60

Mathematical Programming 19 (1980) 121–136.
North-Holland Publishing Company

# CHARACTERIZATIONS OF OPTIMAL SCALINGS OF MATRICES*

Uriel G. ROTHBLUM

*School of Organization and Management, Yale University, New Haven, CT 06520, U.S.A.*

Hans SCHNEIDER

*Department of Mathematics, University of Wisconsin, Madison, WI 53706, U.S.A.*

Received 22 June 1978
Revised manuscript received 4 December 1979

A scaling of a non-negative, square matrix $A \neq 0$ is a matrix of the form $DAD^{-1}$, where $D$ is a non-negative, non-singular, diagonal, square matrix. For a non-negative, rectangular matrix $B \neq 0$ we define a scaling to be a matrix $CBE^{-1}$ where $C$ and $E$ are non-negative, non-singular, diagonal, square matrices of the corresponding dimension. (For square matrices the latter definition allows more scalings.) A measure of the goodness of a scaling $X$ is the maximal ratio of non-zero elements of $X$. We characterize the minimal value of this measure over the set of all scalings of a given matrix. This is obtained in terms of cyclic products associated with a graph corresponding to the matrix. Our analysis is based on converting the scaling problem into a linear program. We then characterize the extreme points of the polytope which occurs in the linear program.

*Key words*: Scalings, Optimal Scaling, Cyclic Products, Matrices.

## 1. Introduction

Let $A \neq 0$ be an $n \times n$, non-negative matrix. A scaling of $A$ is a matrix of the form $DAD^{-1}$ where $D \in D_+^n$ and $D_+^n$ is the set of $n \times n$ non-negative, non-singular, diagonal matrices. In this paper we consider the measure of a scaling $DAD^{-1}$ given by

$$\alpha(D) \equiv \max_{A_{ij}>0,\ A_{kp}>0} \frac{D_i A_{ij} D_j^{-1}}{D_k A_{kp} D_p^{-1}}. \tag{1.1}$$

We characterize $\alpha \equiv \inf_{D \in D_+^n} \alpha(D)$ in terms of cyclic products for cycles of a directed graph associated with $A$. We then apply our results to characterize scalings of rectangular matrices. Let $0 \neq B$ be an $n_1 \times n_2$, non-negative matrix. We use the above results to characterize

* This research was partially supported by National Science Foundation grants ENG 76–15599, MCS 76–06374 and MCS 78–01087 and a grant from the Graduate School of the University of Wisconsin. This work also relates to the Department of the Navy Contract No. ONR 00014–76–C–0085 issued under ONR contract authority NR 047–006.

$$\beta \equiv \inf_{C \in D_+^m, E \in D_+^n} \max_{B_{ij}>0, B_{kp}>0} \frac{C_i B_{ij} E_j^{-1}}{C_k B_{kp} E_p^{-1}}.$$

In this case the characterization is particularly simple:

$$\beta = \max \sqrt[|\pi|]{\gamma(\pi)}$$

where the maximum is taken over all polygonal products $\pi$ is the half length of the polygon $\pi$ and $\gamma(\pi)$ is the cyclic product of the elements of the polygon (see Section 5 for precise definitions and more details). M. v. Golitscheck has pointed out to us that our characterization of $\beta$ is related to some theorems found in Diliberto and Straus [7], Aumann [1, 2], Golitscheck [10, 11] and Bank [19].

The purpose of this paper is to develop characterizations of $\alpha$ and $\beta$ for given matrices $A$ and $B$ (as above), respectively. In Section 2 we show how to compute $\alpha$ by solving a linear program. For this purpose we use a well-known minmax theorem for polyhedra due to Wolfe [18]. For the sake of completeness, we include a proof of the theorem (communicated to us privately by B.C. Eaves) in the Appendix. We show that $\log \alpha$ is the maximum of a linear function over a polytope $P$. We then give algebraic characterizations (in Section 3) and geometric characterizations (in Section 4) of the extreme points of $P$. Corresponding characterizations of optimal scalings of rectangular matrices are then developed in Section 5.

The characterization of the extreme points of the polytope $P$ is important since it enables us to characterize the basic feasible solutions of the corresponding linear program. Of course, the explicit ennumeration of the extreme points is typically not an efficient computational method. However, it might be possible to use our characterizations to develop variants of the simplex method which use the special structure of the basic feasible solutions to accelerate computation.

There are many models in which a matrix can be replaced by any one of its scalings without changing the character of the problem (e.g., linear programming). So, one would like to find a scaling of a matrix which is efficient in some way (e.g., Bauer [3, 4], Curtis and Reid [6], Fulkerson and Wolfe [8], Hamming [12], Orchard-Hays [13], Saunders and Schneider [15], and Tomlin [16]). The first to consider the ratios of the non-zero elements of a non-negative matrix and to use the maximum ratio so obtained as a measure of the goodness of the scaling of the matrix were Fulkerson and Wolfe[8]. Of course, when the matrix has negative elements one can consider absolute values (e.g., Saunders and Schneider [15] where a different measure is used). Fulkerson and Wolfe developed an algorithm for computing the quantity $\beta$ above. One motivation for using the Fulkerson–Wolfe criterion is the fact that high ratios cause difficulties in linear programming. A limitation of their measure is that it does not depend continuously on the elements of the matrix; namely, replacing a zero by a small positive number will change the measure drastically. Other measures have also been used. For example, Saunders and Schneider [15] used a maximal element measure whereas Curtis and Reid [6] used a least square approach. Tomlin [16] has compared the Fulkerson–Wolfe criterion with Curtis–Reid's in numerical experiments which favored the latter. Tomlin summarizes the issue by saying that "The scaling of linear programming problems remains a rather poorly understood subject (as indeed it does for linear equations). Although many scaling techniques have been proposed, the rationale behind them is not always evident and very few numerical results are available."

Fulkerson and Wolfe [8], Curtis and Reid [6] and Tomlin [15] contributed algorithms and numerical results. Like Bauer [3, 4] and Saunders and Schneider [15], our main aim is to develop theoretical analysis of a specific scaling criterion, though we do have some computational results (in Section 2). Our paper is certainly not an endorsement of the Fulkerson–Wolfe method over others. It is offered in the hope that our theoretical analysis will eventually lead to a better understanding of the relationship between the theoretical properties and the numerical efficiency of scaling criteria.

## 2. Conversion of the multiplicative scaling problem into a linear program

In this section we convert the multiplicative scaling problem into an additive one. We then convert the additive problem into a linear program.

We shall find it convenient to look at the graph associated with a given matrix. We first need some definitions. A (directed) *graph* $G$ is an ordered pair $(G_v, G_a)$ where $G_a \subseteq G_v \times G_v$. Elements of $G_v$ are called *vertices* and elements of $G_a$ are called *arcs*. In this paper we shall typically have $G_v \subseteq \{1, \dots, n\} \equiv \langle n \rangle$ for some positive integer $n$. In this case we order $G_a$ lexicographically, viz., $G_a = (a_1, \dots, a_m)$ where $m$ denotes the number of elements of $G_a$. If $G$ is a graph, the vertex–arc incidence matrix, denoted $\Gamma(G)$, is the $n \times m$ matrix defined by

$$\Gamma(G)_{iq} = \begin{cases} 1 & \text{if } a_q = (i,j) \text{ for some } i \neq j \in G_v, \\ -1 & \text{if } a_q = (j,i) \text{ for some } j \neq i \in G_v, \\ 0 & \text{otherwise.} \end{cases}$$

(Observe that $\Gamma_{iq} = 0$ if $a_q = (i,i)$.)

Let $R$ be the real field and let $R_+$ be the set of non-negative reals. By $R^{nm}$ (resp., $R_+^{nm}$) we denote the set of all $n \times m$ matrices with elements in $R$ (resp., $R_+$). As usual, $R^n$ (resp., $R_+^n$) will stand for $R^{n1}$ (resp., $R_+^{n1}$). Throughout, we use subscripts for coordinates. Next, let $D_+^n$ denote the set of non-singular diagonal matrices in $R_+^{nn}$. For $D \in D_+^n$ let $D_i \equiv D_{ii}$, $i = 1, \dots, n$.

Let $A \in R^{nn}$. The (directed) *graph associated* with $A$, written $G(A)$, has $G(A)_v = \langle n \rangle$ and $G(A)_a = \{(i,j) \in \langle n \rangle \times \langle n \rangle \mid A_{ij} \neq 0\}$. Also, the incidence matrix associated with $A$, written $\Gamma(A)$, is the matrix $\Gamma[G(A)]$.

Throughout this paper, let $0 \neq A \in R_+^{nn}$ be a fixed matrix. Let $G \equiv G(A)$ and $\Gamma = \Gamma(A)$. For $D \in D_+^n$ we consider

$$\alpha(D) \equiv \max_{A_{ij}>0,\, A_{kp}>0} \frac{D_i A_{ij} D_j^{-1}}{D_k A_{kp} D_p^{-1}}.$$

Define $d \in R^n$ and $w \in R^m$ by

$$d_i = \log D_i, \quad i = 1, \dots, n,$$

$$w_q = \log A_{ij}, \quad \text{if } a_q = (i,j) \in G_a.$$

Then, if $(i,j) = a_q$

$$\log D_i A_{ij} D_j^{-1} = (w^{\mathrm{T}} + d^{\mathrm{T}} \Gamma)_q = (w^{\mathrm{T}} + d^{\mathrm{T}} \Gamma) e^q$$

where $e^q$ is the $q$th unit vector in $R^m$, i.e.,

$$e_r^q = \begin{cases} 1 & \text{if } r = q, \\ 0 & \text{if } r \neq q \text{ and } r = 1, \dots, m. \end{cases}$$

So,

$$\log \alpha(D) = \max_{q,r=1,\dots,m} (w^{\mathrm{T}} + d^{\mathrm{T}} \Gamma)(e^q - e^r)$$

$$= \max_{x \in F} (w^{\mathrm{T}} + d^{\mathrm{T}} \Gamma) x = \max_{x \in Q} (w^{\mathrm{T}} + d^{\mathrm{T}} \Gamma) x$$

where

$$F = \{e^q - e^r \mid q,r = 1, \dots, m\}$$

and

$$Q = \text{convex hull of } F.$$

Notice that $O \in F$, so

$$\log \alpha(D) \geq 0 \quad \text{for all } D \in D_+^n. \tag{2.2}$$

We next obtain an explicit representation of $Q$. We need two additional definitions. First, let $e \in R^m$ be the vector all of whose coordinates are one. Also, for $x \in R^m$, let $\|x\| = \sum_{i=1}^m |x_i|$, i.e., $\| \,\|$ denotes the $l_1$, norm in $R^m$.

**Theorem 1.**

$$Q = \{x \in R^m \mid e^{\mathrm{T}} x = 0 \text{ and } \|x\| \leq 2\}. \tag{2.3}$$

**Proof.** Let $Q'$ be the right-hand side of (2.3). It is easily seen that $Q'$ is convex and $e^{qr} \equiv e^q - e^r \in Q'$ for all $q,r = 1, \dots, m$. This assures that $Q \subseteq Q'$.

We next show that $Q' \subseteq Q$. Let $x \in Q'$. We prove that $x \in Q$ by induction on the number of non-zero coordinates of $x$, which we denote $\lambda(x)$. If $\lambda(x) = 0$, then $x = 0 \in Q$. If $\lambda(x) = 2$, then $x$ is proportional to some $e^{qr}$, $q,r = 1, \dots, m$ with $q \neq r$, and

$$x = (0.5 + 0.25\|x\|)(e^q - e^r) + (0.5 - 0.25\|x\|)(e^r - e^q) \in Q.$$

Of course, $\lambda(x)=1$ is impossible.

Assume that for some $t \geq 3$ every $z \in Q'$ with $\lambda(z) < t$ belongs to $Q$ and consider $x \in Q'$ having $\lambda(x) = t$. Let $\tau = \min_{x_q \neq 0}|x_q| = |x_{q_0}|$. Since $\lambda(x) = t \geq 3$, $\|x\| > 2\tau$. Also, since $e^{\mathrm{T}}x = 0$, there exists $r_0$, $1 \leq r_0 \leq m$, such that $x_{q_0}x_{r_0} < 0$. Let $y \equiv x - x_{q_0}e^{q_0r_0}$. Notice that $e^{\mathrm{T}}y = 0$ and $\|y\| = \|x\| - 2\tau > 0$. So, $z \equiv 2y/\|y\| \in Q'$. Also observe that $\lambda(z) < \lambda(x) = t$. So, by our induction hypothesis $z \in Q$. It is easily seen that $x = 0.5\|y\|z + (1 - 0.5\|y\|)u$ where $u \equiv x_{q_0}e^{q_0r_0}/(1-0.5\|y\|)$. Since $0 \leq \|y\| \leq \|x\| \leq 2$ and $Q$ is convex, it suffices to prove that $u \in Q$. But, $e^{\mathrm{T}}u = 0$ and (as $\|x\| \leq 2$)

$$\|u\| = \frac{2|x_{q_0}|}{1-0.5\|y\|} = \frac{2\tau}{1+\tau-0.5\|x\|} \leq 2.$$

So $u \in Q'$. But, as $\lambda(u) = 2$, this assures $u \in Q$, completing our proof.

We next examine the quantity of chief interest in this paper:

$$\alpha \equiv \inf_{D \in D_+^n} \alpha(D) = \inf_{D \in D_+^n} \max_{A_{ij}>0, A_{kp}>0} \frac{D_iA_{ij}D_j^{-1}}{D_kA_{kp}D_p^{-1}} \tag{2.4}$$

It follows from (2.1) and (2.2) that

$$\log \alpha = \inf_{d \in R^n} \max_{x \in Q} (w^{\mathrm{T}} + d^{\mathrm{T}}\Gamma)x \geq 0.$$

For each $d \in R^n$, $\max_{x \in Q}(w^{\mathrm{T}} + d^{\mathrm{T}}\Gamma)x < \infty$. So the above infmax is finite and we can apply the minmax theorem of Wolfe [18] (found in the Appendix). Combining this theorem with the fact that if $\Gamma x \neq 0$, then $\inf_{d \in R^n} w^{\mathrm{T}}\Gamma x = -\infty$ to get that

$$\log \alpha = \sup_{x \in Q} \inf_{d \in R^n} (w^{\mathrm{T}} + d^{\mathrm{T}}\Gamma)x = \sup_{x \in P} w^{\mathrm{T}}x$$

where

$$P \equiv \{x \in Q \mid \Gamma x = 0\}. \tag{2.5}$$

Notice that $0 \in P$, so $P \neq \emptyset$. Also note that $P$ is a bounded polytope as it is the intersection of the bounded polytope $Q$ and a subspace. In particular it follows that $\max_{x \in P} w^{\mathrm{T}}x = \max_{u \in U} w^{\mathrm{T}}u$, where $U$ is the set of extreme points of $P$. We now collect the results of this section into a theorem.

**Theorem 2.** *Let $0 \neq A \in R_+^{nn}$ and let $\Gamma = \Gamma[G(A)]$. Let $P = \{x \in R^m \mid e^{\mathrm{T}}x = 0, \|x\| \leq 2, \Gamma x = 0\}$. Then*

$$\log \alpha = \inf_{D \in D_+^n} \max_{A_{ij}>0, A_{kp}>0} \left(\log \frac{D_iA_{ij}D_j^{-1}}{D_kA_{kp}D_p^{-1}}\right) = \sup_{x \in P} w^{\mathrm{T}}x = \sup_{x \in U} w^{\mathrm{T}}x \tag{2.6}$$

*where $w \in R^m$ is the vector whose coordinates are given by $\log A_{ij}$ for $(i,j) \in G(A)$, taken in lexicographical order and $U$ is the set of extreme points of $P$.*

## 3. An algebraic characterization of the extreme points of $P$

The purpose of this section is to obtain an algebraic characterization of the extreme points of the polytope $P$ given by (2.5).

We first need some definitions. We say that a graph $G' = (G'_v, G'_a)$ is a *subgraph* of $G = G(A) = (G_v, G_a)$, if $G'_v \subseteq G_v$ and $G'_a \subseteq G_a$. The *subgraph* of $G$ *associated* with a vector $x \in R^m$, written $G(x)$, is the subgraph of $G$ having

$$G(x)_a = \{a_q \in G_a \mid x_q \neq 0\}$$

and

$$G(x)_v = \{i \in G_v \mid (i,j) \in G(x)_a \text{ or } (j,i) \in G(x)_a \text{ for some } j \in G_v\}.$$

We shall call $x \in R^m$ a *circulation for the subgraph* $G'$ or $G$, if $\Gamma x = 0$ and $G(x)_a \subseteq G'_a$. The term *circulation* (when no subgraph is mentioned) will refer to a circulation for $G$ itself, i.e., $x$ is a circulation if $\Gamma x = 0$. The set of all circulations for a given subgraph $G'$ of $G$ is clearly a subspace of $R^m$. This subspace will be called the *circulation space* of $G'$ and will be denoted $C(G')$. We shall also use the abbreviated notation $C(x)$ for $C[G(x)]$ where $x \in R^m$. The dimension of these subspaces will be denoted $\dim C(G')$ and $\dim C(x)$, respectively.

We say that two vectors $x$ and $y$ in $R^m$ *conform* if for every $q = 1, \dots, m$, $x_q y_q \geq 0$. It is easily seen that $x$ and $y$ conform if and only if $\|x + y\| = \|x\| + \|y\|$.

A *cycle of* a subgraph $G'$ of $G$ is a vector $z \in R^m$ with $G(z)_a \subseteq G'$ such that

$$z \text{ is a non-zero circulation,} \tag{3.1}$$

$$z_q \in \{1, -1, 0\} \quad \text{for all } q = 1, \dots, m \tag{3.2}$$

and

$$\text{if } y \text{ is a circulation and } G(y)_a \subset G(z)_a, \text{ then } y = 0. \tag{3.3}$$

The term "*cycle*" (when no subgraph of $G$ is mentioned) will refer to a cycle of $G$ itself. Of course, if $z$ is a cycle, so is $-z$. Also, it is well-known that the circulation space of any subgraph $G'$ of $G$ is spanned by the set of cycles of $G'$ (cf. [5, p. 90]). Observe that for a circulation $u$,

$$\dim C(u) = 1 \quad \text{if and only if } u \text{ is a non-zero multiple of a cycle.} \tag{3.4}$$

Also, for two cycles $z^1$ and $z^2$

$$G(z^1) \neq G(z^2) \quad \text{if and only if } z^1 \text{ and } z^2 \text{ are linearly independent.} \tag{3.5}$$

We are now ready to begin our consideration of the polytope

$$P = \{x \in R^m \mid \Gamma x = 0,\ e^{\mathrm{T}} x = 0 \text{ and } \|x\| \leq 2\}.$$

**Theorem 3.** *The polytope $P \neq \{0\}$ if and only if either there exists a cycle $z$ for $G(A)$ having $e^{\mathrm{T}}z = 0$ or there exists two linearly independent cycles for $G(A)$.*

**Proof.** Suppose $P \neq \{0\}$. Then there exists a circulation $x \neq 0$ with $e^{\mathrm{T}}x = 0$. Since the circulation space is spanned by the set of cycles it follows that for some $k \geq 1$ there exist independent cycles $z^1, \ldots, z^k$ and non-zero real numbers $\xi_1, \ldots, \xi_k$ such that $x = \sum_{i=1}^{k} \xi_i z^i$. If $k = 1$, $e^{\mathrm{T}}z^1 = 0$ and if $k \geq 2$, then $z^1$ and $z^2$ are two independent cycles. This proves one direction of our theorem.

Next assume that for some cycle $z$, $e^{\mathrm{T}}z = 0$. Then $0 \neq 2z/\|z\| \in P$. Finally let $z^1$ and $z^2$ be two linearly independent cycles having $e^{\mathrm{T}}z^i \neq 0$ for $i = 1,2$. Then for some real numbers $\beta_1, \beta_2$, $x \equiv \beta_1 z^1 + \beta_2 z^2$ has $e^{\mathrm{T}}x = 0$, $\|x\| = 2$. Hence $0 \neq x \in P$, which proves that $P \neq \{0\}$ and completes our proof.

From now on we shall always assume that $P \neq \{0\}$. We shall identify the extreme points of $P$ by proving a sequence of lemmas.

**Lemma 1.** *Suppose $P \neq \{0\}$. Then $u \in P$ is an extreme point of $P$ if and only if $\|u\| = 2$ and there exist no linearly independent conforming vectors $x^1, x^2 \in R^m$, having $e^{\mathrm{T}}x^i = 0$, $\Gamma x^i = 0$ for $i = 1,2$ and $u = x^1 + x^2$.*

**Proof.** Suppose $u$ is an extreme point of $P$. We first show that $u \neq 0$. Since $P \neq \{0\}$, there exists $0 \neq x \in P$. Notice that $-x \in P$ and $0 = 0.5x + 0.5(-x)$. So, 0 is not an extreme point of $P$. So, $u \neq 0$. We next show that $\|u\| = 2$. Suppose $\|u\| < 2$, then for sufficiently small $\epsilon > 0$, $(1+\epsilon)u \in P$ and $(1-\epsilon)u \in P$. Since $u = 0.5(1+\epsilon)u + 0.5(1-\epsilon)u$ it follows that $u$ is not an extreme point. So, $\|u\| = 2$. Next, let $x^1, x^2$ be linearly independent conforming vectors in $R^m$ having $\Gamma x^i = 0$, $e^{\mathrm{T}}x^i = 0$ for $i = 1,2$, and $u = x^1 + x^2$. Then $\|x^1\| + \|x^2\| = \|x^1 + x^2\| = \|u\| = 2$ and $u = \frac{1}{2}\|x^1\|u^1 + \frac{1}{2}\|x^2\|u^2$ where $u^i = 2x^i/\|x^i\| \in P$ implying that $u$ is not an extreme point of $P$. We have thereby shown that any extreme point of $P$ satisfies the conditions of the lemma.

Next suppose $u$ satisfies the conditions of the lemma; in particular $\|u\| = 2$. If $u$ is not an extreme point of $P$, then some $u^1, u^2 \in P$, $u^1 \neq u^2$ and positive numbers $t_1, t_2$ with $t_1 + t_2 = 1$ such that $u = t_1 u^1 + t_2 u^2$. Thus,

$$2 = \|u\| \leq t_1\|u^1\| + t_2\|u^2\| \leq 2.$$

We deduce that $\|u^i\| = 2, i = 1,2,$ and that for $x^i = t_i u^i$, $i = 1,2$, $\|x^1 + x^2\| = \|x^1\| + \|x^2\|$. Thus $x^1$ and $x^2$ conform. Since $\|u^1\| = \|u^2\| = 2$, $u^1 \neq u^2$ and $u^1$ and $u^2$ conform, we get that $x^1, x^2$ are linearly independent. Also, $e^{\mathrm{T}}x^i = 0$ and $\Gamma x^i = 0$ for $i = 1,2$. So $u$ does not satisfy the conditions of the lemma.

**Lemma 2.** *Suppose $u \in P$, $\|u\| = 2$ and* $\dim C(u) = 1$. *Then $u$ is an extreme point of $P$.*

**Proof.** Since $\dim C(u) = 1$, by (3.4) there is a cycle $z$ such that all elements in $C(u)$ are scalar multiples of $z$. Let $u = x^1 + x^2$, where $x^1, x^2$ conform and for $i = 1,2$, $\Gamma x^i = 0$ and $e^T x^i = 0$. Then $G(x^i)_a \subseteq G(u)_a$ whence $x^i \in C(u)$, $i = 1,2$. So $x^i$, $i = 1,2$, is a scalar multiple of $z$. Thus, $x^1$ and $x^2$ are linearly dependent. The conclusion of this lemma now follows from Lemma 1.

**Lemma 3.** *Suppose $u \in P$, $\|u\| = 2$ and* $\dim C(u) = 2$. *If $e^T z \neq 0$, for some cycle $z$ in $C(u)$, then $u$ is an extreme point of $P$.*

**Proof.** Let $u = x^1 + x^2$, where $x^1, x^2$ conform and for $i = 1,2$, $\Gamma x^i = 0$ and $e^T x^i = 0$. Then $G(x^i)_a \subseteq G(u)_a$ whence $x^i \in C(u)$. Suppose $x^1 \neq 0$. Since $e^T z \neq 0 = e^T x^1$ and $x^1 \neq 0$ the vectors $x^1$ and $\dot{z}$ are linearly independent. Since $x^2 \in C(u)$ and $\dim C(u) = 2$, there exist $\beta, \gamma \in R$ for which $x^2 = \beta x^1 + \gamma z$. But as $e^T x^2 = 0$ we have that $\gamma = 0$. Hence $x^1$ and $x^2$ are linearly dependent. This conclusion is immediate when $x^1 = 0$. It now follows from Lemma 1 that $u$ is an extreme point of $R$.

**Lemma 4.** *Suppose $u \in P$, $\|u\| = 2$ and* $\dim C(u) = 2$. *If every cycle $z$ in $C(u)$ satisfies $e^T z = 0$, then $u$ is not an extreme point of $P$.*

**Proof.** Since the cycles in $C(u)$ span $C(u)$, there exists a cycle $z \in C(u)$ such that $z$ and $u$ are linearly independent. For $\epsilon > 0$ which is sufficiently small, $u^1 \equiv u + \epsilon z$ and $u^2 \equiv u - \epsilon z$ conform and are linearly independent. Since $u = 0.5(u^1 + u^2)$, $\Gamma u^i = 0$ and $e^T u^i = 0$, $i = 1,2$, the conclusion follows from Lemma 1.

**Lemma 5.** *Suppose $u \in P$, $\|u\| = 2$ and* $\dim C(u) > 2$. *Then $u$ is not an extreme point of $P$.*

**Proof.** It follows from the assumptions that there exist $x, y \in C(u)$ such that $u, x$ and $y$ are linearly independent. There exist $\xi, \eta \in R$ for which $v = \xi x + \eta y$ has $e^T v = 0$. For sufficiently small $\epsilon > 0$, $u^1 \equiv u + \epsilon v$ and $u^2 \equiv u - \epsilon v$ conform. Clearly $u$ and $v$ are linearly independent and hence, so are $u^1$ and $u^2$. Of course, for $i = 1,2$, $\Gamma u^i = 0$ and $e^T u^i = 0$. But, $u = 0.5(u^1 + u^2)$. It follows from Lemma 1 that $u$ is not an extreme point of $P$.

We next collect the results of Lemmas 2–5 to obtain an algebraic characterization of the extreme points of $P$.

**Theorem 4.** *Suppose $P \neq \{0\}$. Then $u$ is an extreme point of $P$ if and only if $u \in P$, $\|u\| = 2$ and either*

$$\dim C(u) = 1 \tag{3.6}$$

*or*

$$\dim C(u) = 2 \text{ and some cycle } z \in C(u) \text{ satisfies } e^{\mathrm{T}} z \neq 0. \tag{3.7}$$

## 4. A geometric characterization of the extreme points of $P$

The purpose of this section is to obtain a geometric characterization of the extreme points of $P$. We first need some additional definitions.

A *path from $i$ to $j$*, where $i,j = 1, \dots, n$, is a vector $z \in R^m$ such that

$$z_q \in \{1, -1, 0\} \quad \text{for all } q = 1, \dots, m \tag{4.1}$$

and

$$\Gamma z = e^i - e^j, \tag{4.2}$$

where here, $e^i$ and $e^j$ denote the corresponding unit vectors in $R^n$. (Notice that in Section 2 we considered the unit vectors in $R^m$.) A given path $z$ does not necessarily determine uniquely the points $i$ and $j$, e.g., the path $z = 0$ is a path from $i$ to $i$ for every $i = 1, \dots, n$.

We next give the well-known geometric interpretation of paths and cycles (e.g. [5, p. 8]). Namely, a vector $z \in R^m$ is a path from $i$ to $j$ if and only if there exists a positive integer $k$ and $i_1, \dots, i_k$ in $G(z)_v$ such that $i_1 = i$, $i_k = j$ and for every $r = 1, \dots, k-1$ precisely one element of $(i_r, i_{r+1})$ and $(i_{r+1}, i_r)$ is in $G(z)_a$ and each element of $G(z)_a$ has this form: if $a_q = (i_r, i_{r+1})$ we have $z_q = 1$ and if $a_q = (i_{r+1}, i_r)$ we have $z_q = -1$. Of course, a vector $z \in R^m$ is a *cycle* if in addition $i_1, \dots, i_{k-1}$ are distinct and $i_1 = i_k$.

By considering subgraphs of subgraphs of $G$, it is easily seen that the set of subgraphs of $G$ is partially ordered. It follows that this partially ordered set is a lattice, where for subgraphs $G^1$ and $G^2$ of $G$, *we have* $G^1 \vee G^2 = (G_v^1 \cup G_v^2, G_a^1 \cup G_a^2)$ and $G^1 \wedge G^2 = (G_v^1 \cap G_v^2, G_a^1 \cap G_a^2)$.

Given a subgraph $G'$ of $G$, we say that $i$ and $j$ in $G'_v$ are *connected* if there exists a path $y$ from $i$ to $j$ having $G(y)_a \subseteq G'_a$. Of course, the connectedness relation is an equivalence relation. Thus for every subgraph $G'$ of $G$ we get a partition of $G'_v$ into equivalence classes. The number of components in this partition will be denoted $p(G')$. Using the geometric interpretation of paths one can see that for every subgraph $G'$ of $G$ with $p(G') = p$, there exist subgraphs $G^1, \dots, G^p$ of $G$ such that

$$G = \bigvee_{i=1}^{p} G^i \quad \text{and} \quad G^i \wedge G^j = (\emptyset, \emptyset) \quad \text{for } i \neq j,\ i,j = 1, \dots, p$$

where $\{G_v^i \mid i = 1, \dots, p\}$ are the equivalence classes of the connectedness relation on $G'_v$. A subgraph $G'$ of $G$ is called *connected* if every $i,j \in G'_v$ are connected, or equivalently, if either $p(G') = 1$ or $G' = (\emptyset, \emptyset)$.

Let $|H|$ denote the number of elements in the set $H$. No confusion should arise with the fact that this notation is also used to denote absolute values of real numbers. Recall the well-known formula (e.g. [5, p. 16]) which states that for every subgraph $G'$ of $G$

$$\dim C(G') = |G'_a| - |G'_v| + p(G'). \tag{4.3}$$

**Lemma 6.** *Let $z^1$ and $z^2$ be linearly independent cycles for which $G(z^1)_v \cap G(z^2)_v \neq \emptyset$. Then*

$$\dim C[G(z^1) \vee G(z^2)] = p[G(z^1) \wedge G(z^2)] + 1. \tag{4.4}$$

**Proof.** Let $G^1, \ldots, G^p$ be the connected components of $G(z^1) \wedge G(z^2)$. Intuitively, it follows from the geometric interpretation of paths and cycles that both $z^1$ and $z^2$ may be decomposed into the above connected components $G^1, \ldots, G^p$ (of $G(z^1) \wedge G(z^2)$) together with $p$ "disjoint paths" each of which has no arcs in common with $\vee_{t=1}^p G^t$ and only the two "end vertices" are in $\bigcup_{t=1}^p G^t_v$. More precisely, there exist paths $y^{it}$, $i = 1,2$, $t = 1, \ldots, p$ such that for $i = 1,2$

$$G(z^i) = \bigvee_{t=1}^{p} [G^t \vee G(y^{it})]$$

and for $t = 1, \ldots, p$

$$G(y^{it}) \wedge \left( \bigvee_{t=1}^{p} G^t \right) = (H^{it}, \emptyset)$$

where $|H^{it}| = 2$. (Notice that some $G^t_a$ may be empty.) So,

$$[G(z^1) \vee G(z^2)]_a = \sum_{t=1}^{p} (|G^t_a| + |G(y^{1t})_a| + |G(y^{2t})_a|). \tag{4.5}$$

Also notice that

$$\begin{aligned} |[G(z^1) \vee G(z^2)]_v| &= |G(z^1)_v| + |G(z^2)_v| - |[G(z^1) \wedge G(z^2)_v| \\ &= \sum_{i=1}^{2} \sum_{t=1}^{p} (|G^t_v| + |G(y^{it})_v| - 2) - \sum_{t=1}^{p} |G^t_v| \\ &= \sum_{t=1}^{p} (|G^t_v| + |G(y^{1t})_v| + |G(y^{2t})_v|) - 4p. \end{aligned} \tag{4.6}$$

It is easily seen that for $t = 1, \ldots, p$ and $i = 1,2$

$$|G(y^{it})_a| - |G(y^{it})_v| = -1 \tag{4.7}$$

and

$$|G^t_a| - |G^t_v| = -1. \tag{4.8}$$

Also, as $G(z^1)_v \wedge G(z^2)_v \neq \emptyset$ and $z^1$ and $z^2$ are cycles we get that

$$p[G(z^1) \vee G(z^2)] = 1. \tag{4.9}$$

It follows from (4.3) with $G' = G(z^1) \vee G(z^2)$ and (4.5)–(4.9) that

$$\dim C[G(z^1) \vee G(z^2)] = \sum_{i=1}^{p}\left[|G_a^i| - |G_v^i| + \sum_{j=1}^{2}(|G_a^{ij}| - |G_v^{ij}|)\right] + 4p + 1$$
$$= -3p + 4p + 1 = p + 1.$$

**Lemma 7.** *Let $z^1$ and $z^2$ be cycles such that $G(z^1) \neq G(z^2)$. Then* $\dim C[G(z^1) \vee G(z^2)] = 2$ *if and only if $G(z^1) \wedge G(z^2)$ is connected.*

**Proof.** First recall that (3.5) implies that $z^1$ and $z^2$ are linearly independent. Using the geometric interpretation of paths and cycles, one can verify that if $G(z^1)_v \cap G(z^2)_v = \emptyset$, then $G(z^1)$ and $G(z^2)$ are the only two connected components of $G(z^1) \vee G(z^2)$. In this case $G(z^1) \wedge G(z^2) = (\emptyset, \emptyset)$ is connected and (4.3) with $G' = G(z^1) \vee G(z^2)$ combined with the fact that $|G(z)_a| = |G(z)_v|$ for every cycle $z$ imply that

$$\dim C[G(z^1) \vee G(z^2)] = |G(z^1)_a| + |G(z^2)_a| - |G(z^1)_v| - |G(z^2)_v| + 2 = 2.$$

If $G(z^1)_v \cap G(z^2)_v \neq \emptyset$, then $G(z^1) \wedge G(z^2)$ is connected if and only if $p[G(z^1) \wedge G(z^2)] = 1$ and the result of the corollary follows directly from Lemma 6.

By a result of Tutte [17, p. 23] every circulation $u$ is a linear combination with positive coefficients of cycles which conform with $u$. We next use this result to prove that $C(u)$ has a basis of cycles which conform with $u$. This result might be known but we prove it here for completeness.

**Lemma 8.** *Let $u$ be a circulation. Then $C(u)$ has a basis of cycles which conform with $u$.*

**Proof.** It is enough to show that $C(u)$ is contained in the span of the set of cycles which conform with $u$, which we denote $C^*(u)$. Let $x \in C(u)$. Then $u$ and $u + \epsilon x$ conform for all sufficiently small $\epsilon > 0$. Hence, by Tutte's result both $u$ and $u + \epsilon x$ are in $C^*(u)$. It immediately follows that $x \in C^*(u)$.

We are now ready to obtain the geometric characterization of the extreme points of $P$.

**Theorem 5.** *Suppose $P \neq \{0\}$. Then $u$ is an extreme point of $P$ if and only if $u \in P$, $\|u\| = 2$ and either*

$$u = \alpha x \quad \text{for some } \alpha > 0 \text{ and cycle } x \text{ having } e^{\mathrm{T}}x = 0, \tag{4.10}$$

*or*

$$u = \xi_1 z^1 + \xi_2 z^2 \quad \text{for some } \xi_1 > 0,\ \xi_2 > 0 \text{ and conforming cycles } z^1 \text{ and } z^2 \text{ for which } G(z^1) \wedge G(z^2) \text{ is connected and } e^{\mathrm{T}}z^1 > 0 > e^{\mathrm{T}}z^2. \tag{4.11}$$

**Proof.** Suppose $u \in P, \|u\| = 2$ and $u$ satisfies (4.10). Then by (3.4), $\dim C(u) = 1$, whence $u$ is an extreme point by Theorem 4. Suppose $u \in P, \|u\| = 2$ and $u$ satisfies (4.11)). Then $z^1, z^2 \in C(u)$ and $G(z^1) \neq G(z^2)$. So Lemma 7 assures that $\dim C(u) = \dim[G(z^1) \vee G(z^2)] = 2$. Since $e^T z^1 \neq 0$ Theorem 4 implies that $u$ is an extreme point of $P$.

Conversely, let $u$ be an extreme point of $P$. By Theorem 4, $u \in P$, $\|u\| = 2$ and either (3.6) or (3.7) hold. If (3.6) holds, i.e., $\dim C(u) = 1$, then (4.10) follows directly from (3.4). Alternatively, assume that (3.7) holds, i.e., $\dim C(u) = 2$ and for some cycle $z \in C(u)$, $e^T z \neq 0$. Lemma 8 implies that there exist two cycles $z^1$ and $z^2$ both conforming with $u$ such that $z^1$ and $z^2$ form a basis of $C(u)$. Let $u = \xi_1 z^1 + \xi_2 z^2$. Since $u$ is not a scalar multiple of a cycle (in which case (3.4) implies that $\dim C(u) = 1$), $\xi^1 \neq 0$ and $\xi^2 \neq 0$. Since $z^1$ and $z^2$ are independent $G(z^1)_a \setminus G(z^2)_a \neq \emptyset$ and $G(z^2)_a \setminus G(z^1)_a \neq \emptyset$. The above and the fact that $u$ conforms with both $z^1$ and $z^2$ imply that $\xi^1 > 0$ and $\xi^2 > 0$. Since $e^T z \neq 0$ for some $z \in C(u)$ either $e^T z^1 \neq 0$ or $e^T z^2 \neq 0$. But $0 = e^T u = \xi_1 e^T z^1 + \xi_1 e^T z^2$ and $\xi^i > 0, i = 1{,}2$. *Hence* $e^T z^1 \neq 0, e^T z^2 \neq 0$ and $e^T z^1$ and $e^T z^2$ have opposite signs. Also, $G(z^1) \neq G(z^2)$ (as $z^1$ and $z^2$ are linearly independent) and $\dim C[G(z^1) \vee G(z^2)] = \dim C(u) = 2$. So Lemma 7 assures that $G(z^1) \wedge G(z^2)$ is connected, completing our proof that $u$ satisfies (4.11).

We shall next combine Theorem 5 with eq. (2.6) to obtain an explicit characterization of $\alpha$. We first need some additional definitions.

For $a_q = (i,j)$ let $A_q = A_{ij}$. The *cyclic product* of a cycle $z \in C(G)$, denoted $\gamma(z)$, is defined by

$$\gamma(z) = \prod_{q \in G_a} A_q^{z_q}.$$

Of course, the cyclic product of $z$ is simply the product of all $A_q$ for arcs $a_q$ with $z_q = 1$, divided by the product of all $A_q$ for arcs $q$ with $z_q = -1$. Also, let

$$C^0(G) = \{z \in C(G) \mid e^T z = 0\}$$

and

$$C^2(G) = \{(z^1, z^2) \mid z^1, z^2 \in C(G), e^T z^1 > 0 > e^T z^2 \text{ where } z^1 \text{ and } z^2 \text{ conform}\}.$$

It is easy to verify that for $(z^1, z^2) \in C^2(G)$ there exist unique real numbers $\xi_1, \xi_2$ satisfying

$$\xi_1 e^T z^1 + \xi_2{}^T z^2 = 0,$$
$$\xi_1 \|z^1\| + \|\xi_2\| z^2 = 2,$$
$$\xi_1, \xi_2 > 0.$$

The solution of the above system (for fixed $(z^1, z^2) \in C^2(G)$) will be denoted $\xi_1(z^1, z^2)$, $\xi_2(z^1, z^2)$. We are now ready for the explicit characterization of $\alpha$.

**Theorem 6.** *Let $0 \neq A \in R^{nn}$ and let $\alpha$ be given by* (2.4). *Then $\alpha = 1$ if $G^0(G) = C^2(G) = \emptyset$. Otherwise,*

$$\alpha = \max\left\{ \max_{z \in C(G)} \sqrt[\|z\|/2]{\gamma(z)},\ \max_{(z^1, z^2) \in C^2(G)} \gamma(z_1)^{\xi_1(z^1, z^2)} \gamma(z_2)^{\xi_2(z^1, z^2)} \right\}.$$

**Proof.** Taking exponentials, it follows from (2.6) that

$$\alpha = \max_{u \in U} e^{w^{\mathrm{T}} u},$$

where $w_q = A_q$ for $q = 1, \ldots, m$ and $U$ is the set of extreme points of $P$. Our conclusion now follows directly from Theorem 5.

## 5. Rectangular scaling

Let $0 \neq B \in R_+^{n_1 n_2}$ and consider the scaling problem of computing

$$\beta \equiv \inf_{C \in D_+^{n_1}, E \in D_+^{n_2}} \max_{B_{ij} > 0, B_{kp} > 0} \frac{C_i B_{ij} E_j^{-1}}{C_k B_{kp} E_p^{-1}}. \tag{5.1}$$

One can easily see that $\beta$ equals the right-hand side of (2.4), with

$$A \equiv \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \in R^{n \times n}, \tag{5.2}$$

where $n = n_1 + n_2$, and the zero matrices are of the appropriate size. Thus the rectangular scaling problem reduces to the previously discussed problem. It follows that our results apply; in particular, Theorems 2–6 hold. We next use the special structure of $A$ given in (5.2) to further strengthen Theorems 3–6. The graph $G = G(A)$ is easily defined in terms of $B$:

$$(i,j) \in G_a \quad \text{if and only if } i \leq n_1 < j \text{ and } B_{i,j-n_1} > 0,$$

and hence $G$ is bipartite. (For further details, see Saunders and Schneider [14, 15].) It follows easily that every circulation $x$ has $e^{\mathrm{T}} x = 0$ and we immediately obtain a strengthened form of Theorems 3–6. To this end, let $P$ be the polytope defined by (2.5) for $A$ given by (5.2).

**Theorem 3′.** *The polytope $P \neq \{0\}$ if and only if $C[G(A)] \neq \{0\}$.*

**Theorem 4′.** *Suppose $P \neq \{0\}$. Then $u$ is an extreme point of $P$ if and only if $u \in P$, $\|u\| = 2$ and* $\dim C(u) = 1$.

**Theorem 5′.** *Suppose $P \neq \{0\}$. Then $u$ is an extreme point of $P$ if and only if $u = \alpha x \in P$ where $x$ is a cycle of $G(A)$ and $\alpha = 2\|x\|^{-1}$.*

A *polygon* $\pi$ *of* $B$ *of length* $|\pi| = k$ (cf., Saunders and Schneider [14]) is defined to be a pair of sequences $(i_1, \ldots, i_k)$, $i_r \in \langle n_1 \rangle$ and $(j_1, \ldots, j_k)$, $j_r \in \langle n_2 \rangle$, each of distinct integers such that $B_{i_r j_r} > 0$ and $B_{i_{r+1} j_r} > 0$ (with $i_{k+1} = i_1$), $r = 1, \ldots, k$. Of course, the cycles of $G$ correspond to polygons of $B$. For a polygon $\pi$ put

$$\gamma(\pi) = \frac{B_{i_1 j_1} B_{i_2 j_2} \cdots B_{i_k j_k}}{B_{i_2 j_1} B_{i_3 j_2} \cdots B_{i_1 j_k}}.$$

We now obtain

**Theorem 6′.** *Let* $0 \neq B \in R_+^{n_1 n_2}$ *and let* $\beta$ *be given by* (5.1). *Then* $\beta = 1$ *if* $B$ *has no polygons. Otherwise,*

$$\beta = \max_{\pi \in \Pi} \sqrt[|\pi|]{\gamma(\pi)}$$

*where* $\Pi$ *is the set of polygons of* $B$.

## Appendix: A minmax theorem for polyhedra

The purpose of this appendix is to give a proof for Wolfe's [18] minmax theorem for polyhedra with a binlinear function (communicated to us privately by B.C. Eaves).

**Theorem.** *Let* $\emptyset \neq X \subseteq R^n$ *and* $\emptyset \neq Y \subseteq R^m$ *be polyhedra sets,* $Q \in R^{nm}$, $\eta = \inf_{x \in X} \sup_{y \in Y} x^T Q y$ *and* $\delta = \sup_{y \in Y} \inf_{x \in X} x^T Q y$. *Then, either* $\eta = \delta$ *or* $\eta = +\infty$ *and* $\delta = -\infty$. *Also, if* $\eta$ *(or equivalently* $\delta$*) are finite, then there exists* $x_0 \in X$ *and* $y_0 \in Y$ *such that*

$$\eta = \sup_{y \in Y} x_0^T Q y = x_0^T Q y_0 = \inf_{x \in X} x^T Q y_0 = \delta. \tag{A.1}$$

**Proof.** We prove the theorem for polyhedra $X$ and $Y$ having a representation

$$X = \{x \in R^n \mid Ax \geq a, x \geq 0\} \tag{A.2}$$

and

$$Y = \{y \in R^m \mid By \leq b, y \geq 0\}, \tag{A.3}$$

where $A \in R^{kn}$, $B \in R^{tm}$, $a \in R^k$ and $b \in R^t$ and $k$ and $t$ are some positive integers. The proof of the general case follows similarly.

Let $X^* = \{x \in X \mid \sup_{y \in Y} x^T Q y < \infty\}$ and $Y_* = \{y \in Y \mid \inf_{x \in X} x^T Q y > -\infty\}$. If $\eta < \infty$ (resp., $\delta > -\infty$), we have $X^* \neq \emptyset$ (resp., $Y_* \neq \emptyset$). We shall next assume that $\eta < \infty$ and $\delta > -\infty$. By the duality theory of linear programming, (e.g., Gale [9, p. 78]), if $x \in X^*$, then

$$\sup_{y\in Y} x^{\mathrm T}Qy = \max_{\substack{By\le b\\ y\ge 0}} (x^{\mathrm T}Q)y = \min_{\substack{\lambda^{\mathrm T}B\ge x^{\mathrm T}Q\\ \lambda\ge 0}} \lambda^{\mathrm T}b \tag{A.4}$$

and if $x\in X > X^*$, then $\{\lambda\in R^t \mid \lambda^{\mathrm T}B\ge x^{\mathrm T}Q, \lambda\ge 0\}=\emptyset$. So, (A.2), (A.3) and (A.4) assure that

$$\begin{aligned}\eta &= \inf_{x\in X}\sup_{y\in Y} x^{\mathrm T}Qy = \inf_{x\in X^*}\sup_{y\in Y} x^{\mathrm T}Qy \\ &= \inf_{x\in X^*}\min_{\substack{\lambda^{\mathrm T}B\ge x^{\mathrm T}Q\\ \lambda\ge 0}} \lambda^{\mathrm T}b = \inf_{\substack{x\in X,\\ \lambda^{\mathrm T}B\ge x^{\mathrm T}Q\\ \lambda\ge 0}} \lambda^{\mathrm T}b \\ &= \inf_{\substack{(\lambda^{\mathrm T},x^{\mathrm T})\begin{pmatrix}B & 0\\ -Q & A^{\mathrm T}\end{pmatrix}\ge(0,a^{\mathrm T})\\ \lambda,x\ge 0}} (\lambda^{\mathrm T},x^{\mathrm T})\begin{pmatrix}b\\0\end{pmatrix}.\end{aligned} \tag{A.5}$$

$$\begin{aligned}\delta &= \sup_{y\in Y}\inf_{x\in X} x^{\mathrm T}Qy = \sup_{y\in Y_*}\inf_{x\in X} x^{\mathrm T}Qy \\ &= \sup_{y\in Y_*}\min_{\substack{x^{\mathrm T}A^{\mathrm T}\ge a\\ x\ge 0}} x^{\mathrm T}(Qy) = \sup_{y\in Y_*}\max_{\substack{A^{\mathrm T}\mu\le Qy\\ \mu\ge 0}} a^{\mathrm T}\mu \\ &= \sup_{\substack{y\in Y\\ A^{\mathrm T}\mu\le Qy\\ \mu\ge 0}} a^{\mathrm T}\mu = \sup_{\substack{\begin{pmatrix}B & 0\\ -Q & A^{\mathrm T}\end{pmatrix}\begin{pmatrix}y\\ \mu\end{pmatrix}\le\begin{pmatrix}b\\0\end{pmatrix}\\ y,\mu\ge 0}} (0,a^{\mathrm T})\begin{pmatrix}y\\ \mu\end{pmatrix}.\end{aligned} \tag{A.6}$$

Another application of the duality theorem assures that $\eta=\delta$, and that if $\eta$ and $\delta$ are finite, then there exist $x_0\in X^*$ (resp., $y_0\in Y_*$) such that

$$\sup_{y\in Y} x_0^{\mathrm T}Qy = \eta = \delta = \inf_{x\in X} x^{\mathrm T}Qy_0. \tag{A.7}$$

In particular, (A.7) implies that

$$x_0^{\mathrm T}Qy_0 \le \eta = \delta \le x_0^{\mathrm T}Qy_0, \tag{A.8}$$

and (A.1) follows by combining (A.7) and (A.8).

**Remark.** If the function of $x$ and $y$ given by $(x,y)\to x^{\mathrm T}Qy$ is replaced by $(x,y)\to x^{\mathrm T}Qy+x^{\mathrm T}p+r^{\mathrm T}y$ one can still obtain the results. This can be done by observing that $x^{\mathrm T}Qy+x^{\mathrm T}p+r^{\mathrm T}y=\bar x^{\mathrm T}\bar Q\bar y$

$$\bar x=\begin{pmatrix}x\\ e_{(n)}\end{pmatrix},\quad \bar y=\begin{pmatrix}y\\ e_{(m)}\end{pmatrix}\quad\text{and}\quad \bar Q=\begin{pmatrix}Q & (pe_{(m)}^{\mathrm T})m^{-1}\\ (e_{(n)}r^{\mathrm T})n^{-1} & 0\end{pmatrix}$$

where $e_{(k)}=(1,\dots,1)^{\mathrm T}\in R^k$ for an integer $k$.

**Example.** Let $xQy=x-y$, $X=R$ and $Y=R$. It is easy to see that $\eta=-\delta=\infty$.

## References


[1] G. Aumann, "Über approximative Nomographie, I", *Bayerische Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse, Sitzungsberichte* (1958) 137–155.

[2] G. Aumann, "Über approximative Nomographie, II", *Bayerische Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse, Sitzungsberichte* (1959) 103–109.

[3] F.L. Bauer, "Optimally scaled matrices", *Numerische Mathematik* 5 (1963) 73–87.

[4] F.L. Bauer, "Remarks on optimally scaled matrices", *Numerische Mathematik* 13 (1969) 1–3.

[5] C. Berge, *Graphs and hypergraphs* (North-Holland, Amsterdam, 1973).

[6] A.R. Curtis and J.K. Reid, "On the automatic scaling of matrices for Gaussion elimination", *Journal of the Institute of Mathematics and its Applications* 10 (1972) 118–124.

[7] S.P. Diliberto and E.G. Straus, "On the approximation of a function of several variables by the sum of functions of fewer variables", *Pacific Journal of Mathematics* 1 (1951) 195–210.

[8] D.R. Fulkerson and P. Wolfe, "An algorithm for scaling matrices", *SIAM Review* 4 (1962) 142–146.

[9] D. Gale, *The theory of linear economics models* (McGraw-Hill, New York, 1966).

[10] M. v. Golitscheck, "An algorithm for scaling matrices and computing the minimum cycle mean in a digraph", manuscript (1979).

[11] M. v. Golitscheck, "Approximation of functions of two variables by the sum of two functions of one variable", manuscript (1979).

[12] R.W. Hamming, *Introduction to applied numerical analysis* (McGraw-Hill, New York, 1971).

[13] W. Orchard-Hays, *Advanced linear programming computing techniques* (McGraw-Hill, New York, 1968).

[14] B.D. Saunders and H. Schneider, "Flows on graphs applied to diagonal similarity and diagonal equivalence for matrices", *Discrete Mathematics* 24 (1978) 202–220.

[15] B.D. Saunders and H. Schneider, "Cones, graphs and optimal scalings of matrices", *Linear and Multilinear Algebra* 8 (1979) 121–135.

[16] J.A. Tomlin, "On scaling linear programming problems", *Mathematical Programming Study* 4 (1975) 146–166.

[17] W. Tutte, "A class of Abelian groups", *Canadian Journal of Mathematics* 8 (1956) 13–28.

[18] P. Wolfe, "Determinateness of polyhedral games", in: *Linear inequalities and related systems*, Annals of mathematical studies 38 (Princeton University Press, Princeton, NJ, 1956) 195–198.

[19] R.E. Bank, "An automatic scaling procedure for a D'Yakanov–Gunn iteration scheme", *Linear Algebra and its Applications* 28 (1979) 17–33.