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{n} x_{ij} = a_i \qquad 1 \le i \le m \\ & \sum_{i=1}^{m} x_{ij} = b_j \qquad 1 \le j \le n \\ & x_{ij} \ge 0 \qquad 1 \le i \le m, \quad 1 \le j \le n, \end{aligned} \tag{6}$$

where a_i denotes the supply at origin i, b_j the demand at destination j, c_{ij} the unit transportation cost from origin i to destination j and x_{ij} the amount of commodity being shipped from origin i to destination j.

The analogy between the mass transfer problem and the classical transportation problem is obvious. In the continuous transportation problem, origins are the points where the earth soil has to be removed and destinations are the points where the earth has to be deposited.

We suggest algorithms to compute numerically

$$\gamma = \inf \iint_{\mathbb{R}^2} c(x, y)\, \mathrm{d}\xi(x, y) \tag{7}$$

where ξ is a measure on $\mathbb{R}^2$ with respective marginals μ and ν.

For the following algorithms we suppose that the cost function $c(x,y)$ fulfills the condition (5).

Algorithm 1. Let us consider $\epsilon > 0$ sufficiently small. One can find $M > 0$ sufficiently large that

$$\int_{|x|>M} p(x)\,\mathrm{d}\mu(x) < \epsilon/4.$$

We subdivide the interval $A = [-M, M]$ in m sub-intervals $A_1, A_2, ..., A_m$.

If we denote $A_0 = (-\infty, -M)$, $A_{m+1} = (M, +\infty)$,

$$A_0, A_1, ..., A_m, A_{m+1}$$

constitute a partition of $\mathbb{R}$. Similarly for the measure ν, one can find $N > 0$ sufficiently large that

$$\int_{|y|>N} q(y)\,\mathrm{d}\nu(y) < \epsilon/4.$$

We subdivide the interval $B = [-N, N]$ in n sub-intervals $B_1, B_2, ..., B_n$

If we denote $B_0 = (-\infty, -N)$, $B_{n+1} = (N, +\infty)$,

$$B_0, B_1, ..., B_n, B_{n+1}$$

constitute a partition of $\mathbb{R}$.

Let c_{ij} be an approximated value of $c(x,y)$ on the rectangle $A_i \times B_j$. We solve the following Hitchcock problem

$$\begin{aligned}
\min \quad \gamma_{1,mn} = \ & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
\text{s.t.} \ & \sum_{j=1}^{n} x_{ij} = \mu(A_i) \qquad 1 \le i \le m \\
& \sum_{i=1}^{m} x_{ij} = \nu(B_j) \qquad 1 \le j \le n \\
& x_{ij} \ge 0 \qquad 1 \le i \le m, \quad 1 \le j \le n
\end{aligned}$$

and we wish that for m and n sufficiently large, $\gamma_{1,mn}$ be a good approximation of the optimal value of the continuous transportation problem (7).

It is natural that on a sub-rectangle $A_i \times B_j$, we approximate $c(x,y)$ by $c_{ij} = c(x_i, y_j)$ where $(x_i, y_j) \in A_i \times B_j$.

Throughout this paper we denote:

$$M_{ij} = \sup\{c(x,y) : (x,y) \in A_i \times B_j\},$$

$$m_{ij} = \inf\{c(x,y) : (x,y) \in A_i \times B_j\}.$$

Algorithm 2. We choose M and N such that $\mu(A) = \nu(B)$. On the sub-rectangle $A_i \times B_j$ $1 \leq i \leq m$, $1 \leq j \leq n$, we use the following approximation of $c(x, y)$:

$$c_{ij} = \inf\{c(x, y) : (x, y) \in A_i \times B_j\} = m_{ij}.$$

As the integral of the cost function on $(A \times B)^c$ is negligible, we replace $c(x, y)$ by zero when (x, y) doesn't belong to $A \times B$. We do the same for the sub-rectangles $A_i \times B_j$ where $i \in \{0, m+1\},\ j \in \{0, n+1\}$.

Then we solve the following classical transportation problem:

$$\begin{aligned} \min \quad \gamma_1 = & \sum_{i=1}^{m} \sum_{j=1}^{n} m_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{n} x_{ij} = \mu(A_i) \qquad 1 \leq i \leq m \\ & \sum_{i=1}^{m} x_{ij} = \nu(B_j) \qquad 1 \leq j \leq n \\ & x_{ij} \geq 0 \qquad 1 \leq i \leq m, \quad 1 \leq j \leq n. \end{aligned} \tag{8}$$

Before bounding $\gamma - \gamma_1$, let us establish some results that we will need after considering the two following problems:

$$\gamma_2 = \inf \iint_{A \times B} c(x, y)\, \mathrm{d}\xi(x, y) \tag{9}$$

where ξ is a measure on $A \times B$ with respective marginals $\mu|_A$ and $\nu|_B$, and

$$\begin{aligned} \min \quad \gamma_3 = & \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{n} x_{ij} = \mu(A_i) \qquad 1 \leq i \leq m \\ & \sum_{i=1}^{m} x_{ij} = \nu(B_j) \qquad 1 \leq j \leq n \\ & x_{ij} \geq 0 \qquad 1 \leq i \leq m, \quad 1 \leq j \leq n. \end{aligned} \tag{10}$$

Lemma 3.1. *If $c(x, y) \geq 0$, then $\gamma_1 \leq \gamma_2 \leq \gamma_3$.*

Proof. We first show the inequality $\gamma_1 \le \gamma_2$. Let ξ_2 be an optimal solution of the problem (9). We have

$$\begin{aligned}
\gamma_2 &= \iint_{A\times B} c(x,y)\,\mathrm{d}\xi_2(x,y)\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\iint_{A_i\times B_j} c(x,y)\mathrm{d}\xi_2(x,y)\\
&\ge \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}\iint_{A_i\times B_j}\mathrm{d}\xi_2(x,y)\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}\xi_2(A_i\times B_j).
\end{aligned}$$

Let us set $y_{ij} = \xi_2(A_i \times B_j)$. We have:

$$\sum_{j=1}^{n} y_{ij} = \sum_{j=1}^{n}\xi_2(A_i\times B_j) = \xi_2(A_i\times B) = \mu(A_i),$$

$$\sum_{i=1}^{m} y_{ij} = \sum_{i=1}^{m}\xi_2(A_i\times B_j) = \xi_2(A\times B_j) = \nu(B_j).$$

Hence y_{ij} is a feasible solution of the problem (8). Thus

$$\gamma_2 \ge \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}\xi_2(A_i\times B_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}y_{ij} \ge \gamma_1.$$

Let us now show that $\gamma_2 \le \gamma_3$ Consider (x^1_{ij}) an optimal solution for the problem (10). We associate with (x^1_{ij}) a feasible measure ξ for the problem (9) for which

$$\xi(A_i\times B_j) = x^1_{ij} \ \text{ for all } \ i,j.$$

We have:

$$\begin{aligned}
\gamma_3 &= \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}x^1_{ij}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\iint_{A_i\times B_j} M_{ij}\mathrm{d}\xi(x,y)\\
&\ge \sum_{i=1}^{m}\sum_{j=1}^{n}\iint_{A_i\times B_j} c(x,y)\mathrm{d}\xi(x,y)\\
&= \gamma_2.
\end{aligned}$$

□

For the remainder of this section, h represents the maximum length of all A_i, k represents the maximum length of all B_j, $H = \max\{|x| : x \in A\}$ and $K = \max\{|y| : y \in B\}$.

Lemma 3.2. *If $c(x, y) = (y - x)^2$, then $M_{ij} - m_{ij} \le 2(h + k)(H + K)$.*

Proof. We have,

$$\begin{aligned}(y_{i+1} - x_{i+1})^2 - (y_i - x_i)^2 &= (y_{i+1} - x_{i+1} - y_i + x_i)(y_{i+1} - x_{i+1} + y_i - x_i)\\ &\le 2(h + k)(H + K).\end{aligned} \qquad \square$$

Lemma 3.3. *If $|c(x, y)| \le p(x) + q(y)$, then*

$$\gamma \le \gamma_2 + \int_{A^c} p(x)\,\mathrm{d}\mu(x) + \int_{B^c} q(y)\,\mathrm{d}\nu(y).$$

Proof. Consider ξ_2, an optimal solution for the problem (9). If we set

$$\xi = \xi_2 + \frac{\mu' \otimes \nu'}{\mu'(A^c)},$$

where $\mu' = \mu|_{A^c}$ and $\nu' = \nu|_{B^c}$ we have

$$\gamma \le \gamma_2 + \iint_{(A\times B)^c} c(x, y)\,\mathrm{d}\left(\frac{\mu' \otimes \nu'}{\mu'(A^c)}\right) \le \gamma_2 + \int_{A^c} p(x)\,\mathrm{d}\mu(x) + \int_{B^c} q(y)\,\mathrm{d}\nu(y). \quad \square$$

Lemma 3.4. *If the cost function c is superadditive then $\gamma_2 \le \gamma < \infty$.*

Proof. Let ξ be the Fréchet bound which is an optimal solution for problem (7) yielding $\gamma = \iint_{\mathbb{R}^2} c(x, y)\,\mathrm{d}\xi$. We set $\xi_2 = \xi|_{A\times B}$. ξ_2 is a feasible solution for problem (9). Thus

$$\gamma = \iint_{\mathbb{R}^2} c(x, y)\,\mathrm{d}\xi \ge \iint_{A\times B} c(x, y)\,\mathrm{d}\xi_2 \ge \gamma_2.$$

We notice that if the cost function c is represented by the distance $(y - x)^2$, we have $(y-x)^2 \le 2(x^2+y^2)$. Hence $0 \le c(x, y) \le p(x)+q(y)$ with $p(x) = 2x^2 \ge 0$, $q(y) = 2y^2 \ge 0$. If we know that $\int x^2\,\mathrm{d}\mu(x) < \infty$ and $\int y^2\,\mathrm{d}\nu(y) < \infty$, then the condition (5) is fulfilled. $\square$

Theorem 3.1. *If $c(x, y) = (y - x)^2$, $\int_{|x|>M} x^2\,\mathrm{d}\mu(x) < \epsilon/8$ and $\int_{|y|>N} y^2\,\mathrm{d}\nu(y) < \epsilon/8$ then*

$$|\gamma - \gamma_1| \le \epsilon/2 + 2(h + k)(H + K)\mu(X).$$

Proof. By Lemma 3.2 we get the inequality $M_{ij} - m_{ij} \le 2(h + k)(H + K)$. Consequently

$$\gamma_3 - \gamma_1 \le 2(h + k)(H + K)\mu(X). \tag{11}$$

By Lemma 3.3, we have $\gamma \leq \gamma_2 + 2\int_{A^c} x^2 \,\mathrm{d}\mu(x) + 2\int_{B^c} y^2 \,\mathrm{d}\nu(y) \leq \gamma_2 + \epsilon/2$ and

$$\gamma \leq \gamma_2 + \epsilon/2. \tag{12}$$

By Lemma 3.4 and inequality (12), we have $|\gamma - \gamma_2| \leq \epsilon/2$. From Lemma 3.1 and inequality (11), we get $|\gamma_2 - \gamma_1| \leq 2(h+k)(H+K)\mu(X)$. From the last two inequalities, we get the conclusion. □

Algorithm 3. M and N are chosen such that $\mu(A) = \nu(B)$. The cost function $c(x, y)$ is approximated on the sub-rectangles $A_i \times B_j \;\; 1 \leq i \leq m \;, 1 \leq j \leq n$ by

$$c_{ij} = \sup\{c(x, y) : (x, y) \in A_i \times B_j\}.$$

On other sub-rectangles of the subdivision that are outside of $A \times B$, we approximate the cost function by zero as before. We set

$$a_i = \mu(A_i) \;\; 1 \leq i \leq m,$$

$$b_j = \nu(B_j) \;\; 1 \leq j \leq n.$$

We solve the associated Hitchcock problem. We denote by γ^1 its optimal value. Reasoning as before yields that, under same hypotheses on the cost function,

$$|\gamma - \gamma^1| \leq \epsilon/2 + 2(h+k)(H+K). \tag{13}$$

Remark. For each algorithm, we have to choose two numbers M and N such that $\int_{-M}^{M} p(x)\,\mathrm{d}\mu(x) < \epsilon/4$, $\int_{-N}^{N} q(y)\,\mathrm{d}\nu(y) < \epsilon/4$ and $\int_{-M}^{M} p(x)\,\mathrm{d}\mu(x) = \int_{-N}^{N} q(y)\,\mathrm{d}\mu(y)$. If M and N are sufficiently large, the first two conditions are satisfied. If the two measures μ and ν are continuous, it is not difficult to fulfill the third condition: if $\int_{-M}^{M} p(x)\,\mathrm{d}\mu(x) < \int_{-N}^{N} q(y)\,\mathrm{d}\mu(y)$ increase M until we get the required equality and if $\int_{-M}^{M} p(x)\,\mathrm{d}\mu(x) > \int_{-N}^{N} q(y)\,\mathrm{d}\mu(y)$ increase N until we get the required equality. For measures μ and ν which have a discrete part, the previous algorithms should be sligthly modify.

4. Theoretical solution of 4 particular problems

For the problems treated in this section, the given measures μ, ν both have the same distribution (normal), or different distributions (one normal and the other uniform), the cost function is $c(x, y) = (y - x)^2$.

We consider two probability measures defined on the real line whose distribution functions are F and G. The first represents a normal probability distribution $X = N(\mu_1, \sigma_1^2)$ with mean μ_1 and variance ${\sigma_1}^2$, the second represents a normal probability distribution $Y = N(\mu_2, \sigma_2^2)$ with mean μ_2 and variance ${\sigma_2}^2$. We consider the cost function $c(x, y) = (y - x)^2$.

Problem 1. The first problem treated consists in determining the value

$$\gamma = \inf \int (y-x)^2 \mathrm{d}\xi(x,y) \tag{14}$$

where ξ is a measure defined on $\mathbb{R}^2$ whose respective marginals are the two previous measures.

The primal problem (14) is equivalent to the following problem:

$$\gamma = -\sup \int -(y-x)^2 \mathrm{d}\xi(x,y) \tag{15}$$

where ξ is feasible for the problem (14).

Since $-(y-x)^2$ is a superadditive function, the upper Fréchet bound $\xi^*(x,y) = \min\{F(x), G(y)\}$ is an optimal solution for the problem (15) and hence for (14). We notice that the optimal value is theoretically equal to $\int -(y-x)^2 \mathrm{d}\xi^*(x,y)$. As direct evaluation of this value is not practically possible, we notice that solving the problem (14) is equivalent to solving the following problem:

- If $X = N(\mu_1, \sigma_1^2)$ and $Y = N(\mu_2, \sigma_2^2)$ are two normal random variables, how should we choose a random vector $Z = (X, Y)$ in order to minimize $E((X-Y)^2)$?

The response, we suggest, is to set

$$Y = \mu_2 + \sigma_2/\sigma_1(X - \mu_1).$$

We have $Y = a + bX$ where $b = \sigma_2/\sigma_1$ and $a = (\mu_2\sigma_1 - \mu_1\sigma_2)/\sigma_1$. In this case, it is true that Y is a normal random variable $N(\mu_2, \sigma_2^2)$ when X is a normal random variable $N(\mu_1, \sigma_1^2)$.

To prove that $Z = (X, Y)$ is an optimal solution, we show that ξ, the joint distribution of Z, is equal to ξ^*. Let $(x, y) \in \mathbb{R}^2$,

$$\begin{aligned}
\xi(x,y) &= Pr(X \leq x, X \leq (y-a)/b) \\
&= \min\{F(x), F((y-a)/b)\} \\
&= \min\{F(x), G(y)\} = \xi^*(x,y).
\end{aligned}$$

Let us now determine the optimal value of (14). We have:

$$\begin{aligned}
\gamma &= \int_{\mathbb{R}^2} (y-x)^2 \mathrm{d}\xi^*(x,y) \\
&= \int_{\mathbb{R}} y^2 \mathrm{d}G(y) + \int_{\mathbb{R}} x^2 \mathrm{d}F(x) - 2\int_{\mathbb{R}^2} xy \mathrm{d}\xi(x,y) \\
&= \int_{\mathbb{R}} y^2 \mathrm{d}G(y) + (1-2b)\int_{\mathbb{R}} x^2 \mathrm{d}F(x) - 2a\int_{\mathbb{R}} x \mathrm{d}F(x) \\
&= (\sigma_1 - \sigma_2)^2 + (\mu_1 - \mu_2)^2.
\end{aligned}$$

Problem 2. Consider now the Kantorovitch primal problem in the case where the two measures defined on the real line correspond to distinct distributions.

In this problem we determine the minimum value

$$\gamma = \inf \int (y-x)^2 \mathrm{d}\xi(x,y) \tag{16}$$

where ξ is a measure defined on $\mathbb{R}^2$ whose respective marginals are the measures that correspond respectively to a normal random variable $X = N(\mu, \sigma^2)$ and a uniform random variable Y on $[a, b]$. Let's $F(x)$ represent the distribution function of X and $G(y)$ the distribution function of Y.

$$F(x) = \int_{-\infty}^{x} \exp((t-\mu)^2/\sigma^2)\, \mathrm{d}t \,/\, \sqrt{2\pi\sigma^2}$$

and

$$G(y) = \begin{cases} 0 & : \quad y < a \\ (y-a)/(b-a) & : \quad a \le y \le b \\ 1 & : \quad y > b. \end{cases}$$

As before, since $-(y-x)^2$ is a superadditive function, the upper Fréchet bound $\xi^*(x, y) = \min\{F(x),\ G(y)\}$ is an optimal solution of problem (16). In order to determine explicitly the optimal value of the primal problem (16), we notice that solving (16) is equivalent to determining the random vector $Z = (X, Y)$ that minimizes $E((X-Y)^2)$ with X a normal distribution whose distribution function is $F(x)$ and Y the uniform distribution on $[a, b]$ whose distribution function is $G(y)$. If $Y = (b-a)F(X) + a$, then Y is a uniform random variable with distribution function $G(y)$ and $Z = (X, Y)$ minimizes $E((X-Y)^2)$. Let us first show that Y has a uniform distribution with $G(y)$ as its distribution function. In fact

$$\begin{aligned} Pr(Y \le y) &= Pr(X \le F^{-1}((y-a)/(b-a))) \\ &= (y-a)/(b-a) \\ &= G(y). \end{aligned}$$

To prove that $Z = (X, Y)$ minimizes $E((X-Y)^2)$, we show that ξ the joint distribution of Z is equal to ξ^*. Let $(x, y) \in \mathbb{R}^2$. We have:

$$\begin{aligned} \xi(x, y) &= Pr(X \le x, X \le F^{-1}((y-a)/(b-a))) \\ &= \min\{F(x), F[F^{-1}((y-a)/(b-a))]\} \\ &= \xi^*(x, y). \end{aligned}$$

We determine the optimal value of the problem (16). We have:

$$\begin{aligned}\gamma &= \int_{\mathbb{R}^2} (y-x)^2 \mathrm{d}\xi^*(x,y)\\ &= \int_{\mathbb{R}} y^2 \mathrm{d}G(y) + \int_{\mathbb{R}} x^2 \mathrm{d}F(x) - 2\int_{\mathbb{R}^2} xy\mathrm{d}\xi(x,y)\\ &= (b^2+ab+a^2)/3 + \mu^2 + \sigma^2 - 2\int_{\mathbb{R}^2} xy\mathrm{d}\xi(x,y).\end{aligned}$$

Let us evaluate $\int_{\mathbb{R}^2} xy\mathrm{d}\xi(x,y)$.

$$\begin{aligned}\int_{\mathbb{R}^2} xy\mathrm{d}\xi(x,y) &= \int_{\mathbb{R}} x[a+(b-a)F(x)]\mathrm{d}F(x)\\ &= a\mu + (b-a)I/(2\pi\sigma^2)\end{aligned}$$

where $I = \int_{\mathbb{R}}[\int_{-\infty}^{x} \exp(-(t-\mu)/\sigma)^2/2)\mathrm{d}t]x\exp(-(x-\mu)/\sigma)^2/2)\mathrm{d}x$, integration by parts gives $I = \pi\mu\sigma^2 + \sqrt{\pi}\sigma^3$. By using this value of I, we get

$$\int_{\mathbb{R}^2} xy\mathrm{d}\xi(x,y) = \mu(a+b)/2 + \sigma(b-a)/(2\sqrt{\pi}).$$

Thus

$$\gamma = (b^2+ab+a^2)/3 + \mu^2 + \sigma^2 - \mu(a+b) - \sigma(b-a)/\sqrt{\pi}. \tag{17}$$

Problem 3. As a third problem, we determine the optimal value

$$\gamma = \sup \int (y-x)^2 \mathrm{d}\xi(x,y) \tag{18}$$

where ξ is a feasible measure for problem (14). Since $-(y-x)^2$ is a superadditive function, the lower Fréchet bound $\xi_*(x,y) = \max(0, F(x)+G(y)-1)$ is an optimal solution for the primal problem (18). Problem (18) is equivalent to finding a random vector $Z = (X,Y)$ that maximizes $E((X-Y)^2)$ where $X = N(\mu_1, \sigma_1^2)$ and $Y = N(\mu_2, \sigma_2^2)$ are two normal random variables.

Consider the random variable

$$Y = \mu_2 - \sigma_2/\sigma_1(X-\mu_1).$$

We have $Y = a + bX$ where $b = -\sigma_2/\sigma_1$ and $a = (\mu_2\sigma_1 + \mu_1\sigma_2)/\sigma_1$. Thus $Y = N(\mu_2, \sigma_2^2)$ when $X = N(\mu_1, \sigma_1^2)$. In order to prove that $Z = (X,Y)$ is a maximal random vector, we show that the joint distribution ξ of Z is equal to ξ^*.

Let $(x,y) \in \mathbb{R}^2$,

$$\begin{aligned}\xi(x,y) &= \max\{0, F(x) - F((y-a)/b)\}\\ &= \max\{0, F(x)+G(y)-1\} = \xi^*(x,y).\end{aligned}$$

The optimal value of (18) is:

$$\begin{aligned}\gamma &= \int_{\mathbb{R}^2} (y-x)^2 \mathrm{d}\xi^*(x,y),\\ &= \int_{\mathbb{R}} y^2 \mathrm{d}G(y) + \int_{\mathbb{R}} x^2 \mathrm{d}F(x) - 2\int_{\mathbb{R}^2} x(a+bx)\mathrm{d}F(x),\\ &= (\sigma_1+\sigma_2)^2 + (\mu_1-\mu_2)^2.\end{aligned}$$

Problem 4. The last problem treated in this section consists of determining the maximal value

$$\gamma = \sup \int (y-x)^2 \mathrm{d}\xi(x,y) \tag{19}$$

where ξ is a feasible measure for problem (16) This problem is equivalent to the problem of finding a random vector $Z = (X, Y)$ that maximizes $E((X-Y)^2)$ with X a normal random variable whose distribution function is $F(x)$ and Y a uniform random variable whose distribution function is $G(y)$.

Set $Y = (a-b)F(X) + b$. We show that Y is a uniform random variable with distribution function $G(y)$.

$$\begin{aligned}Pr(Y \le y) &= Pr(X \ge F^{-1}((y-b)/(a-b)))\\ &= (y-a)/(b-a) \ = \ G(y).\end{aligned}$$

Also the random vector $Z = (X, Y)$ maximizes $E((X-Y)^2)$. In fact,

$$\begin{aligned}\xi(x,y) &= Pr(X \le x, X \ge F^{-1}((y-b)/(a-b)))\\ &= \max\{0, F(x) - (y-b)/(a-b)\}\\ &= \max\{0, F(x) + G(y) - 1\} = \xi^*(x,y).\end{aligned}$$

The optimal value in problem (19) is:

$$\begin{aligned}\gamma &= \int_{\mathbb{R}^2} (y-x)^2 \mathrm{d}\xi^*(x,y),\\ &= (b^2+ab+a^2)/3 + \mu^2 + \sigma^2 - 2\int_{\mathbb{R}^2} xy \mathrm{d}\xi(x,y),\\ &= (b^2+ab+a^2)/3 + \mu^2 + \sigma^2 - \mu(a+b) + \sigma(b-a)/\sqrt{\pi}.\end{aligned}$$

5. Examples of numerical solution

We consider four examples of numerical solutions, where the intervals A, B are respectively $[\mu_1 - 5\sigma_1, \mu_1 + 5\sigma_1]$ and $[\mu_2 - 5\sigma_2, \mu_2 + 5\sigma_2]$. We divide them to $2n$ equal sub-intervals. The partitions $A_1, ..., A_{2n}$ of A and $B_1, ..., B_{2n}$ of B are chosen such that none of the rectangles $A_i \times B_j$ $i, j \in \{1, ..., 2n\} \times \{1, ..., 2n\}$ are met in their interior by the coordinate axes. This entails that M_{ij} and m_{ij} are found in the four corner points of the rectangle $A_i \times B_j$. The cost matrix, the supplies and

TABLE 1. Values of the deviation $\gamma - \gamma_1$ for examples 1–3 solved by algorithm 2.

n	Example 1	Example 2	Example 3
10	1.0000	0.9461	1.7647
20	0.7500	0.7343	1.2422
40	0.4375	0.4681	0.7386
50	0.3600	0.3934	0.6115
100	0.1900	0.2171	0.3273
200	0.0975	0.1140	0.1693
250	0.0784	0.0921	0.1363
400	0.0494	0.0584	0.0861

TABLE 2. Regressions lines for examples 1–3.

	Example 1	Example 2	Example 3
a	-2.1741	-1.9861	-2.6933
b	0.8472	0.7830	0.8420
r	0.9939	0.9910	0.9954

the demands which are the data of classical transportation problems are computed by a simple program. This program determines first the value of the cumulative function for any real. Those values allow us to determine the supplies and the demands. Assuming that the intervals $[\mu_1 - 5\sigma_1, \mu_1 + 5\sigma_1]$, $[\mu_2 - 5\sigma_2, \mu_2 + 5\sigma_2]$ are subdivided into sub-intervals such that $c(x, y)$ attains its minimum on a corner of the sub-rectangles, we have constructed a function that allows us to obtain the cost matrix. We then solve the Hichcock problem with any efficient program. After using different discretizations, we finish by determining linear regression lines by the method of least squares for the order of the approximations.

The four examples studied are in the category of problem (14) which was solved theoretically in the section of theoretical solution, the cost function being $c(x, y) = (y - x)^2$. In each example the measure μ corresponds to $N(0, 1)$. The measure ν corresponds respectively to $N(1, 1)$, $N(0, 2)$, $N(1, 2)$ and $N(0, 4)$ in the first, second, third and fourth example. The respective theoretical optimal values γ are 1, 1, 2 and 9.

The first three examples were numerically solved with algorithm 2. The first table compares the precision $\gamma - \gamma_1$ where γ_1 is the minimal transportation cost given by algorithm 2 for increasing values n of discretization for each example.

The second table gives the best linear line $y = a + bx$ with its correlation coefficient r according to the method of least squares for the points $(\ln(\gamma - \gamma_1), \ln n)$.

We remark that in the numerical solution of the first three examples by Algorithm 2, the North-West method always provides the optimal solution.

The fourth example was numerically solved by Algorithm 3. We note that in Algorithm 3 during the discretization the cost matrix in the Hitchcock problem

TABLE 3. Error $\gamma_1 - \gamma$ for example 4 and computing time in seconds (Algorithm 3).

n	$\gamma_1 - \gamma$	computing time
10	10.7902	<0.01
20	4.4156	<0.01
40	1.9919	0.07
50	1.5608	0.09
100	0.7485	0.60
200	0.3664	4.54
250	0.2919	8.63
400	0.1812	37.21

doesn't necessarily satisfy the Monge condition which is the equivalent of the superadditivity property to the discrete case, modulo a sign reversal convention, even though the cost function satisfies the inequality $\frac{\partial^2 c}{\partial x \partial y} \leq 0$. In other words the North-West method will not always give the optimal solution. As a matter of fact, in example 4, the North-West method provides an initial solution, but not the optimal solution.

The third table provides the error $\gamma_1 - \gamma$ as given by Algorithm 3 for increasing values n of discretization and the time of computation. The Hitchcock problem was solved by an efficient implementation of the network simplex method [8]. (Characteristics of the computer: Sun Ultra 60, 2 CPU, speed 300 MHZ, 769 Mo of memory.)

6. Analysis of the convergence rate

Consider μ a measure induced by a normal random variable $X = N(\mu_1, \sigma_1^2)$ and ν a measure induced by a normal random variable $Y = N(\mu_2, \sigma_2^2)$. We show in the following paragraphs that if we consider $L = \sqrt{2 \ln n}$, $A = [\mu_1 - L\sigma_1, \mu_1 + L\sigma_1]$ $B = [\mu_2 - L\sigma_2, \mu_2 + L\sigma_2]$ and $A_1, A_2, ..., A_{2n-1}, A_{2n}$ a partition of A where the A_i have lengths $L\sigma_1/n$ and $B_1, B_2, ..., B_{2n-1}, B_{2n}$ a partition of B with the B_i that have lengths $L\sigma_2/n$, then the solution of problem (8) gives a value γ_1 which approaches the optimal value γ of problem (7) as n approaches infinity.

We recall the definition $f(x) \sim g(x)$ (f and g are asymptotic) if and only if

$$\lim_{x \to \infty} f(x)/g(x) = 1.$$

Let us first determine the asymptotic behavior of

$$I(L) = \int_{|x-\mu_1| > L\sigma_1} x^2 \, \mathrm{d}\mu(x) + \int_{|y-\mu_2| > L\sigma_2} y^2 \, \mathrm{d}\nu(y).$$

After some simple computations, we get

$$\sqrt{2\pi}I(L) = ({\mu_1}^2+{\mu_2}^2+{\sigma_1}^2+{\sigma_2}^2)\int_{|t|\geq L}\exp(-t^2/2)\,\mathrm{d}t+({\sigma_1}^2+{\sigma_2}^2)2L\exp(-L^2/2).$$

But as given in [1], $\int_{|t|\geq L}\exp(-t^2/2)\,\mathrm{d}t \sim 2\exp(-L^2/2)/L$; hence

$$I(L) \sim \sqrt{2/\pi}({\sigma_1}^2+{\sigma_2}^2)L\exp(-L^2/2).$$

If we set $L = \sqrt{2\ln n}$, we get

$$I(L) \sim 2/\sqrt{\pi}({\sigma_1}^2+{\sigma_2}^2)\sqrt{\ln n}/n.$$

We now show that γ_1 approaches the optimal value γ when n approaches infinity. By Theorem 3.1,

$$|\gamma-\gamma_1| \leq 1/n + 2(h+k)(H+K).$$

Since $h = L\sigma_1/n \simeq \sqrt{2\ln n}\,\sigma_1/n$ and $k = L\sigma_2/n \simeq \sqrt{2\ln n}\,\sigma_2/n$, we get

$$|\gamma-\gamma_1| \leq [1+2\sqrt{2\ln n}(\sigma_1+\sigma_2)(H(n)+K(n)]/n$$

where

$$H(n) = \max\{|\mu_1-\sqrt{2\ln n}\,\sigma_1|, |\mu_1+\sqrt{2\ln n}\,\sigma_1|\},$$
$$K(n) = \max\{|\mu_2-\sqrt{2\ln n}\,\sigma_2|, |\mu_2+\sqrt{2\ln n}\,\sigma_2|\}.$$

Thus

$$|\gamma-\gamma_1| \leq C(\ln n)/n \tag{20}$$

where C only depends on μ_1, σ_1, μ_2, σ_2 and N, but not on n.

The error $\gamma-\gamma_1$ does not exceed $C\ln n/n$ if $2n$ is the number of sub-intervals used in the discretization for the computation of numerical solution of problem (8). In order to compare this result with Table 2 of Section 5, we can find the regression line $y = \alpha+\beta x$ corresponding to the points $(\ln n, \ln(n\ln n))$, for $n = 10, 20, 40, 50, 100, 200, 250, 400$. we get $\alpha = -0.3294, \beta = 0.7477$ with a correlation coefficient of 99.9%. The parameter β is not too much distant from the values b of examples of Table 2.

7. Conclusion

Finding the numerical solution of a continuous program is not always an easy task. This is the case with the Kantorovitch problem of mass transfer. To our knowledge, very few authors have solved numerically the mass transfer problem. Besides Levin and Milyutin [13], we can cite Anderson and Philpott [2] and also Anderson and Nash [3]. Their solutions use the duality theory. As other authors, Dubuc and Tanguay [6] only discuss cases where the intervals of $\mathbb{R}$ and the cost function are bounded.

We note that we are aware that the proposed algorithms admit a natural generalization for multidimensional problems, but we do not know their rate of convergence.

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables.* Washington, D.C. (1964).

[2] E.J. Anderson and A.B. Philpott, Duality and an algorithm for a class of continuous transportation problems. *Math. Oper. Res.* **9** (1984) 222–231.

[3] E.J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces. Theory and Application.* John Wiley & Sons, Chichester (1987).

[4] P.E. Appell, *Mémoire sur les déblais et les remblais des systèmes continus ou discontinus.* Mémoires présentés par divers savants **29**, 2e série (1887) 181–208.

[5] P.E. Appell, *Le problème géométrique des déblais et remblais.* Gauthier-Villars, Paris (1928).

[6] S. Dubuc and M. Tanguay, Déplacement de matériel continu unidimensionnel à moindre coût. *RAIRO Rech. Oper.,* **20** (1986) 139–161.

[7] M. Fréchet, Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon* **14** (1951) 53–77.

[8] M.D. Grigoriadis, An efficient implementation of the network simplex method. Netflow in Pisa (Pisa, 1983). *Math. Program. Stud.* **26** (1986) 83–111.

[9] F.L. Hitchcock, The distribution of a product from several sources to numerous localities. *J. Math. Phys.* **20** (1941) 224–230.

[10] W. Hoeffding, Masstabinvariante Korrelations-theorie. *Schr. Math. Inst. Univ. Berlin* **5** (1940) 181–233.

[11] L. Kantorovitch, On the translocation of masses. *Doklady Akad. Nauk. SSSR* **37** (1942) 199–201.

[12] H.G. Kellerer, Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** (1984) 399–432.

[13] V.L. Levin and A.A. Milyutin, The Problem of Mass Transfer with a Discontinuous Cost Function and the Mass Statement of the Duality for Convex Extremal Problems. *Uspehi Mat. Nauk.* **34** (1979) 3–68.

[14] G. Monge, Mémoire sur la théorie des déblais et des remblais. *Mém. Math. Phys. Acad. Royale Sci., Paris* (1781) 666–704.

[15] S.T. Rachev and L. Rüschendorf, Solution of some transportation problems with relaxed or additional constraints *SIAM J. Control Optim.* **32** (1994), 673–689.

[16] A.H. Tchen, Inequalities for distributions with given marginals *Ann. Prob.* **8** (1980) 814–827.