Short communication

Remarks on Afriat’s theorem and the Monge–Kantorovich problem

Alexander V. Kolesnikov\textsuperscript{a,}\textsuperscript{*}, Olga V. Kudryavtseva\textsuperscript{b}, Tigran Nagapetyan\textsuperscript{c}

\textsuperscript{a} Higher School of Economics, Moscow, Russia
\textsuperscript{b} Moscow State University, Departments of Economics, Moscow, Russia
\textsuperscript{c} Fraunhofer ITWM, Kaiserslautern, Germany

\section{A R T I C L E I N F O}

\textbf{Keywords:}
Afriat’s theorem
Revealed preferences
Utility functions
Monge–Kantorovich problem
Optimal transportation
Cyclical monotonicity

\section{A B S T R A C T}

The famous Afriat’s theorem from the theory of revealed preferences establishes necessary and sufficient conditions for the existence of utility function for a given set of choices and prices. The result on the existence of a \textit{homogeneous} utility function can be considered as a particular fact of the Monge–Kantorovich mass transportation theory. In this paper we explain this viewpoint and discuss some related questions.

© 2013 Elsevier B.V. All rights reserved.

1. Afriat’s theorem

The first description of the concept of the revealed preferences can be find in the work of Samuelson (1938), where he presented the weak axiom of revealed preferences. The strong axiom of revealed preferences (SARP) was introduced by Houthakker (1950). It was shown by Afriat (1967) that SARP is a necessary and sufficient condition for the existence of an appropriate utility function for a finite set of choices and prices observed (this is the rationalization of the preferences relations). Later Varian (1982, 1983) extended the method of Afriat (1967) by providing tests for homothetic and additive separability and rationalizing models of behavior.

The connection between Afriat’s theorem and the Monge–Kantorovich problem is known (see, for instance, Levin, 1997, Shananin, 2009, Pospelova and Shananin, 1998 for the connection with the so-called Monge–Kantorovich optimal transportation problem). One can also mention the shortest path problem, which is known to be related to Afriat’s theorem after Varian (1982, 1983). This problem has a solution under assumption of the absence of negative cycles, which in turn can be viewed as a “cyclical monotonicity” assumption. See Chapter 9 in Ahuja et al. (1993) for the description of the shortest path problem in terms of the linear programing duality. See also Rochet (1987), where the relation with Rockafellar’s cyclical monotonicity theorem is discussed. Nevertheless, the authors find that an instructive and short description of this relation is somehow missing in the literature.

We fill this gap and, applying some recent results on the Monge–Kantorovich problem, give a complete characterization of the rationalizable data sets from the “transportational” viewpoint. Some related results based on duality, linear programing, cyclical monotonicity, etc., were obtained in Hostel et al. (2004), Brown and Calamiglia (2007), Kannai (2004), Dievert (2012) and Figalli et al. (2011). See also Ekeland and Galichon (2010) for another variational interpretation of Afriat’s theorem. For an account in the Monge–Kantorovich problem the reader is referred to Bogachev and Kolesnikov (2012) and Villani (2003).

In the standard model we have \(m\) different goods and \(n\) observations represented by vectors \(X^i \in \mathbb{R}^m_+\), \(1 \leq i \leq n\)

\[ X^i = (x^i_1, \ldots, x^i_m) \]

with corresponding vectors of prices

\[ P^i = (p^i_1, \ldots, p^i_m) \in \mathbb{R}^m_+ \]
This means that the quantity $x_i$ of the $k$-th good was bought at the price $p_i$. Thus the total amount of money spent by the $i$-th customer equals to 

$$(X', P') = \sum_{j=1}^{n} x_i' p_j.$$ 

**Remark 1.1.** Just for the sake of simplicity we deal with the space $\mathbb{R}^m_+(\mathbb{R}_+ = (0, +\infty))$ of vectors with positive coordinates (zero price and zero consumed amount of any good is prohibited).

A general tool of many classical models in economics is the so-called utility function $u : \mathbb{R}^m_+ \to \mathbb{R}$. Given an utility function $u$ we say that a customer prefers $X \in \mathbb{R}^m_+$ to $Y \in \mathbb{R}^m_+$ iff $u(X) \geq u(Y)$.

**Remark 1.2.** It is a standard and natural assumption in the utility function theory that $u$ is homogeneous: 

$$u(tX) = tu(X), \quad X \in \mathbb{R}^m_+, \quad t \in \mathbb{R}_+.$$ 

In our paper the utility functions we deal with are always homogeneous (except of Section 3).

Under which assumptions on a given data set there exists a utility function that is consistent with this set of observations (choices)? This was the problem solved by Afriat. Let us describe a systematic approach based on natural modeling of customer’s behavior. We usually assume that a given fixed price vector $P$ the customer always chooses the most preferable combination of goods $X$, i.e. $u$ attains its maximal value on the set $(Y : \langle Y, P \rangle \leq \langle X', P' \rangle, \ Y \in \mathbb{R}^m_+)$. 

**Definition 1.3.** We say that the set $\{X_i, P_i\}, \ 1 \leq i \leq n$ admits an utility function $u$ (or $u$ rationalizes this set) if $u(Y) < u(X)$ for every $Y \in \mathbb{R}^m_+$ satisfying $\langle Y, P \rangle \leq \langle X', P' \rangle$.

**Remark 1.4.** Let $u$ be continuous. Then this definition has a simple geometrical meaning: every hyperplane $\{Y : \langle P, Y - X \rangle = 0\}$ is supporting to the set $\{Y : u(Y) \geq u(X)\}$.

Necessary and sufficient condition for the existence of $u$ for a given data set was obtained in Afriat (1967) (see Theorem 2.10 below).

2. Monge–Kantorovich problem

**Remark 2.1.** In contrast to the previous section, we denote below the finite sets in $\mathbb{R}^m \times \mathbb{R}^m$ by $(x_i, y_i)$ instead of $(X', P')$.

In the modern formulation of the Monge–Kantorovich problem one considers a couple of probability measures $\mu$ and $\nu$ on $\mathbb{R}^m$ and a cost function $c(x, y)$. 

**Definition 2.2.** Denote by $P_{\mu, \nu}$ the set of probability measures on $X \times Y = \mathbb{R}^m \times \mathbb{R}^m$ satisfying 

$$P_{\mu, \nu} = \mu \times \nu.$$ 

Here $P_{\mu, \nu}X, P_{\mu, \nu}Y$ are projections of $P$ onto $X, Y$ respectively, i.e. measures defined by 

$$P_{\mu, \nu}(A) = P(A \times Y), \quad P_{\mu, \nu}(B) = P(X \times B).$$ 

The measure $P$ on $\mathbb{R}^m \times \mathbb{R}^m$ solves the Monge–Kantorovich problem if it satisfies the following properties:

1. $P \in P_{\mu, \nu}$
2. $P$ is the minimum of the functional $K(P) = \int c(x, y) \ dP$.

Interpreting $c(x, y)$ as a transportation cost of some production unit from the point $x$ to the point $y$, the integral $\int c(x, y) \ dP$ equals to the total cost of transportation. The measures $\mu$ and $\nu$ are initial and final distributions of the total production respectively.

We give the following example. Let $P$ be the uniform distribution on a discrete set $(x_i, y_i), \ 1 \leq i \leq n$, i.e. $P(x_i, y_i) = \frac{1}{n}$. We consider $P$ to be a candidate to solve the Monge–Kantorovich problem for $\mu = P_{\mu, \nu}X, \ \nu = P_{\mu, \nu}Y$. The total cost equals to 

$$\frac{1}{n} \sum_{i=1}^{n} c(x_i, y_i).$$ 

Now let $\sigma \in S_n$ be any permutation of indices. Take a measure $P_{\sigma}$ which is the uniform distribution on the set $S_n = \{x_\sigma(i)\}, \ 1 \leq i \leq n$.

Note that $P_{\sigma}$ still has the same projections. The new total transportation cost equals to $\frac{1}{n} \sum_{i=1}^{n} c(x_\sigma(i), y_{\sigma(i)})$. Thus, a necessary condition for being optimal in the Monge–Kantorovich sense is the following inequality between total costs:

$$\sum_{i=1}^{n} c(x_i, y_i) \leq \sum_{i=1}^{n} c(x_\sigma(i), y_{\sigma(i)}). \quad (1)$$ 

This observation leads to the following definition:

**Definition 2.3.** A set $A \subset X \times Y$ is called $c$-monotone if every finite subset $(x_i, y_i), \ 1 \leq i \leq n \subset A$ and every permutation $\sigma \in S_n$ satisfies (1).

It is well-known that every permutation $\sigma$ can be decomposed into a product of several cyclical permutations, i.e. permutations of the type $\sigma(i_1) = i_2, \ \sigma(i_2) = i_3, \ldots, \sigma(i_{n-1}) = i_n, \ \sigma(i_n) = i_1$.

This immediately gives us that $c$-monotonicity is equivalent to $c$-cyclical monotonicity.

**Definition 2.4.** A set $A \subset X \times Y$ is called $c$-cyclically monotone if every finite subset $(x_i, y_i), \ 1 \leq i \leq n \subset A$ satisfies 

$$\sum_{i=1}^{n} c(x_i, y_i) \leq \sum_{i=1}^{n} c(x_i, y_{i+1}) \quad (2)$$

with the agreement $y_n = y_1$.

**Definition 2.5.** We say that $(x, y) \in \mathbb{R}^m \times \mathbb{R}^m$ belongs to $c$-superdifferential of a function $u : \mathbb{R}^m \to \mathbb{R}$ if $u(z) \leq u(x) + c(z, y) - c(x, y)$ for every $z \in \mathbb{R}^m$.

The following theorem gives a full characterization of the solutions to the Monge–Kantorovich problem.

**Theorem 2.6.** Let $X = Y$ be a complete, separable, metric space, $\mu$ and $\nu$ be Borel probability measures thereon. Assume that $c(x, y) : X \times Y \to [0, +\infty)$ is a lower semi-continuous nonnegative cost function and $\int c(x, y) \ d\mu(x) \ d\nu(y) < \infty$. Let $\pi$ be a Borel probability measure on $X \times Y$ such that $P_{\mu, \nu} \pi = \mu, P_{\nu, \mu} \pi = \nu$. Then the following statements are equivalent:

1. $\pi$ is the solution to the Monge–Kantorovich problem with marginals $\mu, \nu$ and the cost function $c$;
2. there exists a $c$-cyclically monotone set $\Gamma$ satisfying $\pi(\Gamma) = 1$;
(3) there exists a function \( u \) and a set \( I' \) satisfying \( \pi(I') = 1 \) such that \( I' \) is contained in the \( c \)-superdifferential of \( u \).

**Remark 2.7.** It is easy to check that the theorem holds also for \( c \) uniformly bounded from below: \( c(x, y) \geq K, \ K \in \mathbb{R} \). We will use the theorem in the case of continuous cost functions only.

The facts collected in this theorem are the cornerstones of the Monge–Kantorovich theory. The equivalence of (2) and (3) for \( c(x, y) = (x - y)^2 \) was proved by Rockafellar (see Rockafellar, 1970). The relation (3) \( \implies \) (2) is elementary. Indeed, one has for \((x, y) \in I'\)

\[
u(x_{i+1}) - \nu(x_i) \leq c(x_{i+1}, y_i) - c(x_i, y_i).
\]

Summing up these inequalities one obtains the desired cyclical monotonicity property. The equivalence of (2) and (3) for general \( c \) was obtained by Rüschendorf (1996). The implication (1) \( \implies \) (2) is very well-known and was apparently discovered for the first time by Smith and Knott (1987) for \( c = (x - y)^2 \). The implication (2) \( \implies \) (1) is relatively recent and was proved in sufficient generality in Schachermayer and Teichmann (2009).

**Remark 2.8.** One can always assume that the function \( u \) from the item (3) of Theorem 2.6 is defined on the whole \( \mathbb{R}^m \) and is \( c \)-concave, i.e. there exists a function \( \varphi(y) \) such that \( u(x) = \inf_{y \in \mathbb{R}^m} c(x, y) - \varphi(y) \).

Let us discuss connections of the Monge–Kantorovich theory with the revealed preferences. The key observation here is the following proposition.

**Proposition 2.9.** The set of data \( \{(X_i, P_i)\} \) admits a positive homogeneous utility function \( u \) if and only if it is contained in the \( c \)-superdifferential of the function \( v = \log u \) for \( c(x, y) = \ln(x, y) \).

**Proof.** Note that the relation

\[
\langle P_i, X_i \rangle < \langle P_i, X_i' \rangle \implies \langle u(Z) \rangle < \langle u(X_i') \rangle
\]

for a positive homogeneous function \( u \) is equivalent to the following inequality:

\[
\frac{u(X_i)}{\langle P_i, X_i \rangle} \geq \frac{u(Z)}{\langle P_i, Z \rangle}.
\]

Indeed, (3) follows immediately from (4).

Assume that (3) holds. Take any \( X_i \) and \( Z \) and find positive \( \lambda \) such that \( \langle P_i, \lambda \cdot Z \rangle = \langle P_i, X_i' \rangle \). Then it follows from (3) that

\[
u(X_i) - \nu(Z) \geq \log(\langle P_i, X_i' \rangle) - \log(\langle P_i, Z \rangle).
\]

\[
\text{using that } u \text{ is homogeneous and } \lambda = \frac{\langle P_i, X_i' \rangle}{\langle P_i, Z \rangle}, \text{ we immediately get (4)}.
\]

We finish the proof with the observation that (4) is equivalent to the inequality

\[
\nu(X_i) - \nu(Z) \geq \log(\langle P_i, X_i' \rangle) - \log(\langle P_i, Z \rangle)
\]

for \( v = \log u \) and every \( Z \). The proof is complete. \( \square \)

We are already ready to get Afriat’s theorem from Theorem 2.6. To this end we identify the set \( \{X_i\}, \ 1 \leq i \leq n \) with the probability measure

\[
\mu = \frac{1}{n} \delta_{X_i},
\]

where \( \delta_{X_i} \) is the Dirac measure concentrated in \( X_i \). Similarly

\[
v = \frac{1}{n} \delta_{P_i}.
\]

Finally, let \( \pi \) be a measure on \( \mathbb{R}^m_+ \times \mathbb{R}^m_+ \) defined by

\[
\pi = \frac{1}{n} \delta_{(X_i, P_i)}.
\]

**Theorem 2.10** (Generalized Afriat’s Theorem, Discrete Case). Let \( (X_i, P_i) \subset \mathbb{R}^m_+ \times \mathbb{R}^m_+ \) be a finite set, \( c(x, y) = \ln(x, y) \). The following statements are equivalent:

1. \( \pi \) is the solution to the Monge–Kantorovich problem with marginals \( \mu, \nu \) and the cost function \( c \);
2. (the set \( \{X_i, P_i\} \) is \( c \)-cyclically monotone;
3. the set \( \{X_i, P_i\} \) admits a positive homogeneous utility function \( u \).

**Proof.** By Proposition 2.9 (3) is equivalent to the property that the set \( \{X_i, P_i\} \) is included in the \( c \)-superdifferential of \( u = \log u \).

Hence the statement is a particular case of Theorem 2.6. \( \square \)

**Remark 2.11.** The cyclical monotonicity for \( c(x, y) = \ln(x, y) \) (property 2) is equivalent to the following inequality for any \( k \) different indexes \( i_1, i_2, \ldots, i_k \):

\[
\langle \langle P_{i_1}, X_{i_1} \rangle, \langle P_{i_2}, X_{i_2} \rangle, \ldots, \langle P_{i_k}, X_{i_k} \rangle \rangle \leq \langle \langle P_{i_1}, X_{i_1} \rangle, \langle P_{i_2}, X_{i_2} \rangle, \ldots, \langle P_{i_k}, X_{i_k} \rangle \rangle.
\]

(5)

The latter is known as a homogeneous axiom of revealed preferences (HARP).

**Remark 2.12.** In the homogeneous case HARP is equivalent to SARP (see, Varian, 2006).

3. Monge–Kantorovich problem in the non-homogeneous case

One can ask the following natural question. Assume that we are given a non-homogeneous rationalizable discrete data \( (X_i, P_i) \). Whether exists a cost function \( c(x, y) \) such that the corresponding utility function \( u \) is a potential for some Monge–Kantorovich with the cost function \( c(x, y) \)? The answer is yes, but \( c \) highly depends on the data set in general (unlike the homogeneous case where one can always set \( c = \ln(x, y) \)). Indeed, it is known that for every rationalizable \( \{X_i, P_i\} \), there exist positive numbers \( s_i \) such that the following system of inequalities has a solution:

\[
y_j - y_i \leq s_i \langle P_i, X_i' - X_i \rangle.
\]

(6)

This is the most difficult step in the proof of the general Afriat’s theorem (see Fostel et al., 2004 for relatively short arguments).

We set

\[
u(X_i) = y_i, \quad c(X_i, P_i) = s_i \langle P_i, X_i \rangle.
\]

One can extend \( u \) to \( \mathbb{R}^m_+ \):

\[
u(x) = \min_{1 \leq i \leq n} \{s_i \langle P_i, x - X_i \rangle \}.
\]

Clearly, (6) means that \( \{X_i, P_i\} \) is included in the \( c \)-superdifferential of \( u \). By Theorem 2.6 the data set \( \{X_i, Y_i\} \) is a support of a measure \( \pi \) solving some optimal transportation problem for the cost function \( c \).

4. Continuous case and optimal transportation

In this section we deal only with the cost function \( c(x, y) = \ln(x, y) \).

**Theorem 2.10** has a natural generalization to the non-discrete case. Consider a non-finite (even non-countable) data of observations

\[
D = \{(x_i, y_i) \subset \mathbb{R}^m \times \mathbb{R}^m, \ i \in I\}.
\]

As we have seen in the previous section, it is convenient to deal with probability measures on \( D \). Thus we assume that a probability measure \( \pi \) on \( S \) is given. All the statements below are formulated up to a set of zero measure. The projection of \( \pi \) are denoted by \( \mu \) and \( \nu \) respectively.
Just for the technical reasons and for the sake of simplicity we will assume in this section the following:

**Assumption.** There exists a compact set \( K \subset \mathbb{R}^m \times \mathbb{R}^m \) such that \( \pi(K) = 1 \).

**Remark 4.1.** Under this assumption the cost function \( c(x, y) \) is continuous on the support of \( \pi \). This makes applicable all the theorems from the previous section.

**Definition 4.2.** We say that \( \pi \) admits a utility function \( u \) if and only if for \( \pi \)-almost all \((x, y)\) and every \( z \subset \mathbb{R}^m \) one has
\[

u(z) < u(x) 
\]
provided \( (x, z) < (x, y) \).

The following result is just the continuous version of Theorem 2.10 and the proof follows the same arguments.

**Theorem 4.3** (Generalized Afriat’s Theorem, Continuous Case). Let \( c(x, y) = \ln(x, y) \). The following statements are equivalent:

1. \( \pi \) is the solution to the Monge–Kantorovich problem with marginals \( \mu, \nu \) and the cost function \( c \).
2. There exists a \( c \)-cyclically monotone set \( F \) satisfying \( \pi(F) = 1 \).
3. \( \pi \) admits a positive homogeneous utility function \( u \).

Let us make an important remark on the structure of the optimal solutions. Let \( S = S^{m-1} = \{ x \in \mathbb{R}^m \setminus \{0\} : |x| = 1 \} \) be the \( m-1 \)-dimensional sphere of radius 1. Let \( P_3 \) be the projection on \( S^{m-1} : \)
\[
P_3(x) = \frac{x}{|x|} \in \mathbb{R}^m.
\]
In the same way we set
\[
P_2(y) = \frac{y}{|y|} \in \mathbb{R}^m.
\]
We denote by \( \mu_S = \mu \circ P_3^{-1} \) the projection of \( \mu \) onto \( S \), i.e., a measure on \( S \) which is defined by the formula
\[
\mu_S(A) = \mu(P_3^{-1}(A)).
\]
Here \( A \subset S \) is an arbitrary Borel set and \( P_3^{-1}(A) = \{ z : P_3(z) \in A \} \) is the preimage of \( A \) under \( P_3 \). In the same way we define \( \nu_S \) and \( \pi_{S \times S} = \pi \circ (P_3^{-1}(x), P_3^{-1}(y)) \).

It is clear that the marginals \( \mu \) and \( \nu \) of the problem of minimizing \( \int \ln(x, y) \, d\pi \) is equivalent to the problem of minimizing of \( \int \ln(x, y) \, d\tau \). Indeed, this follows from the relation
\[
\int \ln \frac{x}{|x|} \frac{y}{|y|} \, d\tau = \int \ln(x, y) \, d\tau - \int \log |x| \, d\mu - \int \log |y| \, d\nu
\]
and the fact that the quantities \( \int \log |x| \, d\mu, \int \log |y| \, d\nu \) are fixed. This means that \( \pi \) is \( c \)-optimal if and only if its projection \( \pi_{S \times S} \) on \( S \times S \) is optimal for the marginals \( \mu_S, \nu_S \) and the cost function \( \ln(x, y) \).

Now let us assume that \( \pi_S \) and \( \nu_S \) have densities with respect to the surface measure \( \sigma \) on \( S \):
\[
\nu_S = f \cdot \sigma, \quad \nu_S = g \cdot \sigma.
\]
Then it is well-known (see Villani, 2003; Bogachev and Kolesnikov, 2012 and the references therein) that there exists a mapping \( T : S \to S \) with the following property:
\[
\pi(F) = 1, \quad F = \{(x, T(x)), x \in S \}.
\]
In particular, \( \pi \)-almost all points \((x_i, y_i)\) satisfy the relation \( y_i = T(x_i) \) and \( \nu_S \) is the image of \( \mu_S \) under \( T \) in the following sense:
\[
\nu_S(T(A)) = \mu_S(A),
\]
where \( T(A) = \{ y : y = T(x) \) for some \( x \in A \) \) for every Borel set \( A \subset S \). The mapping \( T \) is called optimal transportation mapping. It can also be identified with the inverse demand function.

It is easy to understand the relation of \( T \) with the utility function \( u \). If \( u \) is differentiable at the point \( x \) (this fails actually only on a set of \( \mu \)-measure zero), then the hyperplane \( L \) given by the equation \((z - x_i, y_i) = 0\) touches the level set \( u \) exactly at the point \( x_i \). Hence \( \nabla u(x_i) \) is the normal vector of \( L \) satisfying
\[
\nabla u(x_i) - \nabla u(x_i) = -\langle x_i, y_i \rangle.
\]

**Conclusion:** \( \nu_S \) is the image of \( \mu_S \) under the mapping \( x \to \nabla u(x_i) \cdot x \in S \).

We note that \( T(x) \) coincides with the normal vector to the surface \((y : y = u(x)) \) taken at the point \( x \). It follows from Remark 2.8 that this surface can be assumed convex (meaning that the set \((y : y \geq u(x)) \) is convex). This provides a relation with the so-called Alexandrov’s problem.

For a convex surface \( F \subset \mathbb{R}^m \) we consider its normal mapping into the sphere \( S : F \ni x \mapsto N(x), \) where \( N(x) \) is the normal to the tangent plane to \( F \) at the point \( x \). Suppose that the origin is inside of \( F \). Then \( F \) can be parameterized by means of a radial function: \( F = (r(x)) \) \( x \in S \). Let us define a mapping \( T_3 : S \to S \), \( T_3(x) = N(r(x)) \).

**Definition 4.4.** Let \( \mu, \nu \) be a couple of probability measures \( \mu, \nu \) on \( S \). We say that a convex surface \( F \) is a solution to an Alexandrov’s problem if \( \nu \) is the image of \( \mu \) under \( T_3 \).

A generalized version of this problem was posed and solved by A.D. Alexandrov in Alexandrov (1996). Rewriting this problem analytically one gets a kind of Monge–Ampère equation which involves the Gauss curvature of \( F \). It was shown by Oliker (2007) (see also recent development in Bertrand, 2012) that the Monge–Kantorovich problem for the function \( c(x, y) = -\log(x, y) \) can be used to construct the solution to Alexandrov’s problem. Note that our situation is almost the same; the only difference is the sign of the cost function.

**Remark 4.5.** It is easy to see that the whole theory concerning revealed preferences can be extended in the same way if instead of the standard scalar product one considers any function \( b(x, y) \) which is homogeneous in both variables: \( t b(x, y) = b(tx, ty), t \geq 0 \). Namely, given a data set \( \{(x_i, y_i)\} \) one tries to find a homogeneous function \( u \) with the property
\[
b(x_i, y_i) > b(z, y_i) \implies u(z) < u(x_i).
\]
This problem can be reduced to the optimal transportation problem for the cost function \( c(x, y) = \log b(x, y) \).

5. **Negative cycles and potential fields**

Let us consider again a general cost function \( c(x, y) \) and a couple of probability measures \( \mu, \nu \). We assume that \( \mu \) and \( \nu \) have densities with respect to the Lebesgue measure. Let \( T \) be the optimal transport of \( \mu \) onto \( \nu \). We will assume that both \( c(x, y) \) and \( T \) are sufficiently regular. It follows from the general Kantorovich duality statement and from Theorem 2.6(3) as well that \( T \) and \( u \) are related by the formula
\[
\nabla_s c(x, T(x)) = \nabla u(x)
\]
Warshall–Floyd algorithm allows finding a shortest path between every pair of vertices in a directed graph. It starts with initializing the distance matrix with the direct edge weights and then iteratively updates the distances between all pairs of vertices until no further updates are needed. This process ensures the absence of negative cycles, which is equivalent to finding the shortest path. The algorithm is of great importance in various fields, including computer science and operations research.

The absence of negative cycles in a graph is crucial for the correctness of the Warshall–Floyd algorithm. If negative cycles exist, the algorithm may not correctly compute the shortest path, as the path lengths can become infinite.

Acknowledgments

This study was carried out within “The National Research University Higher School of Economics” Academic Fund Program in 2012–2013, research grant No. 11-01-0175. The first author was supported by RFBR projects 10-01-00518, 11-01-90421-Ukr-f-a, and the program SBF 701 at the University of Bielefeld. The second author was supported by RFBR project 12-06-33030a. We thank Professor Atsushi Kaji for his interest and stimulating discussions.

References

A. V. Kolesnikov et al. / Journal of Mathematical Economics 49 (2013) 501–505

(see, for instance formula (2.63) in Villani, 2003). In particular, \( V_c(x, T(x)) \) is a potential vector field and the integral

\[
\int \nabla_c(x, T(x)) \, d\gamma = u'(\gamma(1)) - u'(\gamma(0))
\]

along any smooth path \( \gamma : [0, 1] \to \mathbb{R}^d \) depends on \( \gamma(0) \) and \( \gamma(0) \) only.

This observation can be interpreted as a continuous analog of the well-known statement from optimization combinatorics: a discrete graph admits a shortest path for every couple of vertices if and only if it has no negative cycles. Given a finite data \((x_i, y_i)\) let us endow every edge \((x_i, x_j)\) of the directed graph with weights \(x_{ij}, \ 1 \leq i, j \leq n\), with the “distance” \(a_{ij} = c(x_i, y_i) - c(x_j, y_j)\) (the number \(a_{ij}\) is allowed to be negative). The sequence \(x_1, x_2, \ldots, x_n\) is called negative cycle if

\[
\sum_{k=1}^{n} a_{x_k x_{k+1}} < 0, \quad i_{n+1} = i_1.
\]

It follows immediately from the definition that the absence of the negative cycles is equivalent to the \(c\)-cyclical monotonicity of the data set. In the absence of negative cycles every two vertices admit a shortest path joining them. A classical computational algorithm for finding the shortest path based on dynamical programing principle is the \(\text{Warshall–Floyd}\) algorithm.

Now let us assume that we have a continuous \(c\)-cyclically monotone data set \(D = \{(x, p(x)), x \in X \subset \mathbb{R}^d\}\), where \(p : X \to \mathbb{R}^d\) is a sufficiently regular mapping, and a smooth path \(\gamma : [0, 1] \to X\) with \(\gamma(0) = \gamma(1)\). Let us pick numbers \(x_i = \gamma(i/n), 1 \leq i \leq n\). By the \(c\)-cyclical monotonicity

\[
0 \leq -\frac{1}{n} \sum_{i=1}^{n} \left( c(x_{i+1}, y_i) - c(x_i, y_i) \right).
\]

Passing to the limit \(n \to \infty\) we get \(\int \nabla_c(x, p(x)) \, d\gamma \geq 0\). Running the cycle in the opposite direction we get in the same way

\[
-\int \nabla_c(x, p(x)) \, d\gamma \geq 0.
\]

Finally we get a continuous version of the “absence of negative cycles” principle: if \(D\) is \(c\)-cyclically monotone, then

\[
\int \nabla_c(x, p(x)) \, d\gamma = 0.
\]

Thus every “continuous cycle” is zero in the smooth setting. Hence

\[
\nabla_c(x, p(x)) = \nabla u
\]

for some potential \(u\). Clearly, all the paths joining two points \(x_0, x_1\) has the same “length” \(\int \nabla_c(x, p(x)) \, d\gamma\).

In particular, we get for \(c = \log(x, y)\) that

\[
\frac{p(x)}{\langle x, p(x) \rangle} = \nabla u(x).
\]

This relation has been studied systematically in Shananin (2009) and Pospelova and Shananin (1998) from the viewpoint of the theory of index numbers.

References