Characteristic Functions of Directed Graphs and Applications to Stochastic Equilibrium Problems

Yu. Nesterov *

February 3, 2003

*CORE, Catholic University of Louvain, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium; e-mail: nesterov@core.ucl.ac.be

Abstract

In this paper we introduce the notions of characteristic and potential functions of directed graphs and study their properties. The main motivation for our research is the stochastic equilibrium traffic assignment problem, in which the drivers choose their routes with some probabilities. Since the number of the strategies in this game is very big, we need to find an efficient way of computation of the expected arc flows in the network. We show that the characteristic functions of the graphs are very useful in this respect. Using this technique we can form and solve numerically the equilibrium traffic assignment problem in a reasonable computational time. As a byproduct of our results we show that the spectral radius of a matrix with non-negative elements admits a convex parametrization as a function of its entries.

1 Introduction

Let us start by giving some motivation for this paper. In many models of Game Theory with independent players it is natural to introduce some uncertainty in the decision making process. If a player needs to choose one of M strategies with the actual cost values $\{c_r\}_{r=1}^M$, then his/her perceptions may be affected by some random errors. As a result, the value of the chosen strategy may be different from the optimal value $\min\limits_{1\le r\le M} c_r$. One way to treat such a situation is as follows. We assume that the player compares the values c_r affected by some random additive errors ϵ_r. In this case, for any variant r we can speak about some probability p_r for the variant to be chosen by the player. Of course, these probabilities depend on distribution of ϵ_r and usually it is difficult to obtain a closed-form expression for them. However, in some cases this can be done. For example, if we assume that the distribution for all ϵ_r is double exponential, we arrive at the following expression (see, e.g. [3]):

$$p_i = e^{-c_i/\mu} / \sum_{r=1}^{M} e^{-c_r/\mu}, \quad i = 1, \dots, M, \tag{1.1}$$

where $\mu > 0$ is a distribution parameter. Rule (1.1) is called the *logit model*. One of the most important features of this relation is that it can be written as the gradient of a convex function dependent on $c = (c_1, \dots, c_M)^T$. Indeed, let us introduce a potential function

$$\psi(c) = \mu \ln \left(\sum_{r=1}^{M} e^{-c_r/\mu} \right). \tag{1.2}$$

Let $\nabla\psi(c)$ denote the gradient of the function $\psi(\cdot)$. And let $p = (p_1, \dots, p_M)^T$. Then (1.1) can be written as

$$p + \nabla\psi(c) = 0. \tag{1.3}$$

Note that from the computational point of view, the representation (1.2), (1.3) is acceptable only if the number of strategies M is relatively small. However, in some important classes of models this is definitely not the case. For example, in many transportation models the strategy of a player is a route in a network. Since the number of such routes grows exponentially with the network size, the number of different strategies of the player is very big[1] even for relatively small networks. Thus, in such a situation any model based on the direct formulas (1.2), (1.3) becomes computationally intractable.

The question we are trying to answer in this paper is as follows:

> *Is it possible to have some simple closed-form expressions for the potential function of the type (1.2) when the set of strategies corresponds to some set of routes in the general network?*

We show that this question has a positive answer for some natural "large" sets of routes, like "*all routes in the network*", or "*all routes with uniformly bounded length*". In order to prove this claim, we develop a special technique based on *characteristic functions of directed graphs*. This technique is presented in Section 2. As a byproduct of our results we prove that a spectral radius of a non-negative matrix admits a convex parametrization as a function of elements of the matrix. In Section 3 we present closed-form expressions

[1] Or even infinite, if we allow routes with cycles.

for the expected flows in the network when the behavior of drivers is described by a logit model. In Section 4 we apply the above technique in order to derive a stochastic version of the Stable Dynamic Model (see [5], [6]) for equilibrium traffic assignment. In Section 5 we discuss the possibilities for efficient computation of the gradients of the potential functions. This is necessary in numerical schemes for finding the solutions of the stochastic equilibrium problems. Finally, in the last Section 6 we present some simple stochastic assignment models with incomplete data about the origin-destination flows.

In this paper we often work with log-convex functions. Note that log-convexity is a stronger property than usual convexity. Since such functions are not very common in the literature, we present some simple properties of log-convex functions in an Appendix.

Remark. When the paper was ready for publication, we discovered that the potential function for all (cyclic) routes in the network was already proposed in [1, 2]. However, the way of its derivation differs from our analysis. In any case, the notion of characteristic function and all results related to the sets of routes with uniformly bounded length seem to be new.

2 Characteristic functions of directed graphs

Consider a network $\mathcal{N}$, which consists of n nodes and m directed arcs. Denote by $\mathcal{A}$ the set of all arcs in $\mathcal{N}$. For arc $\alpha \in \mathcal{A}$, we introduce a *travel cost* value $t^{(\alpha)}$. In the majority of applications it is natural to assume that $t^{(\alpha)} \geq 0$. However, in this section that assumption is not necessary. We treat the set of all travel cost values as a column vector $t \in R^m$. In some applications it is important to have the possibility to link two nodes by several parallel arcs. Denote by $\mathcal{A}[i,j]$ the set of all arcs which go from node i to node j. For any $\alpha \in \mathcal{A}$ we denote its root node by $o(\alpha)$ and its end node by $d(\alpha)$.

Let us describe first the objects we are going to work with. Consider a finite sequence of arcs $r = \{\alpha_s\}_{s=1}^l$ with the following property:

$$d(\alpha_s) \quad = \quad o(\alpha_{s+1}), \quad s = 1, \dots, l-1.$$

We call this sequence the *route* in $\mathcal{N}$ connecting the nodes $o(\alpha_1)$ and $d(\alpha_l)$. The value l is called the *length* of the route; sometimes for this length we use notation $L(r)$. Note that such a route can contain cycles. Further, for a route r we can introduce its *travel cost function*:

$$c_r(t) = \sum_{\alpha \in r} t^{(\alpha)}.$$

Note that $c_r(t)$ is a linear function of t.

Let us fix two nodes p and k and denote by $\mathcal{R}$ some *set of routes* in $\mathcal{N}$ (finite or infinite), which connect these two nodes. We can introduce formally the following *characteristic function*:

$$g_{\mathcal{R}}(t) = \sum_{r \in \mathcal{R}} e^{-c_r(t)}. \tag{2.1}$$

If $\mathcal{R} = \emptyset$, we set $g_{\mathcal{R}}(t) \equiv 0$. Thus, if this function is well defined, either $g_{\mathcal{R}}(t) > 0$ for all values of t, or it is identically equal to zero. Let us mention some simple properties of characteristic functions.

- If $\mathcal{R} = \mathcal{R}_1 \bigcup \mathcal{R}_2$ then $g_{\mathcal{R}}(t) = g_{\mathcal{R}_1}(t) + g_{\mathcal{R}_2}(t)$, $\mathrm{dom}\, g_{\mathcal{R}} = \mathrm{dom}\, g_{\mathcal{R}_1} \bigcap \mathrm{dom}\, g_{\mathcal{R}_2}$.
- If $\mathcal{R}_1 \subseteq \mathcal{R}_2$, then $g_{\mathcal{R}_1}(t) \leq g_{\mathcal{R}_2}(t)$, $\mathrm{dom}\, g_{\mathcal{R}_1} \supseteq \mathrm{dom}\, g_{\mathcal{R}_2}$.
- If $\bar{t} \in \mathrm{dom}\, g_{\mathcal{R}}$ and $t \geq \bar{t}$ then $t \in \mathrm{dom}\, g_{\mathcal{R}}$.

Note that for any finite $\mathcal{R}$ the characteristic function is well defined and it is a continuous infinitely times differentiable function of $t \in R^m$.

Finally, for the above set of routes $\mathcal{R}$ we can also formally define the corresponding *potential function*:

$$\psi_{\mathcal{R}}(t) = \ln g_{\mathcal{R}}(t).$$

Define $\psi_{\emptyset}(t) \equiv -\infty$. For $\mathcal{R} = \mathcal{R}_1 \bigcup \mathcal{R}_2$ we have $\psi_{\mathcal{R}}(t) = \ln \left(e^{\psi_{\mathcal{R}_1}(t)} + e^{\psi_{\mathcal{R}_2}(t)} \right)$.

Lemma 1 *If $\mathcal{R} \neq \emptyset$, then $\psi_{\mathcal{R}}(t)$ is a convex function of t.*

Proof:
In other words, we need to prove that $g_{\mathcal{R}}(t)$ is log-convex. Indeed, any non-zero characteristic function $g_{\mathcal{R}}(t)$ is a sum of functions of the form $e^{-c_r(t)}$. Since $c_r(t)$ is linear in t, each of these components is log-convex in t. Therefore, the characteristic function itself is also log-convex in t (see Item 3 in Appendix). □

There is an interesting relation between the potential functions and the shortest path distances in the network $\mathcal{N}$. Denote by $SP_{\mathcal{R}}(t)$ the shortest path distance computed for the routes $r \in \mathcal{R}$:

$$SP_{\mathcal{R}}(t) = \inf\{c_r(t) : \ r \in \mathcal{R}\}.$$

Let $\langle \cdot, \cdot \rangle$ be the standard inner product in R^m:

$$\langle x, y \rangle = \sum_{i=1}^{m} x^{(i)} y^{(i)}, \quad x, y \in R^m.$$

For a function $\phi(t)$, $t \in R^m$, we use notation $\nabla \phi(t)$ in order to denote its gradient with respect to $\langle \cdot, \cdot \rangle$. Note that the gradient is a vector in R^m.

Lemma 2 *Let $\mathcal{R}$ be non-empty and finite. Then for any $t, \bar{t} \in R^m$ we have*

$$\langle \nabla \psi_{\mathcal{R}}(\bar{t}), t \rangle \leq -SP_{\mathcal{R}}(t), \tag{2.2}$$

$$SP_{\mathcal{R}}(t) = -\lim_{\mu \to 0} \mu \psi_{\mathcal{R}}(\bar{t} + t/\mu) = -\lim_{\mu \to 0} \langle \nabla \psi_{\mathcal{R}}(\bar{t} + t/\mu), t \rangle. \tag{2.3}$$

Proof:
In order to prove (2.2), note that $c_r(t)$ is a linear homogeneous function of t. Therefore

$$\langle \nabla \psi_{\mathcal{R}}(\bar{t}), t \rangle = -\sum_{r \in \mathcal{R}} c_r(t) e^{-c_r(\bar{t})} / \left[\sum_{r \in \mathcal{R}} e^{-c_r(\bar{t})} \right] \leq -SP_{\mathcal{R}}(t).$$

Further, $g_{\mathcal{R}}(\bar{t}+t/\mu) = \sum\limits_{r\in\mathcal{R}} e^{-c_r(\bar{t})-c_r(t)/\mu}$ and the set $\mathcal{R}$ is finite. Let us fix some $t \in R^m$. Denote $c^* = SP_{\mathcal{R}}(t)$ and $\mathcal{R}^* = \{r \in \mathcal{R}: \ c_r(t) = c^*\}$. Then

$$
\begin{aligned}
-\mu \ln g_{\mathcal{R}}(\bar{t}+t/\mu) &= -\mu \ln \sum_{r\in\mathcal{R}} e^{-c_r(\bar{t})-c_r(t)/\mu} \\
&= -\mu \ln \left[\sum_{r\in\mathcal{R}^*} e^{-c_r(\bar{t})-c^*/\mu} + \sum_{r\in\mathcal{R}\setminus\mathcal{R}^*} e^{-c_r(\bar{t})-c_r(t)/\mu} \right] \\
&= c^* - \mu \ln \left[\sum_{r\in\mathcal{R}^*} e^{-c_r(\bar{t})} + \sum_{r\in\mathcal{R}\setminus\mathcal{R}^*} e^{-c_r(\bar{t})-(c_r(t)-c^*)/\mu} \right].
\end{aligned}
$$

Note that $c_r(t) > c^*$ for all $r \in \mathcal{R}\setminus\mathcal{R}^*$. Therefore, taking in the last expression the limit in $\mu \to 0$ we get the first equality in (2.3). The second one follows from l'Hopital's rule. □

Finally, let us describe some properties of the gradient of the potential functions. (These properties are non-trivial only if the set $\mathcal{R}$ is infinite.). For each route $r \in \mathcal{R}$ we can introduce its *incidence vector* $a_r \in R^m$. Each entry of this vector corresponds to an arc $\alpha \in \mathcal{A}$ and it is equal to the number of times the arc α is presented in this route. Hence, the definition (2.1) can be rewritten as follows:

$$g_{\mathcal{R}}(t) = \sum_{r\in\mathcal{R}} e^{-\langle a_r,t\rangle}.$$

Lemma 3 *Let $t \in \mathrm{int}\,(\mathrm{dom}\, g_{\mathcal{R}})$. Then the gradient of the potential function $\psi_{\mathcal{R}}(t)$ is well defined by the following expression*

$$\nabla\psi_{\mathcal{R}}(t) = -\tfrac{1}{g_{\mathcal{R}}(t)} \sum_{r\in\mathcal{R}} a_r e^{-\langle a_r,t\rangle}, \tag{2.4}$$

where the series is absolutely convergent. Moreover, for such t we have:

$$A\nabla\psi_{\mathcal{R}}(t) = e_k - e_p, \tag{2.5}$$

where e_i is the ith coordinate vector in R^n and A in an $(n\times m)$ matrix with the following elements:

$$A^{(i,\alpha)} = \begin{cases} 1, & \text{if } o(\alpha)=i, \\ -1, & \text{if } d(\alpha)=i, \\ 0, & \text{otherwise.} \end{cases}, \quad i=1,\dots,n, \ \alpha=1,\dots,m.$$

Proof:
Indeed, for $t \in \mathrm{int}\,(\mathrm{dom}\, g_{\mathcal{R}})$ the series in (2.1) is absolutely convergent. Moreover, since all a_r have non-negative components and all terms in (2.1) are convex in t, the partial sums of (2.4) converge. Therefore the series (2.4) converge to the gradient of function $\psi_{\mathcal{R}}(t)$. Further, since route r connects nodes p and k we have $Aa_r = e_p - e_k$. Hence, we get (2.5) from the fact that the series (2.4) converges. □

In the rest part of this section we are going to derive closed-form expressions for some natural classes of routes in $\mathcal{N}$.

For two nodes p and k denote by $\mathcal{R}^l_{p,k}$, $(l \geq 1)$, the set of all routes in $\mathcal{N}$ of the length l, which connect nodes p and k. Note that for some values of l the set of such routes may be empty. It is important that the above sets can be represented in a recursive way. Indeed, the sets of routes of length one are as follows:

$$\mathcal{R}^1_{p,k} = \mathcal{A}[p,k], \quad p,k = 1,\ldots,n.$$

And for each $l \geq 1$ we clearly have

$$\mathcal{R}^{l+1}_{p,k} = \bigcup_{i\in\mathcal{I}(k)} \bigcup_{r\in\mathcal{R}^l_{p,i}} \bigcup_{\alpha\in\mathcal{A}[i,k]} \{r \cup \alpha\} \quad p,k = 1,\ldots,n, \tag{2.6}$$

where $\mathcal{I}(k) = \{i : \ \mathcal{A}[i,k] \neq \emptyset\}$.

Since the number of elements in $\mathcal{R}^l_{p,k}$ is finite, the corresponding characteristic functions are well defined. It is convenient to consider these functions as entries of some $(n \times n)$-matrix $G_l(t)$:

$$[G_l(t)]^{(k,p)} = g_{\mathcal{R}^l_{p,k}}(t), \quad k,p = 1,\ldots,n.$$

In order to describe the analytical structure of this matrix function, we need to introduce the *incidence matrix function* $E(t)$. That is an $(n \times n)$-matrix with the following entries:

$$[E(t)]^{(i,j)} = \begin{cases} \sum\limits_{\alpha\in\mathcal{A}[j,i]} e^{-t^{(\alpha)}}, & \text{if } \mathcal{A}[j,i] \neq \emptyset, \\ 0, & \text{if } \mathcal{A}[j,i] = \emptyset, \end{cases} \qquad i,j = 1,\ldots,n.$$

Theorem 1 *For any $l \geq 1$ we have $G_l(t) = \underbrace{E(t)\cdots E(t)}_{l \text{ times}} = E^l(t)$.*

Proof:
Let us fix some node p and consider the following vector function

$$a^l_p(t) = G_l(t)e_p,$$

where e_p is the pth coordinate vector in R^n. We prove that

$$a^l_p(t) = E^l(t)e_p, \quad l \geq 1, \tag{2.7}$$

by induction. Indeed, for $l = 1$ we have

$$\left[a^1_p(t)\right]^{(k)} = \begin{cases} \sum\limits_{\alpha\in\mathcal{A}[p,k]} e^{-t^{(\alpha)}}, & \mathcal{A}[p,k] \neq \emptyset, \\ 0, & \mathcal{A}[p,k] = \emptyset. \end{cases} \qquad k = 1,\ldots,n.$$

Thus, $a^1_p(t) \equiv E(t)e_p$. Assume now that (2.7) is true for some $l \geq 1$. Then, in view of definition of the travel cost function $c_r(t)$ and relation (2.6), for any $k = 1,\ldots,n$ we have

$$\begin{aligned} \left[a^{l+1}_p(t)\right]^{(k)} &= \sum_{r\in\mathcal{R}^{l+1}_{p,k}} e^{-c_r(t)} = \sum_{i\in\mathcal{I}(k)} \sum_{r\in\mathcal{R}^l_{p,i}} \sum_{\alpha\in\mathcal{A}[i,k]} e^{-t^{(\alpha)}-c_r(t)} \\ &= \sum_{i\in\mathcal{I}(k)} \sum_{\alpha\in\mathcal{A}[i,k]} e^{-t^{(\alpha)}} a^l_{i,p}(t) = \sum_{i\in\mathcal{I}(k)} [E(t)]^{(k,i)} \cdot \left[a^l_p(t)\right]^{(i)}. \end{aligned}$$

Thus, $g_p^{l+1}(t) = E(t) g_p^l(t)$ and (2.7) follows. Note that (2.7) is equivalent to the statement of the lemma. □

As a corollary of the above theorem, we have the following statement.

Property 1 *An element $[E^l(t)]^{(k,p)} > 0$ if and only if there exists a route of length l connecting the nodes p and k. The value $[E^l(0)]^{(k,p)}$ is equal to the number of different routes of length l from node p to node k.*

Let us look now at the *cumulative* set of routes:

$$\widehat{\mathcal{R}}_{p,k}^L = \bigcup_{l=1}^{L} \mathcal{R}_{p,k}^l.$$

Then for two nodes p and k the *cumulative characteristic function* of order L is defined as follows:

$$g_{\widehat{\mathcal{R}}_{p,k}^L}(t) = \sum_{r \in \widehat{\mathcal{R}}_{p,k}^L} e^{-c_r(t)} = \sum_{l=1}^{L} g_{\mathcal{R}_{p,k}^l}(t). \tag{2.8}$$

Let us introduce the matrix function $\widehat{G}_L(t)$ with the entries

$$\left[\widehat{G}_L(t)\right]^{(k,p)} = g_{\widehat{\mathcal{R}}_{p,k}^L}(t), \quad k,p = 1,\ldots,n.$$

Then, in view of Theorem 1, for any $L \geq 1$ we have the following representation

$$\widehat{G}_L(t) = \sum_{l=1}^{L} E^l(t). \tag{2.9}$$

Thus, in view of (2.9) these matrix functions can be formed in the following recursive way:

$$\left.\begin{array}{rl} \widehat{G}_1(t) & = E(t), \\ \widehat{G}_{L+1}(t) & = E(t)(I + \widehat{G}_L(t)) = \widehat{G}_L(t) + E^{L+1}(t), \ L = 1,\ldots. \end{array}\right\} \tag{2.10}$$

Note that $\widehat{G}_{L+1}^{(k,p)}(t) \geq \widehat{G}_L^{(k,p)}(t) \geq 0$ for all $L \geq 1$, $k,p = 1,\ldots,n$.

Most of the properties of cumulative characteristic and potential function follow directly from our previous results. Denote by $\sigma(p,k)$ the minimal arc distance between the nodes p and k:

$$\sigma(p,k) = \max\{L : \ \mathcal{R}_{p,k}^l = \emptyset \ \forall l < L\}.$$

Property 2 *An element $[\widehat{G}_L(t)]^{(k,p)}$ is a positive function if and only if $L \geq \sigma(p,k)$. The value $[\widehat{G}_L(0)]^{(k,p)}$ is equal to the total number of different routes from p to k, whose length does not exceed L.*

Proof:
Indeed, if $L \geq \sigma(p,k)$ then there exists a route $r \in \mathcal{R}_{p,k}^l$ with $l \leq L$. Therefore the corresponding characteristic function is strictly positive. Conversely, if $L < \sigma(p,k)$, then $\mathcal{R}_{p,k}^l = \emptyset$ for all $l \leq L$ and therefore the corresponding characteristic function is identically zero. □

Property 3 *All non-zero entries of the matrix $\widehat{G}_L(t)$ are log-convex functions of t.* □

Property 4 *Let $L \geq \sigma(p,k)$. Then for any s and t in R^m we have*

$$\langle \nabla \psi_{\widehat{R}^L_{p,k}}(s), t\rangle \leq -SP_{\widehat{R}^L_{p,k}}(t). \tag{2.11}$$

Moreover, for any $\bar{t}$ and t from R^m

$$SP_{\widehat{R}^L_{p,k}}(t) = -\lim_{\mu\to 0} \mu \psi_{\widehat{R}^L_{p,k}}(\bar{t} + t/\mu) = -\lim_{\mu\to 0} \langle \nabla \psi_{\widehat{R}^L_{p,k}}(\bar{t} + t/\mu), t\rangle. \tag{2.12}$$

Proof:
Indeed, the function $\psi_{\widehat{R}^L_{p,k}}(\cdot)$ is well defined if and only if $L \geq \sigma(p,k)$. Therefore we can apply Lemma 2. □

Finally, let us mention some properties of the asymptotic sets of routes:

$$\widetilde{\mathcal{R}}_{p,k} = \bigcup_{l=1}^{\infty} \mathcal{R}^l_{p,k}.$$

For two nodes p and k the *asymptotic characteristic function* is defined as follows:

$$g_{\widetilde{\mathcal{R}}_{p,k}}(t) = \sum_{r\in\widetilde{\mathcal{R}}_{p,k}} e^{-c_r(t)} = \sum_{l=1}^{\infty} g_{\mathcal{R}^l_{p,k}}(t). \tag{2.13}$$

Let us introduce the matrix function $\widetilde{G}(t)$ with the entries

$$\left[\widetilde{G}(t)\right]^{(k,p)} = g_{\widetilde{\mathcal{R}}_{p,k}}(t), \quad k,p = 1,\ldots,n.$$

Then, in view of Theorem 1, we have

$$\widetilde{G}(t) = \sum_{l=1}^{\infty} E^l(t). \tag{2.14}$$

Thus, the convergence of the series in (2.14) depends on the spectral radius of the matrix $E(t)$:

$$\rho(t) = \max_{1\leq i\leq n} |\lambda_i(E(t))|,$$

where $\lambda_i(\cdot)$ are the eigenvalues of corresponding matrix. We need the following auxiliary statement.

Theorem 2 *Function $\rho(t)$ is log-convex in t.*

Proof:
Note that all entries of the matrix $E(t)$ are non-negative. Therefore, in view of the minimax representation of the spectral radius of non-negative matrix, we have:

$$\begin{aligned}\rho(t) &= \min_{x\in R^n_+} \max_{1\le k\le n} \tfrac{1}{x^{(k)}}(E(t)x)^{(k)} = \min_{x\in R^n_+} \max_{1\le k\le n} \tfrac{1}{x^{(k)}} \left[\sum_{i\in\mathcal{I}(k)} \sum_{\alpha\in\mathcal{A}[i,k]} e^{-t^{(\alpha)}} x^{(i)} \right] \\ &= \inf_{y\in R^n} \max_{1\le k\le n} \sum_{i\in\mathcal{I}(k)} \sum_{\alpha\in\mathcal{A}[i,k]} e^{-t^{(\alpha)}+y^{(i)}-y^{(k)}}.\end{aligned}$$

Thus, $\rho(t) = \inf\limits_{y\in R^n} \max\limits_{1\le k\le n} \phi_k(t,y)$ for some functions $\phi_k(t,y)$. Note that each function $\phi_k(t,y)$ is log-convex in (t,y). Therefore $\phi(t,y) = \max\limits_{1\le k\le n} \phi_k(t,y)$ is log-convex in (t,y). Hence, the function $\rho(t) = \inf\limits_{y\in R^n} \phi(t,y)$ is log-convex in t. □

Remark 1 *Note that an arbitrary matrix A with non-negative elements can be represented as $A = E(t)$ with an appropriate $E(\cdot)$ and some $t \in R^m$. Thus, Theorem 2 proves, in particular, that the spectral radius of a non-negative matrix admits a log-convex parametrization as a function of its entries.*

Now we can describe the main properties of the asymptotic characteristic functions.

Theorem 3 *1. The domain of the matrix function $\widetilde{G}(t)$ is an open convex set:*

$$\mathrm{dom}\, \widetilde{G} = \{t \in R^m :\ \rho(t) < 1\}.$$

The matrix $\widetilde{G}(t)$ becomes unbounded when t approaches the boundary of the domain.
2. If $\bar{t} \in \mathrm{dom}\, \widetilde{G}$ and $t \ge \bar{t}$ then $t \in \mathrm{dom}\, \widetilde{G}$.
3. For any $t \in \mathrm{dom}\, \widetilde{G}$ this matrix function can be represented as follows:

$$\widetilde{G}(t) = (I - E(t))^{-1} - I = E(t)(I-E(t))^{-1} = \lim_{L\to\infty} \widehat{G}_L(t). \tag{2.15}$$

Moreover, for such t the sequence of directional derivatives $\{\mathcal{D}\widetilde{G}_L(t)[h]\}$ also converges:

$$\mathcal{D}\widetilde{G}(t)[h] = \lim_{L\to\infty} \mathcal{D}\widehat{G}_L(t)[h], \quad \forall h \in R^m. \tag{2.16}$$

4. For any $t \in \mathrm{dom}\, \widetilde{G}$ the spectral radius of matrix $\widetilde{G}(t)$ is as follows:

$$\tilde{\rho}(t) \equiv \lambda_{\max}\left(\widetilde{G}(t)\right) = \tfrac{\rho(t)}{1-\rho(t)}. \tag{2.17}$$

The function $\tilde{\rho}(t)$ is log-convex on $\mathrm{dom}\, \widetilde{G}$.

Proof:
Indeed, if $\rho(t) < 1$ then the matrix $\widetilde{G}(t)$ is well defined by (2.14). The spectral radius $\tilde{\rho}(t)$ of this matrix is then given by (2.17). Thus, some elements of $\widetilde{G}(t)$ go to infinity when t approaches the boundary of the domain $\{t :\ \rho(t) < 1\}$. In view of Theorem 2

this domain is an open convex set. Note that the function $\tilde{\rho}(t)$ is log-convex since $\rho(t)$ is log-convex and $1-\rho(t)$ is concave.

The statement of Item 2 is valid since $\rho(t) \leq \rho(\bar{t})$ for any $t \geq \bar{t}$.

In order to prove the statement of Item 3, note that for $t \in \operatorname{dom} \widetilde{G}$ we have $\widetilde{G}(t) = \lim\limits_{L\to\infty} \widehat{G}_L(t)$. Thus, each element of the matrix $\widetilde{G}(t)$ is a convex function in t. Moreover, in view of (2.10) the elements of the matrix $\widehat{G}_L(t)$ increase as L increases. Hence, we have

$$\text{epi } \left[\widetilde{G}(t)\right]^{k,p} = \bigcap_{L=1}^{\infty} \text{epi } \left[\widehat{G}_L(t)\right]^{k,p}, \quad k,p = 1,\dots,n,$$

where $\operatorname{epi}\phi(t)$ denotes the epigraph of the function $\phi(t)$:

$$\operatorname{epi}\phi(t) = \{(t,\tau):\ t \in \operatorname{dom}\phi,\ \tau \geq \phi(t)\}.$$

Since each element of the matrix $\widetilde{G}(t)$ is a continuously differentiable function, we have

$$\nabla \left[\widetilde{G}(t)\right]^{(k,p)} = \lim_{L\to\infty} \nabla \left[\widehat{G}_L(t)\right]^{(k,p)},$$

and (2.15) follows. □

Let us look now at the *asymptotic potential functions*:

$$\psi_{\widetilde{R}_{p,k}}(t) = \ln[\widetilde{G}(t)]^{(k,p)}, \quad p,k = 1,\dots,n.$$

Denote $\sigma(\mathcal{N}) = \max\limits_{1\leq p,k\leq n} \sigma(p,k)$. If $\sigma(p,k) < \infty$ and L is large enough, say $L \geq m$, then for $t \geq 0$ we have

$$SP_{\widehat{R}^L_{p,k}}(t) = SP_{\widetilde{R}_{p,k}}(t).$$

Theorem 4 *1. If $\sigma(\mathcal{N}) < \infty$, then the asymptotic potential is well defined for any pair of nodes p and k.*
2. If the asymptotic potential is well defined for some pair of nodes p and k, then it is a convex function in t. Moreover, if the potential is defined for some $\bar{t} \in R^m$, then for any $t \geq 0$ we have

$$SP_{\widetilde{R}_{p,k}}(t) = -\lim_{\mu\to 0} \mu\psi_{\widetilde{R}_{p,k}}(\bar{t}+t/\mu) = -\lim_{\mu\to 0} \langle \nabla\psi_{\widetilde{R}_{p,k}}(\bar{t}+t/\mu), t\rangle.$$

Proof:
If $\sigma(\mathcal{N}) < \infty$ then in view of Property 2 the matrix $\widehat{G}_L(t)$ is strictly positive for any t and L large enough. Hence, $\widetilde{G}(t)$ is a positive matrix and the asymptotic potential function of any entry is well defined.

Let the asymptotic potential $\psi_{\widetilde{R}_{p,k}}(t)$ be well defined for some pair of nodes p and k and some $\bar{t} \in R^m$. Since for any $t \geq 0$ and $\mu > 0$ we have

$$g_{\widetilde{R}^L_{p,k}}(\bar{t}+t/\mu) \leq g_{\widetilde{R}^L_{p,k}}(\bar{t}),$$

the asymptotic potential is well defined at any point $\bar{t} + t/\mu$ with $\mu > 0$. Note that from l'Hopital rule we get

$$\lim_{\mu \to 0} \mu \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu) = \lim_{\mu \to 0} \langle \nabla \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu), t \rangle,$$

provided that the right-hand side of this equality is well defined. Consider the function $\xi(\mu) = \langle \nabla \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu), t \rangle$. Since the asymptotic potential is a convex differentiable function, we have

$$\xi'(\mu) = -\langle \nabla^2 \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu)t, t \rangle / \mu^2 \le 0.$$

On the other hand, in view of relation (2.16) and inequality (2.11) we obtain

$$\xi(\mu) = \langle \nabla \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu), t \rangle = \lim_{L \to \infty} \langle \nabla \psi_{\widehat{R}^L_{p,k}}(\bar{t} + t/\mu), t \rangle \le - \lim_{L \to \infty} SP_{\widehat{R}^L_{p,k}}(t) = -SP_{\widetilde{R}_{p,k}}(t).$$

Thus, we get the relation

$$\lim_{\mu \to 0} \mu \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu) = \lim_{\mu \to 0} \langle \nabla \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu), t \rangle \le -SP_{\widetilde{R}_{p,k}}(t),$$

in which all limits exist. Let us choose now $L = m$. Note that $m \ge \sigma(p, k)$. Therefore the potential $\psi_{\widehat{R}^L_{p,k}}(\cdot)$ is well defined and

$$\psi_{\widetilde{R}_{p,k}}(s) \ge \psi_{\widehat{R}^L_{p,k}}(s), \quad s \in \operatorname{dom} \psi_{\widetilde{R}_{p,k}}.$$

Hence, in view of relation (2.12),

$$\lim_{\mu \to 0} \mu \psi_{\widetilde{R}_{p,k}}(\bar{t} + t/\mu) \ge \lim_{\mu \to 0} \mu \psi_{\widehat{R}^L_{p,k}}(\bar{t} + t/\mu) = -SP_{\widehat{R}^L_{p,k}}(t) = -SP_{\widetilde{R}_{p,k}}(t).$$

□

In the next sections we will consider the applications of the characteristic functions of graphs in some stochastic models.

3 Stochastic route choice model

Let us look at the particular form of the logit model (1.1) as applied to transportation networks. In this situation a driver has to travel from node p to node k. The set of his/her strategies is some set of routes $\mathcal{R}$. The cost of each route r is given by the travel cost function $c_r(t)$. Thus, the probability $p_r = p_r(t)$ for the driver to choose a route $r \in \mathcal{R}$ for transportation is equal to

$$p_r(t) = e^{-c_r(t)/\mu} / \sum_{q \in \mathcal{R}} e^{-c_q(t)/\mu}, \quad r \in \mathcal{R}. \tag{3.1}$$

Let us assume now that we have continuum drivers. Then we have to speak about a continuous *flow* of drivers travelling from p to k. Denote by d the volume of such a flow.

Then, if each driver applies the rule (3.1) for the route choice, the *expected arc flow vector* f_r, ($f_r = f_r(t) \in R^m$), of drivers on the route r is as follows:

$$f_r(t) = d \cdot e^{-c_r(t)/\mu} / \sum_{q \in \mathcal{R}} e^{-c_q(t)/\mu}, \quad r \in \mathcal{R}, \tag{3.2}$$

Note that the number of routes in $\mathcal{R}$ may be very big.[2] In such a situation we need a compact description of the impact of this random flow on the loading in the network. In order to obtain such a description, let us compute the expected arc flow vector $f(t) \in R^m$ generated by the drivers. Clearly, this arc flow vector $f(t)$ can be written as follows:

$$f(t) = \sum_{r \in \mathcal{R}} f_r(t) \cdot a_r, \tag{3.3}$$

where the values $f_r(t)$ are given by (3.2).

On the other hand, it appears that the expression (3.3) can be rewritten in a very compact way. Let us introduce the following potential function

$$\psi_{\mathcal{R}}(t) = \ln \left(\sum_{r \in \mathcal{R}} e^{-c_r(t)} \right). \tag{3.4}$$

Lemma 4 *For any $\mu > 0$ and $t \in R^m$ such that $t/\mu \in \mathrm{int}\,(\mathrm{dom}\,\psi_{\mathcal{R}})$ we have*

$$f(t) = -d \cdot \nabla \psi_{\mathcal{R}}(t/\mu). \tag{3.5}$$

The observed travel time is then $-d\langle \nabla \psi_{\mathcal{R}}(t/\mu), t\rangle$ and $Af(t) = d \cdot (e_k - e_p)$.

Proof:
By definition, $c_r(t) = \langle a_r, t\rangle$. Therefore $\psi_{\mathcal{R}}(t) = \ln \left(\sum_{r \in \mathcal{R}} e^{-\langle a_r, t\rangle} \right)$, and we get (3.5) by usual differentiation rules. The last statement follows from Lemma 3. □

The relation (3.5) gives us a criterion for stochastic transportation models to be computationally tractable. Indeed, if we are able to compute the value of the potential function $\psi_{\mathcal{R}}(t)$, then usually we can compute its gradient. In this case we can compute the expected arc flow vector $f(t)$ and the expected travel time. Of course, the value of the function $\psi_{\mathcal{R}}(t)$ might be computed directly by the expression (3.4). However, that can be done in a reasonable computational time only when the set $\mathcal{R}$ is relatively small. At the same time, in many practical situations we cannot restrict our model to a small number of routes. Fortunately, as we have seen, the situation is not hopeless. The potential functions $\psi_{\mathcal{R}}(t)$ with $\mathcal{R} = \widehat{\mathcal{R}}^L_{p,k}$ or $\mathcal{R} = \widetilde{\mathcal{R}}_{p,k}$, presented in Section 2, provide us with examples of potential functions, for which the set of routes is very big or even infinite. Nevertheless, for both of these functions their values can be computed in a reasonable time (see (2.10), (2.15)). In the next section we discuss the stochastic equilibrium traffic assignment model based on these potential functions.

[2] Or even infinite. Then we need to assume that the series $\sum_{r \in \mathcal{R}} e^{-c_r(t)/\mu}$ converge.

4 Stochastic traffic assignment

In any equilibrium traffic assignment model we have to make a decision on the following components:

- A behavioral principle, which states how the drivers choose their routes.
- Network performance model.
- Description of demands.

Then, in general, the traffic equilibrium assignment problem is as follows:

> *Find the origin-destination flows and the arc travel time values, which satisfy the demands and which are consistent with the route choice principle and the network performance model.*

Let us show how that works in our situation. In deterministic traffic assignment the behavior of drivers is usually described by the Wardrop principle [9]:

> *Any driver chooses one of the shortest paths for travelling between the origin and destination.*

In stochastic framework this principle is replaced by some game-theoretical model of the route choice. As we have seen in Section 3, in the case of logit model we come to the following principle.

> *Let the arc travel time vector t be fixed. Denote by $\mathcal{R}$ the set of possible routes for some origin-destination pair. And let d be the demand flow for this pair. Then the expected arc flow of these drivers is as follows:*
>
> $$f(t) = -d \cdot \nabla \psi_{\mathcal{R}}(t/\mu). \tag{4.1}$$

Note that in our case we have a continuous flow of drivers. Therefore, even if we consider the behavior of each driver as a game with different probabilities for different routes, the only possible visible result of this game is given by the expression (4.1).[3]

In the equilibrium traffic assignment models the performance of the network usually is described by some travel time function $tt(f)$, which tells us what is the travel time on the arc as a function of the flow of drivers on this arc [8]. In order to prove the existence theorem, we need to assume that $tt(f)$ is a non-decreasing function of f. However, in view of recent criticism of this approach [5, 6], in this paper we employ the following *max-flow* model. Let us assume that for each arc $\alpha \in \mathcal{A}$ we know the maximal possible level of flow $\bar{f}^{(\alpha)}$ and the minimal travel time $\bar{t}^{(\alpha)}$. Let $f^{(\alpha)}$ be the observed flow on this arc. Then the travel time $tt^{(\alpha)}$ on this arc satisfies the following axioms:

$$\text{If } f^{(\alpha)} < \bar{f}^{(\alpha)} \text{ then } tt^{(\alpha)} = \bar{t}^{(\alpha)}. \quad \text{If } f^{(\alpha)} = \bar{f}^{(\alpha)} \text{ then } tt^{(\alpha)} \geq \bar{t}^{(\alpha)}. \tag{4.2}$$

[3]To see that, we can divide the demand flow d on N groups of drivers with demand flows d/N for each group. And let each group of drivers chooses its route independently in accordance with the logit model (1.1). Of course, in this case the resulting flow $f(t)$ will be random. However, it is easy to see that the probability that the resulting flow is different from (4.1) goes to zero as N goes to infinity.

It is convenient to use vector notation for these objects:

$$f = (f^{(1)}, \cdots, f^{(m)})^T, \quad \bar{f} = (\bar{f}^{(1)}, \cdots, \bar{f}^{(m)})^T, \quad \bar{t} = (\bar{t}^{(1)}, \cdots, \bar{t}^{(m)})^T.$$

Finally, denote by $\mathcal{OD}$ the set of origin-destination pairs. For each pair $(p,k) \in \mathcal{OD}$ we fix the *demand flow* $d_{p,k}$ and some *set of routes* $\mathcal{R}_{p,k}$ in $\mathcal{N}$, which connect p and k.

Let us prove that the solution of the *stochastic equilibrium traffic assignment* problem (4.1), (4.2) can be found from the following convex minimization problem:

$$\min \left\{ \langle \bar{f}, t \rangle + \mu \psi(t/\mu) : \; t \geq \bar{t} \right\}, \tag{4.3}$$

where $\mu > 0$ and

$$\psi(t) = \sum_{(p,k) \in \mathcal{OD}} d_{p,k} \cdot \psi_{\mathcal{R}_{p,k}}(t).$$

Note that for $R_{p,k} = \widehat{\mathcal{R}}^L_{p,k}$ or $R_{p,k} = \widetilde{\mathcal{R}}_{p,k}$ the value of this function is computable.

Theorem 5 *Let $\bar{t}/\mu \in \mathrm{dom}\, \psi$. 1. Let the demand flow be implementable by some origin-destination flows $f_{p,k}$ such that*

$$\sum_{(p,k) \in \mathcal{OD}} f_{p,k} < \bar{f}.$$

Then the problem (4.3) is solvable.
2. Let t^* be a solution to (4.3). Then the equilibrium origin-destination arc flows are as follows:

$$f^*_{p,k} = -d_{p,k} \cdot \nabla \psi_{\mathcal{R}_{p,k}}(t^*/\mu), \quad (p,k) \in \mathcal{OD}.$$

These flows satisfy the corresponding demands.
3. The equilibrium arc flow $f^* = \sum\limits_{(p,k) \in \mathcal{OD}} f^*_{p,k}$ satisfy the arc flow bounds. Moreover, the pair (t^*, f^*) satisfy the arc performance hypothesis (4.2).

Proof:
Let us prove the first statement of the theorem. In view of our assumption, for each origin-destination pair (p,k) there exist a subset of routes $\mathcal{R}'_{p,k} \subseteq \mathcal{R}_{p,k}$ and a set of positive values $\{\beta^r_{p,k}\}_{r \in \mathcal{R}'_{p,k}}$ such that

$$\sum_{r \in \mathcal{R}'_{p,k}} \beta^r_{p,k} = d_{p,k}, \quad (p,k) \in \mathcal{OD},$$

and such that for origin-destination flows $f_{p,k} = \sum\limits_{r \in \mathcal{R}'_{p,k}} \beta^r_{p,k} a_r$ we have

$$\hat{f} \equiv \sum_{(p,k) \in \mathcal{OD}} f_{p,k} < \bar{f}.$$

Without loss of generality we can assume that any route in $\mathcal{R}'_{p,k}$ has no cycle. Therefore each of these sets contains a finite number of routes.

Let us prove that we can bound the objective function in (4.3) from below by some strictly increasing linear function. Note that

$$
\begin{aligned}
\langle \hat{f}, t\rangle + \mu\psi(t/\mu) &= \langle \hat{f}, t\rangle + \sum_{(p,k)\in\mathcal{OD}} d_{p,k}\cdot \mu\psi_{\mathcal{R}_{p,k}}(t/\mu) \\
&= \langle \hat{f}, t\rangle + \sum_{(p,k)\in\mathcal{OD}} d_{p,k}\cdot \mu \ln\left(\sum_{r\in\mathcal{R}_{p,k}} e^{-\langle a_r, t\rangle/\mu}\right) \\
&\geq \langle \hat{f}, t\rangle + \sum_{(p,k)\in\mathcal{OD}} d_{p,k}\cdot \mu \ln\left(\sum_{r\in\mathcal{R}'_{p,k}} e^{-\langle a_r, t\rangle/\mu}\right) \\
&= \sum_{(p,k)\in\mathcal{OD}} \left[d_{p,k}\cdot \mu \ln\left(\sum_{r\in\mathcal{R}'_{p,k}} e^{-\langle a_r, t\rangle/\mu}\right) + \sum_{r\in\mathcal{R}'_{p,k}} \beta^r_{p,k}\langle a_r, t\rangle\right].
\end{aligned}
$$

Since each of the sets $\mathcal{R}'_{p,k}$ is finite, we can use the inequality

$$
d\ln\left(\sum_{r=1}^{N} e^{-x_r}\right) + \sum_{r=1}^{N}\beta_r x_r \geq -\sum_{r=1}^{N}\beta_r \ln\frac{\beta_r}{d},
$$

which is valid for all $x\in R^N$ and all positive β_r such that $\sum\limits_{r=1}^{N}\beta_r = d$. Denoting by

$$
\gamma_{p,k} = -\sum_{r\in\mathcal{R}'_{p,k}} \beta^r_{p,k}\ln\frac{\beta^r_{p,k}}{d_{p,k}}, \qquad \gamma = \sum_{(p,k)\in\mathcal{OD}}\gamma_{p,k},
$$

we conclude that

$$
\langle \bar{f}, t\rangle + \mu\psi(t/\mu) = \langle \bar{f}-\hat{f}, t\rangle + \langle \hat{f}, t\rangle + \mu\psi(t/\mu) \geq \langle \bar{f}-\hat{f}, t\rangle + \mu\gamma.
$$

Thus, the level sets of the objective function in the problem (4.3) are bounded. Therefore this problem is solvable.

In order to prove the second statement, note that by assumption $\bar{t}/\mu \in \operatorname{int}\left(\operatorname{dom}\psi_{\mathcal{R}_{p,k}}\right)$ for any $(p,k)\in\mathcal{OD}$. Therefore the gradients $\nabla\psi_{\mathcal{R}_{p,k}}(t/\mu)$ are well defined at any $t\geq\bar{t}$. Thus, this statement follows from Lemma 4.

Finally, we observe that any $t/\mu \geq \bar{t}/\mu$ belongs to $\operatorname{int}(\operatorname{dom}\psi)$. Therefore the solution t^* of the problem (4.3) satisfies the Karush-Kuhn-Tucker conditions:

$$
\begin{aligned}
\bar{f} + \nabla\psi(t^*/\mu) &= s^*, \\
(s^*)^{(\alpha)}\cdot\left((t^*)^{(\alpha)} - \bar{t}^{(\alpha)}\right) &= 0, \quad \alpha = 1,\ldots,m,
\end{aligned}
$$

where $s^*\geq 0$ is the vector of optimal dual multipliers for the constraints $t\geq\bar{t}$. Thus, if $(f^*)^{(\alpha)} < \bar{f}^{(\alpha)}$, we always have $(s^*)^{(\alpha)} > 0$ and therefore $(t^*)^{(\alpha)} = \bar{t}^{(\alpha)}$. □

5 Computational aspects

In the previous section we have seen that the equilibrium solution to the stochastic traffic assignment problem can be found from the convex minimization problem (4.3). In order to solve the latter problem numerically, we have to compute the objective function of the problem and its gradient in a reasonable computational time. Let us show how we can do so when we work with asymptotic or cumulative sets of routes.

5.1 Asymptotic sets of routes

Let us choose $\mathcal{R}_{p,k} = \widetilde{\mathcal{R}}_{p,k}$ for each origin-destination pair (p, k). In order to simplify the notation, assume that we have the full origin-destination matrix $D \in R^{n\times n}$, whose entries $D^{(i,j)}$ are the demand flows from node j to node i. Denote by $\langle \cdot, \cdot \rangle_M$ the standard scalar product in the space of $(n \times n)$-matrices. Then the minimization problem (4.3) can be written as follows:

$$\min\{\langle \bar{f}, t\rangle + \langle D, \mu \ln \Big\{(I - E(t/\mu))^{-1} - I\Big\}\rangle_M : \; t \geq \bar{t}\}, \tag{5.1}$$

where $\ln\{\cdot\}$ denotes a component-wise logarithm of an $(n \times n)$-matrix.

Let us look at the non-trivial part of the objective function in the above problem:

$$F(t) = \langle D, \ln \Big\{(I - E(t))^{-1}\Big\}\rangle_M. \tag{5.2}$$

In order to write down its derivatives we need the following notation. For two $(n \times n)$ matrices B and C we denote by $\{B/C\}$ the following matrix:

$$\{B/C\}^{(i,j)} = B^{(i,j)}/C^{(i,j)}, \quad i, j = 1, \ldots, n.$$

Let us compute the directional derivative of function $F(t)$ along a direction $h \in R^m$. For simplicity we assume that $\sigma(\mathcal{N}) < \infty$ and $\rho(t) < 1$; so the matrix $(I - E(t))^{-1}$ is well defined all its entries are positive.

$$\begin{aligned}
\mathcal{D}F(t)[h] \; &= \langle D, \mathcal{D}\left(\ln\left\{(I - E(t))^{-1}\right\}\right)[h]\rangle_M \\
&= \langle \left\{D/(I - E(t))^{-1}\right\}, \mathcal{D}\left((I - E(t))^{-1}\right)[h]\rangle_M \\
&= \langle \left\{D/(I - E(t))^{-1}\right\}, (I - E(t))^{-1}\mathcal{D}E(t)[h](I - E(t))^{-1}\rangle_M \\
&= -\langle \left\{D/(I - E(t))^{-1}\right\}, (I - E(t))^{-1}B(t)[h](I - E(t))^{-1}\rangle_M,
\end{aligned}$$

where $B(t)[h]$ is an $(n \times n)$-matrix with the entries

$$B(t)[h]^{(i,j)} = \begin{cases} \sum\limits_{\alpha\in\mathcal{A}[j,i]} e^{-t^{(\alpha)}} h^{(\alpha)}, & \mathcal{A}[j, i] \neq \emptyset, \\ 0, & \mathcal{A}[j, i] = \emptyset. \end{cases}$$

Note that the linear operator $B(t)[h]$ maps R^n into the space of $(n \times n)$-matrices $R^{n\times n}$. Denote by $B^*(t)[Y]$ its adjoint operator:

$$\langle B(t)[h], Y\rangle_M = \langle B^*(t)[Y], h\rangle \quad \forall h \in R^n, \; \forall Y \in R^{n\times n}.$$

Then

$$\begin{aligned}
\mathcal{D}F(t)[h] &= -\langle \{D/(I-E(t))^{-1}\}, (I-E(t))^{-1}B(t)[h](I-E(t))^{-1}\rangle_M \\
&= -\langle B(t)[h], (I-E(t))^{-T}\{D/(I-E(t))^{-1}\}(I-E(t))^{-T}\rangle_M \\
&= -\langle B^*(t)\left[(I-E(t))^{-T}\{D/(I-E(t))^{-1}\}(I-E(t))^{-T}\right], h\rangle.
\end{aligned}$$

Thus, we come to the following representation:

$$\nabla F(t) = -B^*(t)\left[(I-E(t))^{-T}\left\{D/(I-E(t))^{-1}\right\}(I-E(t))^{-T}\right]. \tag{5.3}$$

Taking into account the expressions (5.2), (5.3), we conclude that the computational time we need to find the function value $F(t)$ and the gradient $\nabla F(t)$ is proportional to the time, which is necessary for inverting the matrix $(I-E(t))$. In the worst case that needs $O(n^3)$ arithmetic operations. Of course, if this matrix is sparse and the demand matrix D is not filled completely, the above estimate can be significantly improved.

5.2 Cumulative sets of routes

Let show now how we can compute the potential functions for the cumulative sets of routes. That can be useful, for example, if in our model we want to bound the maximal length of the route.

Let us fix some node p. Assume we want to compute the potential functions for the cumulative sets of routes $\widehat{\mathcal{R}}^L_{p,k}$, $k=1,\dots,n$. Let us fix some $\mu>0$. Denote

$$\left.\begin{aligned}
a^l_k(t) &= \mu\psi_{R^l_{p,k}}(t/\mu) \\
b^l_k(t) &= \mu\psi_{\widehat{R}^l_{p,k}}(t/\mu)
\end{aligned}\right\}, \quad k=1,\dots,n,\ l=1,\dots,L.$$

Recall that some of these functions may be identically equal to $-\infty$ (this tells us that the corresponding set of routes is empty).

Let us show that the above functions can be computed by a simple recursion. Indeed,

$$a^1_k(t) = b^1_k(t) = \left\{\begin{array}{ll}
\mu\ln\left(\sum\limits_{\alpha\in\mathcal{A}[p,k]} e^{-t^{(\alpha)}/\mu}\right), & \text{if } \mathcal{A}[p,k]\neq\emptyset, \\
-\infty, & \text{if } \mathcal{A}[p,k]=\emptyset.
\end{array}\right., \quad k=1,\dots,n.$$

And for $l=1,\dots,L-1$ we have:

$$\left.\begin{aligned}
a^{l+1}_k(t) &= \mu\ln\left(\sum_{i\in\mathcal{I}(k)}\sum_{\alpha\in\mathcal{A}[i,k]} e^{(a^l_i(t)-t^{(\alpha)})/\mu}\right) \\
b^{l+1}_k(t) &= \mu\ln\left(e^{b^l_k(t)/\mu} + e^{a^{l+1}_k(t)/\mu}\right)
\end{aligned}\right\}, \quad k=1,\dots,n. \tag{5.4}$$

Note that each step in l of the above process takes $O(m)$ arithmetic operations. Thus, the computation of the values of all functions $b^L_k(t)$, $k=1,\dots,n$, needs $O(Lm)$ operations.

When these values are computed, we can sum up them with coefficients $d_{p,k}$. The resulting value forms a part of the objective function in (4.3). Since we have direct formulae for computing this value, its gradient in t can be also computed in $O(Lm)$ operations. It is interesting that the process (5.4), when $\mu \to 0$, becomes a well known Bellman-Ford scheme for computing the shortest path distances.

On the other hand, the relations (5.4) can be used for barrier representation of the problem (4.3) (see [7]). For that we need to employ the self-concordant barriers for entropy-type functions proposed in [4]. Then we can use the modern interior-point methods for solving (4.3).

Finally, let us mention a situation in which it is very often reasonable to use the latter technique. Assume that the graph of our network $\mathcal{N}$ contains no cycles. This means that there exists a number $\hat{n} < n$ such that $\mathcal{R}^l_{p,k} = \emptyset$ for any $l \geq \hat{n}+1$ and any p and k. In other words, $E^{\hat{n}+1}(t) \equiv 0$ and $\rho(t) \equiv 0$ for any $t \in R^m$. Consequently,

$$(I - E(t))^{-1} = \sum_{l=0}^{\hat{n}} E^l(t).$$

Is such a situation, especially when the number of origin-destination pairs is not very big, the scheme (5.4) looks much more attractive than the direct inversion of the matrix $(I - E(t))$.

6 Stochastic assignment with incomplete data

Very often it is difficult to get a reliable origin-destination matrix, which we need in the problem (4.3). Usually this data is obtained from a sophisticated analysis of the population, employment and the road structure of the region. Since this analysis must be done *before* we find the equilibrium travel time in the network, it can be based only on the free traffic shortest paths distances in the network. However, it is clear that such distance can be very different from the real distances affected by congestion. So, in the case we need to estimate (or, to compute) the origin-destination matrix of our problem, it is natural to try to do that simultaneously with the computation of the equilibrium travel time. In order to do that, we still need some aggregate information about the origin-destination flows and some model, which explains the appearance of an origin-destination pair. Let us look at some typical situation.

Assume we know nothing about the origin-destination flows in our network except one value Φ, which is the total origin-destination flow:

$$\Phi = \sum_{(p,k)\in\mathcal{OD}} d_{p,k}.$$

Let us introduce some additional notation. Denote by $\mathcal{O}$ the set of origins and by $\mathcal{D}$ the set of destinations. Then the set of our origin-destination pairs will be as follows:

$$\mathcal{OD} = \mathcal{O} \times \mathcal{D}.$$

For each $(p,k) \in \mathcal{OD}$ we introduce the set of routes $\mathcal{R}_{p,k}$. Note that the value of the function

$$\theta_{\mathcal{R}_{p,k}}(t) = -\mu\psi_{\mathcal{R}_{p,k}}(t/\mu)$$

is the *expected minimal cost* for travelling from p to k, provided that the drivers of this origin-destination pair apply the logit model (1.1) for the route choice (see [3]).

Let us introduce two weight vectors P and Q in R^m with the following entries:

$$P^{(i)} > 0 \text{ for } i \in \mathcal{O} \text{ and } P^{(i)} = 0 \text{ otherwise,}$$
$$Q^{(j)} > 0 \text{ for } j \in \mathcal{D} \text{ and } Q^{(j)} = 0 \text{ otherwise.}$$

For example, the weights P and Q can reflect the size of population at the origins and the number of jobs at destinations.

Our model for driver's behavior is of two levels. First, we introduce a stochastic demand. We assume that the probability $\pi_{i,k}$ for appearance a driver, which wishes to travel from origin i to a destination k is as follows:

$$\pi_{i,k}(t) = \frac{P_i \cdot Q_k \cdot e^{-\theta \mathcal{R}_{i,k}(t)}}{\sum\limits_{(l,j)\in\mathcal{OD}} P_l \cdot Q_j \cdot e^{-\theta \mathcal{R}_{l,j}(t)}} \tag{6.1}$$

Thus, the expected demand for this origin-destination pair is $\Phi \cdot \pi_{i,k}(t)$. Let us assume now that the drivers of this group apply the logit model (3.1) for the route choice. Then, the expected flow of drivers for the pair (i,k) is

$$f_{i,k}(t) = -\Phi \cdot \pi_{i,k}(t) \cdot \nabla \psi_{\mathcal{R}_{i,k}}(t/\mu), \tag{6.2}$$

(see Lemma 4). Consequently, the total flow in the network is

$$f(t) = \sum_{(i,k)\in\mathcal{OD}} f_{i,k}(t) = -\Phi \cdot \sum_{(i,k)\in\mathcal{OD}} \pi_{i,k}(t) \cdot \nabla \psi_{\mathcal{R}_{i,k}}(t/\mu). \tag{6.3}$$

The equilibrium problem we are going to solve now is as follows.

Find the arc travel time vector t^* and the arc flow vector $f^* = f(t^*)$, which satisfy the max-flow network performance model (4.2). (6.4)

It appears that the solution of this problem can be found from the following optimization problem:

$$\min_{t \geq \bar{t}} \; [\; \langle \bar{f}, t\rangle + \Phi \cdot \mu\psi(t/\mu) \;], \tag{6.5}$$

where $\psi(t) \equiv \psi(P,Q,t) = \ln \left(\sum\limits_{i\in\mathcal{O}} \sum\limits_{j\in\mathcal{D}} P^{(i)} Q^{(j)} g_{\mathcal{R}_{i,j}}(t) \right)$.

References

[1] T. Akamatsu. Cyclic flows, markov process and stochastic traffic assignment. *Transportation Research*, 30B, pp. 369-386, 1996.

[2] T. Akamatsu. Decomposition of path choice entropy in general transport networks. *Transportation Science*, v.31, No. 4, pp. 349-362, November 1997.

[3] S.P. Anderson, A. de Palma and J.-F. Thisse. *Discrete Choice Theory of Product Differentiation.* MIT Press, Cambridge, 1992.

[4] A. Nemirovskii. Polynomial time methods in convex programming. In: J. Renegar, M. Schub, S. Smale, Eds, *The Mathematics of Numerical Analysis*, 1995 AMS-SIAM Summer Seminar in Applied Mathematics, July 17-August 11, 1995, Park City, Utah (vol. 32 of Lectures in Applied Mathematics).

[5] Yu. Nesterov. Stable Flows in Transportation Networks. CORE DP #9907, Université Catholique de Louvain, January 1999. To appear in *Optimization and Engineering.*

[6] Yu. Nesterov and A. de Palma. Stable Dynamics in Transportation Systems. CORE DP 2000/27, Université Catholique de Louvain, April 2000.

[7] Yu. Nesterov and A. Nemirovskii. *Interior-point Polynomial Algorithms in Convex Programming.* SIAM studies in applied mathematics, vol. 13, SIAM, Philadelphia, 1994.

[8] Y. Sheffi. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods.* Prentice-Hall, Englewood Cliffs, New Jersey, 1985.

[9] J.G. Wardrop. *Some Theoretical Aspects of Road Traffic Research*, in Proceedings of the Institute of Civil Engineers, Part II, 1952, pp. 325-378.

Appendix. Properties of log-convex functions

Let $f(x)$ be a real-valued function on R^n. Define

$$\phi(x) = \ln f(x), \quad \mathrm{dom}\,\phi = \{x \in R^n : \; f(x) > 0\}.$$

The function $f(x)$ is called *log-convex* if $\mathrm{dom}\,\phi \neq \emptyset$ and $\phi(x)$ is convex. Let us present some simple properties of log-convex functions.

1. *If $f(x)$ is log-convex then it is also convex on* $\mathrm{dom}\,\phi$*.*

Indeed, let x and y belong to $\mathrm{dom}\,\phi$. Let us choose some $\alpha \in [0,1]$. Then

$$\begin{aligned} \alpha f(x) + (1-\alpha) f(y) &= \alpha e^{\phi(x)} + (1-\alpha) e^{\phi(y)} \geq e^{\alpha\phi(x) + (1-\alpha)\phi(y)} \\ &\geq e^{\phi(\alpha x + (1-\alpha) y)} = f(\alpha x + (1-\alpha) y). \end{aligned}$$

2. *If $f(x)$ is log-convex and $\beta > 0$ then $\beta f(x)$ is log-convex.*

3. *If functions $f_1(x)$, $f_2(x)$ are log-convex then $f(x) = f_1(x) + f_2(x)$ is log-convex.*

First of all note that the function of two variables $\xi(u_1,u_2) = \ln(e^{u_1} + e^{u_2})$ is convex. Indeed,

$$\nabla \xi(u_1,u_2) = \left(\frac{e^{u_1}}{e^{u_1}+e^{u_2}}, \frac{e^{u_2}}{e^{u_1}+e^{u_2}} \right)^T, \quad \nabla^2 \xi(u_1,u_2) = \frac{e^{u_1+u_2}}{(e^{u_1}+e^{u_2})^2} \cdot \left(\begin{array}{c|c} 1 & -1 \\ \hline -1 & 1 \end{array} \right) \succeq 0.$$

Therefore, for any (u_1,u_2), (v_1,v_2) and $\alpha \in [0,1]$ we have

$$\xi(\alpha u_1 + (1-\alpha) v_1, \alpha u_2 + (1-\alpha) v_2) \leq \alpha \xi(u_1,u_2) + (1-\alpha)\xi(v_1,v_2).$$

Let x and y belong to $\mathrm{dom}\,\phi$. Let us choose some $\alpha \in [0,1]$. Denote

$$\begin{aligned} u_1 &= \phi_1(x) = \ln f_1(x), \quad u_2 = \phi_2(x) = \ln f_2(x), \\ v_1 &= \phi_1(y) = \ln f_1(y), \quad v_2 = \phi_2(y) = \ln f_2(y). \end{aligned}$$

Then

$$\begin{aligned} &\ln(f_1(\alpha x + (1-\alpha) y) + f_2(\alpha x + (1-\alpha) y)) = \ln\left(e^{\phi_1(\alpha x + (1-\alpha) y)} + e^{\phi_2(\alpha x + (1-\alpha) y)} \right) \\ &\leq \ln\left(e^{\alpha\phi_1(x) + (1-\alpha)\phi_1(y)} + e^{\alpha\phi_2(x) + (1-\alpha)\phi_2(y)} \right) = \xi(\alpha u_1 + (1-\alpha) v_1, \alpha u_2 + (1-\alpha) v_2) \\ &\leq \alpha \xi(u_1,u_2) + (1-\alpha)\xi(v_1,v_2) = \alpha \ln(f_1(x) + f_2(x)) + (1-\alpha)\ln(f_1(y) + f_2(y)). \end{aligned}$$

4. *If $f(x)$ is log-convex and $p > 0$ then $f^p(x)$ is log-convex.*

5. *If $f(x)$ is concave then $1/f(x)$ is log-convex.*

6. *If $f(x)$ is convex then $e^{f(x)}$ is log-convex.*

7. *If functions $f_1(x)$, $f_2(x)$ are log-convex then $f(x) = \max\{f_1(x), f_2(x)\}$ is log-convex.*

Indeed, $\ln f(x) = \ln \max\{f_1(x), f_2(x)\} = \max\{\ln f_1(x), \ln f_2(x)\}$. Thus, this function is convex as the maximum of two convex functions.

8. *If functions $f_1(x)$, $f_2(x)$ are log-convex then $f(x) = f_1(x) \cdot f_2(x)$ is log-convex.*

9. *Let the function $f(x,y)$ be log-convex in (x,y). Define a new function*

$$F(x) = \inf\{f(x,y) : \; y \in Q\},$$

where Q is a convex set. Then $F(x)$ is a log-convex function.

Indeed, $\ln F(x) = \inf\{\ln f(x,y) : \; y \in Q\}$. Since $\ln f(x,y)$ is convex in (x,y), the function $\ln F(x)$ must be convex in x.