Universität Augsburg

Institut für Mathematik

Norbert Gaffke, Friedrich Pukelsheim

Vector and matrix apportionment problems and separable convex integer optimization

Preprint Nr. 006/2007 — 22. Mai 2007

Institut für Mathematik, Universitätsstraße, D-86 135 Augsburg http://www.math.uni-augsburg.de/

Impressum:

Herausgeber:

Institut für Mathematik
Universität Augsburg
86135 Augsburg
`http://www.math.uni-augsburg.de/forschung/preprint/`

ViSdP:

Friedrich Pukelsheim
Institut für Mathematik
Universität Augsburg
86135 Augsburg

Vector and matrix apportionment problems and separable convex integer optimization

N. Gaffke* and F. Pukelsheim**

*University of Magdeburg
Faculty of Mathematics
PF 4120, D-39016 Magdeburg
Norbert.Gaffke@Mathematik.Uni-Magdeburg.De

**University of Augsburg
Institute of Mathematics
D-86135 Augsburg
Pukelsheim@Math.Uni-Augsburg.De

Abstract

Algorithms for the proportional rounding of a nonnegative vector, and for the biproportional rounding of a nonnegative matrix are discussed. Here we view vector and matrix rounding as special instances of a generic optimization problem that employs an additive version of the objective function of Gaffke and Pukelsheim (2007). The generic problem turns out to be a separable convex integer optimization problem, in which the linear equality constraints are given by a totally unimodular coefficient matrix. So, despite the integer restrictions of the variables, Fenchel duality applies. Our chief goal is to study the implied algorithmic consequences. We establish a general algorithm based on the primal optimization problem. Furthermore we show that the biproportional algorithm of Balinski and Demange (1989), when suitably generalized, derives from the dual optimization problem. Finally we comment on the shortcomings of the alternating scaling algorithm, a discrete variant of the well-known Iterative Proportional Fitting procedure.

Short title: Apportionment and separable integer optimization.

Key words. Totally unimodular matrix – Elementary vector – Graver basis – Convex programming duality – Alternating maximization procedure

1 Introduction

A *separable* objective function is of the form

$$F(x) \,=\, \sum_{e\in E} f_e(x_e) \ ,$$

where $x = (x_e)_{e\in E} \in \mathbb{R}^E$ is a (column) vector variable whose components we label, for convenience, by the elements e of some finite set E, and f_e (for $e \in E$) are real functions of a real variable. By $\mathbb{Z}$ we denote the set of all integers, and by $\mathbb{Z}^E$ the set of all integer vectors in $\mathbb{R}^E$. Let $\mu = (\mu_e)_{e\in E} \in \mathbb{Z}^E$ be a *positive* vector, i.e., its components are positive integers, which will define a componentwise upper bound for the vector variable x. We assume that each function f_e is a *convex* function on the interval $[\,0\,,\,\mu_e]$.

Let A be a given *totally unimodular* $V \times E$ matrix, where V is another finite set, (so the rows of A are labelled by the elements $v \in V$ and the columns of A are labelled by the elements $e \in E$). Recall that total unimodularity of A means that all square submatrices of A have determinants -1, 0, or $+1$. In particular, all the entries of A are in $\{-1, 0, +1\}$. Let $b \in \mathbb{Z}^V$ be given such that linear system

$$Ax \,=\, b\ , \quad 0 \le x \le \mu\,, \tag{1.1}$$

has a solution for $x \in \mathbb{R}^E$ and hence also a solution $x^{(0)} \in \mathbb{Z}^E$, (cf. Schrijver (1999, Theorem 19.3)). Note that '$\le$' between vectors stands for the usual componentwise semi-ordering. So $0 \le x \le \mu$ means $0 \le x_e \le \mu_e$ for all $e \in E$. We will consider the integer extremum problem,

$$\text{minimize} \quad F(x) \;=\; \sum_{e\in E} f_e(x_e) \tag{1.2}$$

$$\text{subject to} \qquad x = (x_e)_{e\in E} \in \mathbb{Z}^E\,,\ 0 \le x \le \mu\,, \quad Ax \;=\; b\;. \tag{1.3}$$

Clearly, only the values of f_e at the integers points in $\{0, 1, \ldots, \mu_e\}$ enter into the problem, and the convexity of f_e enters only by its $\mathbb{Z}$-convexity, (cf. Hemmeke (2003)), i.e., the increments $\Delta f_e(n) \;=\; f_e(n) - f_e(n-1)$ are nondecreasing in $n \in \{1, \ldots, \mu_e\}$. For technical reasons we extend the definition of the increments to $n = 0$ and $n = \mu_e + 1$ by

$$\Delta f_e(n) \;=\; \begin{cases} -\infty & ,\ \text{if } n = 0 \\ f_e(n) \;-\; f_e(n-1) & ,\ \text{if } 1 \le n \le \mu_e \\ +\infty & ,\ \text{if } n = \mu_e + 1 \end{cases} \tag{1.4}$$

So, without loss of generality, we may assume the convex functions f_e to be *piecewise linear*,

$$f_e(t) \;=\; f_e(n-1) \;+\; \Delta f_e(n)\,\big(t - (n-1)\big)\;, \quad \text{if} \;\; n-1 \le t \le n \;\; \text{and} \;\; n \in \{1, \ldots, \mu_e\}\,. \tag{1.5}$$

In fact, since the slopes $\Delta f_e(n)$ are nondecreasing in n, the function f_e from (1.5) is convex on $[\,0\,,\,\mu_e]$. Two special cases are of particular interest.

Vector apportionment problem

A simple special case is given when V a one-point set, $E = \{1, \ldots, p\}$, and $A \;=\; [1, \ldots, 1]$. The contraints in (1.3) then read as

$$x = (x_1, \ldots, x_p)' \in \mathbb{Z}^p\,, \quad 0 \le x \le \mu \quad \sum_{j=1}^{p} x_j \;=\; h\,, \tag{1.6}$$

for a given positive integer h, the "house size". Trivially, consistency of (1.1) means here that $h \le \sum_{j=1}^{p} \mu_j$. A problem of minimizing (1.2) (with $E \;=\; \{1, \ldots, p\}$) subject to (1.6) will be referred to as a *vector apportionment problem*. For this problem, but without upper bounds μ_j, the optimal solutions were characterized in Saaty (1970, p.184), and for special functions f_j the problem was treated by Te Riele (1978) and Thépot (1986). As it is shown in Gaffke and Pukelsheim (2007), proportional rounding of a positive vector $w \;=\; (w_1, \ldots, w_p)' \in \mathbb{R}^p$ can be written as a vector apportionment problem employing functions $f_e \;=\; f_j$ $(1 \le j \le p)$ such that

$$f_j(n) \;=\; \sum_{k=1}^{n} \log \frac{s(k)}{w_j}\;, \quad (n \;=\; 0, 1, \ldots, \mu_j)\,,$$

$$\text{whence} \qquad f_j(0) \;=\; 0 \quad \text{and} \quad \Delta f_j(n) \;=\; \log \frac{s(n)}{w_j}\;,\; (n \;=\; 1, \ldots, \mu_j)\;,$$

where $s(n)$, $n = 1, 2, 3, \ldots$, is a given signpost sequence defining the rounding law, i.e.

$$0 \;<\; s(1) \;<\; s(2) \;<\; s(3) \;<\; \ldots\;, \quad \text{and} \;\; n - 1 \le s(n) \le n \;\; \text{for all } n \ge 1\,.$$

(Actually, this is the case of a *pervious* rounding law in that $s(1) \;>\; 0$; the *impervious* case $s(1) = 0$ can be treated similarly). □

Matrix apportionment problem

Another particular (but more difficult) case is given when $V = \{R_1, \ldots, R_k, C_1, \ldots, C_\ell\}$, a set of size $k + \ell$, where $k \geq 2$ and $\ell \geq 2$, E is a nonempty subset of the set of all (ordered) pairs (i, j), $1 \leq i \leq k$, $1 \leq j \leq \ell$, and $A = (a_{v,e})_{v\in V,\, e\in E}$ is given by

$$a_{v,e} = \begin{cases} 1 & \text{, if } v = R_i \text{ and } e = (i,j) \text{ for some } j \\ 1 & \text{, if } v = C_j \text{ and } e = (i,j) \text{ for some } i \\ 0 & \text{, else} \end{cases} \tag{1.7}$$

That is, A is the vertex-edge incidence matrix of a bipartite (undirected) graph with vertices $R_1, \ldots, R_k$ and $C_1, \ldots, C_\ell$, and there is an edge between R_i and C_j iff $(i, j) \in E$. Thus A is totally unimodular, (cf. Schrijver (1999, Section 19.3, Example 1)). The contraints in (1.3) turn into

$$x = (x_{i,j})_{(i,j)\in E} \in \mathbb{Z}^E , \quad 0 \leq x \leq \mu , \quad x_{i,+} = r_i \ \forall\, i , \quad x_{+,j} = c_j \ \forall\, j , \tag{1.8}$$

where we have used the notation

$$x_{i,+} = \sum_{j\,:\,(i,j)\in E} x_{i,j} \ , \quad x_{+,j} = \sum_{i\,:\,(i,j)\in E} x_{i,j} \ ,$$

and where $b = (r_1, \ldots, r_k, c_1, \ldots, c_\ell)'$, the r_i and c_j being positive integers. Of course, it is assumed that $\sum_{i=1}^k r_i = \sum_{j=1}^\ell c_j = h$, the house size. A problem of minimizing (1.2) under (1.8) will be referred to as a *matrix apportionment problem.* In Gaffke and Pukelsheim (2007) it was shown that biproportional rounding of a nonnegative real matrix $W = \big(w_{i,j}\big)_{\substack{1\leq i\leq k \\ 1\leq j\leq \ell}}$ can be written as a matrix apportionment problem employing $E = \big\{(i,j) : w_{i,j} > 0\big\}$ and functions $f_e = f_{i,j}$ such that

$$f_{i,j}(n) = \sum_{k=1}^{n} \log \frac{s(k)}{w_{i,j}} \ , \quad (n = 0, 1, \ldots, \mu_{i,j}) ,$$

$$\text{whence} \qquad f_{i,j}(0) = 0 \quad \text{and} \quad \Delta f_{i,j}(n) = \log \frac{s(n)}{w_{i,j}} \ , \ (n = 1, \ldots, \mu_{i,j}) \ . \qquad \square$$

There is a considerable body of literature on separable convex programming (integer or continuous) with linear constraints, providing efficient algorithms for solution, (cf. Hochbaum and Shantikumar (1990)). These results are still to be exploited for (bi)proportional rounding purposes. More general nonlinear integer optimization problems are considered in Murota, Saito, Weismantel (2004) and in Hemmecke (2003). We will concentrate on separable convex integer programming problems under totally unimodular linear equations.

Our present paper is organized as follows. In Section 2 a characterization of the optimal solutions to the primal integer problem (1.2)-(1.3) is given offering a basis for the primal algorithm outlined in Section 3. A duality result is derived in Section 4, and a conceptual dual algorithm is formulated in Section 5. In Sections 6 and 7 we concentrate on the two instances mentioned above, vector and matrix apportionment problems. For vector apportionment problems, the dual algorithm coincides with the one of Happacher and Pukelsheim (1996, p. 378; 2000, p. 154), and Dorfleitner and Klein (1999). For matrix apportionment problems, the dual algorithm is akin to the one described by Balinski and Demange (1989), and by Balinski and Rachev (1997, Section 5), see also Balinski (2006) and Rote and Zachariasen (2007). Section 8 is concerned with an alternative dual method, the alternating scaling algorithm, which requires relatively low computational effort. However, in general it may fail to find the optimum due to non-smoothness of the dual objective function. Despite this deficiency, the alternating scaling method is a useful heuristics which provides a nearly optimal solution, and in many instances even an optimal solution.

2 The primal problem

We address problem (1.2)-(1.3) under the assumptions stated in Section 1. For the (totally unimodular) matrix A its nullspace and the orthogonal complement of the latter, which is the range of the transposed A', will be of particular interest,

$$\begin{aligned} \mathcal{N}(A) &= \left\{ x \in \mathbb{R}^E \,:\, Ax = 0 \right\} , \\ \mathcal{R}(A') &= \left\{ y \in \mathbb{R}^E \,:\, \exists\, \lambda \in \mathbb{R}^V \text{ with } y = A'\lambda \right\} . \end{aligned}$$

The *support* of a vector $x = (x_e)_{e \in E} \in \mathbb{R}^E$ is defined by

$$\text{supp}(x) = \left\{ e \in E \,:\, x_e \neq 0 \right\} .$$

Below we will have to further classify the supporting indices of a vector $x = (x_e)_{e \in E} \in \mathbb{R}^E$ by introducing

$$E^+(x) = \left\{ e \in E \,:\, x_e > 0 \right\} \quad \text{and} \quad E^-(x) = \left\{ e \in E \,:\, x_e < 0 \right\} .$$

Let $\mathcal{L}$ be a linear subspace of $\mathbb{R}^E$. An *elementary vector* of $\mathcal{L}$ is defined to be a nonzero vector $z \in \mathcal{L}$ which has minimal support within $\mathcal{L} \setminus \{0\}$, i.e., $0 \neq z \in \mathcal{L}$ and for all $0 \neq x \in \mathcal{L}$:

$$\text{supp}(x) \subseteq \text{supp}(z) \;\text{ implies }\; \text{supp}(x) = \text{supp}(z) ,$$

cf. Rockafellar (1972, pp. 203-204). From the total unimodularity of the matrix A we get:

Lemma 2.1 *If z is an elementary vector of $\mathcal{N}(A)$ then, for some positive scalar γ, the vector $\gamma\, z$ has all components in $\{-1, 0, +1\}$.*

Proof. Let $z = (z_e)_{e \in E}$ be an elementary vector of $\mathcal{N}(A)$. Denote $\widetilde{E} = \text{supp}(z)$ and consider the subvector $\widetilde{z} = (z_e)_{e \in \widetilde{E}} \in \mathbb{R}^{\widetilde{E}}$. Let a^e, $(e \in E)$, be the columns of A and consider the $V \times \widetilde{E}$ submatrix $\widetilde{A}$ with columns a^e, $e \in \widetilde{E}$. Clearly, $\widetilde{A}$ is again totally unimodular. The nullspace of $\widetilde{A}$ has dimension equal to 1 (and consists thus of all scalar multiples of $\widetilde{z}$), which can be seen as follows. Let $\widetilde{x} = (x_e)_{e \in \widetilde{E}} \in \mathbb{R}^{\widetilde{E}}$ with $\widetilde{A}\widetilde{x} = 0$, and let $\widetilde{x} \neq 0$. Then, we augment $\widetilde{x}$ by zero components $x_e = 0$, $e \in E \setminus \widetilde{E}$, to obtain a vector $x \in \mathbb{R}^E$. We have $x \in \mathcal{N}(A)$ and $\text{supp}(x) \subseteq \text{supp}(z)$, and hence $\text{supp}(x) = \text{supp}(z)$. So x is also an elementary vector of $\mathcal{N}(A)$ with the same support as z which implies, (cf. Rockafellar (1972, Lemma 22.4)), that $x = \beta z$ for some nonzero scalar β, and thus $\widetilde{x} = \beta\widetilde{z}$. Hence the nullspace of $\widetilde{A}$ is spanned by $\widetilde{z}$.
Now we identify, for the nullspace of $\widetilde{A}$, another basis vector which has all components equal to ± 1 or zero. Consider the polytope

$$\widetilde{\mathcal{P}} = \left\{ \widetilde{x} = (x_e)_{e \in \widetilde{E}} \in \mathbb{R}^{\widetilde{E}} \,:\, \widetilde{A}\widetilde{x} = 0 \;,\; -1 \leq x_e \leq 1 \;\forall\; e \in \widetilde{E} \right\} .$$

By total unimodularity of $\widetilde{A}$ all of the vertices of $\widetilde{\mathcal{P}}$ are integral, (cf. Schrijver (1999, Theorem 19.3)). Since $\alpha\widetilde{z} \in \widetilde{\mathcal{P}}$ for some nonzero scalar α, we have $\widetilde{\mathcal{P}} \neq \{0\}$, and so there is a nonzero vertex $\widetilde{x}^*$ of $\widetilde{\mathcal{P}}$ with all components equal to ± 1 or zero. In particular, $\widetilde{x}^*$ is an element of the nullspace of $\widetilde{A}$ and thus $\widetilde{x}^* = \gamma\widetilde{z}$ for some nonzero scalar γ. Augmenting $\widetilde{x}^*$ by zero components to obtain a vector x^* of $\mathbb{R}^E$, we have $\gamma z = x^*$ which has all components in $\{-1, 0, +1\}$. If $\gamma < 0$ then the same is true for $(-\gamma)z = -x^*$. □

We will call an elementary vector of $\mathcal{N}(A)$ which has all components equal to ± 1 or zero an *elementary sign vector* of $\mathcal{N}(A)$. Using the results of Graver (1975) it can be shown that the elementary sign vectors of $\mathcal{N}(A)$ constitute the *Graver basis* of A which is defined as follows (and actually refers to *any* integer matrix A). The Graver basis of A consists of all vectors which are minimal in the set of all nonzero integer vectors of $\mathcal{N}(A)$ w.r.t. the semi-ordering "$\preceq$" defined by:

$$x = (x_e)_{e\in E} \preceq y = (y_e)_{e\in E} \iff x_e y_e \geq 0 \text{ and } |x_e| \leq |y_e| \text{ for all } e \in E\,,$$

(cf. Hemmecke (2003), p. 1). A slightly weaker notion we will also use is that of a *sign vector* of $\mathcal{N}(A)$, which is any nonzero vector of $\mathcal{N}(A)$ having all components equal to ± 1 or zero. For a sign vector $z = (z_e)_{e\in E}$ of $\mathcal{N}(A)$ we obviously have

$$E^+(z) = \big\{e \in E : z_e = +1\big\} \quad \text{and} \quad E^-(z) = \big\{e \in E : z_e = -1\big\}\,.$$

Lemma 2.2 *Let $c_e \in \mathbb{R}\cup\{-\infty\}$ and $d_e \in \mathbb{R}\cup\{+\infty\}$ with $c_e \leq d_e$ for all $e \in E$ be given. Then one and only one of the following two alternatives* (a) *and* (b) *holds:*

(a) *There exists a vector $y = (y_e)_{e\in E} \in \mathcal{R}(A')$ with $c_e \leq y_e \leq d_e$ for all $e \in E$.*

(b) *There exists a sign vector z of $\mathcal{N}(A)$ such that*

$$\sum_{e\in E^-(z)} c_e > \sum_{e\in E^+(z)} d_e\,.$$

Moreover, condition (b) *is equivalent to the following condition* (b*)*:*

(b*) *There exists an elementary sign vector z of $\mathcal{N}(A)$ such that*

$$\sum_{e\in E^-(z)} c_e > \sum_{e\in E^+(z)} d_e\,.$$

□

The result of Lemma 2.2 is a fairly direct consequence from Rockafellar (1972, Theorem 22.6), and our Lemma 2.1, (see the proof of Theorem 7.1 in Gaffke and Pukelsheim (2007)). It can also be derived from strong duality in linear programming.
Note that the inequality in (b) and (b*) of Lemma 2.2 in particular implies that $c_e > -\infty$ for all $e \in E^-(z)$ and $d_e < +\infty$ for all $e \in E^+(z)$.

Using Lemmas 2.1 and 2.2 we now derive two (equivalent) characterizations of an optimal solution to problem (1.2)-(1.3). The first shows the elementary sign vectors of $\mathcal{N}(A)$ to constitute a universal test set in the sense of Hemmecke (2003); this follows also from the more general results of that paper (see p. 4 in Hemmecke (2003)). The second characterization is of dual (Lagrangian) type; this is related to a result in Sun, Tsai, and Qi (1993, Proposition 2.3) who deal with the case of a network matrix A. However, we will give a short proof of our next theorem by means of Lemmas 2.1 and 2.2. Recall the definition in (1.4) of the increments $\Delta f_e(n)$, $n \in \{0, 1, \ldots, \mu_e + 1\}$, which are nondecreasing in n.

Theorem 2.3 *Let $x^* = \big(x^*_e\big)_{e\in E}$ be a feasible solution to problem* (1.2)-(1.3)*, (i.e., x^* satisfies* (1.3)*). The following three conditions* (i), (ii)*, and* (iii) *are equivalent:*

(i) *x^* is an optimal solution to problem* (1.2)-(1.3).

(ii) *For all elementary sign vectors z of $\mathcal{N}(A)$ with $E^+(z) \subseteq \{e : x^*_e < \mu_e\}$ and $E^-(z) \subseteq \{e : x^*_e > 0\}$ one has $F(x^*) \leq F(x^* + z)$.*

(iii) *There exists a vector $y^* = (y^*_e)_{e\in E} \in \mathcal{R}(A')$ such that*

$$\Delta f_e(x^*_e) \leq y^*_e \leq \Delta f_e(x^*_e + 1) \quad \forall\, e \in E\,.$$

Proof.
(i) $\Longrightarrow$ (ii) Assume (i). Let $z = (z_e)_{e \in E}$ be an elementary sign vector of $\mathcal{N}(A)$ such that $E^+(z) \subseteq \{e : x_e^* < \mu_e\}$ and $E^-(z) \subseteq \{e : x_e^* > 0\}$. Then $x^* + z$ is again feasible for problem (1.2)-(1.3), and thus $F(x^*) \leq F(x^* + z)$.
(ii) $\Longrightarrow$ (iii) Assume (ii). Let $z = (z_e)_{e \in E}$ be an elementary sign vector of $\mathcal{N}(A)$ such that $E^+(z) \subseteq \{e : x_e^* < \mu_e\}$ and $E^-(z) \subseteq \{e : x_e^* > 0\}$. Then,

$$
\begin{aligned}
0 \leq F(x^* + z) - F(x^*) &= \sum_{e \in E} \big(f_e(x_e^* + z_e) - f_e(x_e^*) \big) \\
&= \sum_{e \in E^+(z)} \big(f_e(x_e^* + 1) - f_e(x_e^*) \big) + \sum_{e \in E^-(z)} \big(f_e(x_e^* - 1) - f_e(x_e^*) \big) \\
&= \sum_{e \in E^+(z)} \Delta f_e(x_e^* + 1) - \sum_{e \in E^-(z)} \Delta f_e(x_e^*) ,
\end{aligned}
$$

which shows that

$$\sum_{e \in E^-(z)} \Delta f_e(x_e^*) \leq \sum_{e \in E^+(z)} \Delta f_e(x_e^* + 1) . \tag{2.1}$$

Inequality (2.1) remains true for *any* elementary sign vector z of $\mathcal{N}(A)$, since if one or both of the inclusions $E^+(z) \subseteq \{e : x_e^* < \mu_e\}$ and $E^-(z) \subseteq \{e : x_e^* > 0\}$ are *not* satisfied then the right hand side of (2.1) becomes $+\infty$ or the left hand side of (2.1) becomes $-\infty$. Now Lemma 2.2 applies to

$$c_e = \Delta f_e(x_e^*) \quad \text{and} \quad d_e = \Delta f_e(x_e^* + 1) , \quad (e \in E) ,$$

and shows that alternative (a) of that lemma must hold, which is condition (iii).
(iii) $\Longrightarrow$ (i) Assume (iii) for some $y^* \in \mathcal{R}(A')$. Let $x = (x_e)_{e \in E}$ be any feasible point to problem (1.2)-(1.3). By the convexity of the functions f_e we have for every $e \in E$,

$$
\begin{aligned}
f_e(x_e) - f_e(x_e^*) &\geq \Delta f_e(x_e^* + 1)\,(x_e - x_e^*) \geq y_e^*(x_e - x_e^*) , \quad \text{if } x_e \geq x_e^* , \\
f_e(x_e) - f_e(x_e^*) &\geq \Delta f_e(x_e^*)\,(x_e - x_e^*) \geq y_e^*(x_e - x_e^*) , \quad \text{if } x_e < x_e^* .
\end{aligned}
$$

Summing over $e \in E$, and observing that $y^* = A'\lambda^*$ for some $\lambda^* \in \mathbb{R}^V$ and $Ax = Ax^* = b$, we obtain

$$F(x) - F(x^*) \geq (A'\lambda^*)'(x - x^*) = \lambda^{*\prime}(Ax - Ax^*) = 0 .$$

Thus, $F(x^*) \leq F(x)$ for every feasible point x to problem (1.2)-(1.3). □

3 A conceptual primal algorithm

Suppose that we have an algorithm, let us call it an *Oracle X*, which decides between the alternatives (a) and (b) of Lemma 2.2. More precisely, for any given input values c_e and d_e, $(e \in E)$, as in Lemma 2.2, suppose that Oracle X either returns a vector $y \in \mathcal{R}(A')$ with $c_e \leq y_e \leq d_e \ \forall\, e \in E$, or it returns a sign vector z of $\mathcal{N}(A)$ such that

$$\sum_{e \in E^-(z)} c_e > \sum_{e \in E^+(z)} d_e .$$

By linear programming methods it should be possible to construct an Oracle X of polynomially (in $\#V + \#E$) bounded running time. For vector and matrix apportionment problems specific Oracles X will be given in Sections 6 and 7. However, a primal algorithm stated next for solving problem (1.2)-(1.3), which is based on an Oracle X, will *not* be polynomial due to an exponentially increasing size of the feasible region (1.3).

Conceptual primal algorithm, (needs an Oracle X).
Start with any feasible point $x = \big(x_e\big)_{e\in E}$ to problem (1.2)-(1.3). Set

$$c_e \;=\; \Delta f_e(x_e) \quad \text{and} \quad d_e \;=\; \Delta f_e(x_e+1) \qquad \forall\, e \in E\,,$$

and apply Oracle X. If the oracle yields a point $y \in \mathcal{R}(A')$ satisfying

$$c_e \;\leq\; y_e \;\leq\; d_e \quad \forall\, e \in E\,,$$

then, by Theorem 2.3, x is optimal. If the oracle yields a sign vector $z = (z_e)_{e\in E}$ of $\mathcal{N}(A)$ such that

$$\sum_{e\in E^-(z)} c_e \;>\; \sum_{e\in E^+(z)} d_e \;,$$

then define a new point by $\widetilde{x} = x + z$. Clearly, $\widetilde{x}$ is feasible to problem (1.2)-(1.3), and

$$\begin{aligned} F(x) - F(\widetilde{x}) \quad &= \quad \sum_{e\in E^-(z)} \Delta f_e(x_e) \;-\; \sum_{e\in E^+(z)} \Delta f_e(x_e+1) \\ &= \quad \sum_{e\in E^-(z)} c_e \;-\; \sum_{e\in E^+(z)} d_e \;>\; 0\;. \end{aligned}$$

Hence $\widetilde{x}$ is strictly better than x, $F(\widetilde{x}) < F(x)$. Replace x by $\widetilde{x}$ and repeat. Since the feasible region (1.3) is finite, the algorithm will terminate with an optimal solution after a finite number of iterations. □

4 The dual problem

Strong duality of convex programming applies to the primal problem (1.2)-(1.3), despite the integer restriction in (1.3). This is due to the total unimodularity of the matrix A. For, as pointed out in Section 1, the (convex) functions f_e may be taken to be the piecewise linear functions from (1.5). Doing so, we consider the relaxed version of the primal problem by removing the integer restriction,

$$\text{minimize} \quad F(x) \;=\; \sum_{e\in E} f_e(x_e) \tag{4.1}$$

$$\text{subject to} \quad x = (x_e)_{e\in E} \in \mathbb{R}^E\,, \quad 0 \leq x \leq \mu\,, \quad Ax \;=\; b\,, \tag{4.2}$$

which is a convex separable piecewise-linear program as studied in Fourer (1985). In fact, by the total unimodularity of A (and since b and μ are integer vectors), an optimal solution to the relaxed problem (4.1)-(4.2) is close to an optimal solution to the integer problem (1.2)-(1.3), and the two problems share the same optimal value. So, the integer problem and the relaxed version are nearly equivalent. This is shown by the following lemma.

Lemma 4.1 *Let f_e $(e \in E)$ be the piecewise linear convex functions from* (1.5)*. If x^* is an optimal solution to the relaxed problem* (4.1)-(4.2) *then there exists a rounding of the non-integer components of x^* to one of the neighboring integers such that the obtained (rounded) point x^{**} is again an optimal solution to problem* (4.1)-(4.2) *and thus also an optimal solution to the primal integer problem* (1.2)-(1.3)*.*

Proof. Let $x^* = (x^*_e)_{e\in E}$ be an optimal solution to problem (4.1)-(4.2), (which exists by compactness of the feasible region (4.2) and by continuity of the objective function F). Let

$a_e^* = \lfloor x_e^* \rfloor$ (the greatest integer not exceeding x_e^*), $\xi_e^* = x_e^* - a_e^*$, $a^* = (a_e^*)_{e\in E}$, and $\xi^* = (\xi_e^*)_{e\in E}$. Then ξ^* belongs to the polytope defined by

$$\mathcal{P} \;=\; \left\{ \xi \in \mathbb{R}^E \,:\, 0 \le \xi \le \sigma \;, \quad A\xi \;=\; d \right\} \;,$$

where $\sigma = (\sigma_e)_{e\in E}$ and d are given by

$$\sigma_e \;=\; \begin{cases} 1 & , \text{ if } a_e^* < x_e^* \\ 0 & , \text{ if } a_e^* = x_e^* \end{cases} \quad , \qquad d \;=\; b - Aa^* \;.$$

Note that d has integer components. Since A is totally unimodular, each vertex of the polytope $\mathcal{P}$ is an integer vector, (cf. Schrijver (1999, Theorem 19.3)), and thus a vector of zeros and ones. The function $\xi \longmapsto F(a^* + \xi)$ is *linear* on $\mathcal{P}$ and therefore attains its minimum at some vertex of $\mathcal{P}$. So there is a vector ξ^{**} of zeros and ones in $\mathcal{P}$ such that

$$F(a^* + \xi^{**}) \;\le\; F(a^* + \xi^*) \;=\; F(x^*) \;.$$

Hence $x^{**} \;=\; a^* + \xi^{**}$ is also an optimal solution to problem (4.1)-(4.2) and x^{**} is an integer vector. □

Consider the conjugate function of the piecewise linear convex function f_e,

$$g_e(t) \;=\; \max\left\{ \xi\, t \;-\; f_e(\xi) \;:\; 0 \le \xi \le \mu_e \right\} \;=\; \max\left\{ n\, t \;-\; f_e(n) \;:\; n = 0, 1, \ldots, \mu_e \right\} \quad \forall\, t \in \mathbb{R}\,. \tag{4.3}$$

More explicitely: g_e is a convex piecewise-linear function on $\mathbb{R}$ whose breakpoints are the slopes of f_e and whose slopes are the breakpoints of f_e (cf. Fourer (1985), Section 4),

$$g_e(t) \;=\; n\, t \;-\; f_e(n) \;, \quad \text{if } \; t \in I_e(n) \text{ and } n \in \{0, 1, \ldots, \mu_e\} \,, \tag{4.4}$$

$$\text{with intervals} \quad I_e(n) \;=\; \begin{cases} (-\infty\,,\, \Delta f_e(1)\,] & , \text{ if } n = 0 \\ [\,\Delta f_e(n)\,,\, \Delta f_e(n+1)\,] & , \text{ if } 1 \le n < \mu_e \\ [\,\Delta f_e(\mu_e)\,,\, \infty) & , \text{ if } n = \mu_e \end{cases} \;. \tag{4.5}$$

The dual objective function is given by, (cf. Fourer (1985), Section 5),

$$G(\lambda) \;=\; b'\lambda \;-\; \sum_{e\in E} g_e(y_e) \;, \quad \text{where } y = (y_e)_{e\in E} = A'\lambda \;, \quad \forall\, \lambda \in \mathbb{R}^V \,, \tag{4.6}$$

and the *dual problem* is to maximize $G(\lambda)$ over $\lambda \in \mathbb{R}^V$. Note that $G(\lambda)$ depends on λ only through $y = A'\lambda \in \mathcal{R}(A')$, since $b = Ax^{(0)}$ for some $x^{(0)} \in \mathbb{R}^E$ and hence $b'\lambda = {x^{(0)}}'y$. Also, by (4.4), we may write G as

$$\begin{aligned} &G(\lambda) \;=\; (b - A\nu)'\lambda \;+\; F(\nu) \;, \quad \text{with} \quad \nu = (\nu_e)_{e\in E} \text{ such that} \\ &\nu_e \in \{0, 1, \ldots, \mu_e\} \text{ and } \; y_e \in I_e(\nu_e) \;\forall\; e \in E \;, \; (\text{where } \; y = A'\lambda)\,). \end{aligned} \tag{4.7}$$

Now, strong duality can directly be verified:

Theorem 4.2 *The minimum value* $\min F(x)$ *of the primal problem* (1.2)-(1.3) *equals the maximum value* $\max G(\lambda)$ *of the dual problem and that maximum value is attained.*
If x^* is a point satisfying (1.3) *and $\lambda^* \in \mathbb{R}^V$, then a necessary and sufficient condition for x^* to be an optimal solution to problem* (1.2)-(1.3) *and λ^* to be a maximizer of G is that $y^* = A'\lambda^*$ satisfies*

$$\Delta f_e(x_e^*) \;\le\; y_e^* \;\le\; \Delta f_e(x_e^* + 1) \quad \forall\, e \in E \;.$$

Proof. Let x be a feasible point to problem (1.2)-(1.3) and let $\lambda \in \mathbb{R}^V$, $y = A'\lambda$. By (4.3) and (4.4)-(4.5), for any $e \in E$,

$$g_e(y_e) \geq x_e y_e - f_e(x_e)$$

with equality if and only if $\Delta f_e(x_e) \leq y_e \leq \Delta f_e(x_e + 1)$. Hence, by (4.6),

$$G(\lambda) \leq b'\lambda - x'y + F(x)$$

with equality if and only if

$$\Delta f_e(x_e) \leq y_e \leq \Delta f_e(x_e + 1) \quad \forall\, e \in E\,. \tag{4.8}$$

But $x'y = x'A'\lambda = (Ax)'\lambda = b'\lambda$, and we have thus obtained: $G(\lambda) \leq F(x)$ with equality if and only if (4.8) holds. Together with Theorem 2.3 the result follows. □

The dual algorithm for maximizing $G(\lambda)$ to be established below utilizes that, by (4.7), the function $G(\lambda)$ is *linear* on each polyhedral subset

$$\Lambda(\nu) = \left\{ \lambda \in \mathbb{R}^V \,:\, (A'\lambda)_e \in I_e(\nu_e)\ \forall\, e \in E \right\},$$

for any fixed $\nu = (\nu_e)_{e \in E}$, $\nu_e \in \{0, 1, \ldots, \mu_e\}$, $(e \in E)$. Solving the linear program of maximizing $G(\lambda)$ over $\Lambda(\nu)$ for a fixed ν will produce a solution $\widehat{\lambda}$, with $\widehat{y}_e = (A'\widehat{\lambda})_e$ hitting the left or the right boundary of $I_e(\nu_e)$ for some (or several) $e \in E$. If $e \in E$ and $\widehat{y}_e$ equals the *left* boundary of $I_e(\nu_e)$, then we are free to replace ν_e by $\nu_e - 1$. If $e \in E$ and $\widehat{y}_e$ equals the *right* boundary of $I_e(\nu_e)$, then we are free to replace ν_e by $\nu_e + 1$. The goal is to assign these changes of the ν_e in such a way that the (integer) vector

$$\theta = \theta(\nu) = b - A\nu$$

decreases in its l^1-norm, $\delta(\theta) = \sum_{v \in V} |\theta_v|$. Then, by repeating the procedure, we will end up with a vector ν^* of integers $\nu_e^* \in \{0, 1, \ldots, \mu_e\}$, $(e \in E)$, such that $\theta^* = 0$, i.e., $A\nu^* = b$, and a vector $\lambda^* \in \Lambda(\nu^*)$. That is, ν^* is feasible for the primal problem (1.2)-(1.3) and $G(\lambda^*) = F(\nu^*)$, hence ν^* and λ^* are optimal solutions to the primal and the dual problem, resp. In fact, the goal can be achieved, in principle, as we show next. Moreover, it turns out that in each linear programming step (for fixed ν) it suffices to compute a *weak Pareto solution* $\widehat{\lambda}$ rather than an *optimal* solution to the linear program

$$\text{maximize} \quad \theta(\nu)'\lambda \quad \text{subject to} \quad \lambda \in \Lambda(\nu)\,.$$

By a *weak Pareto solution* we mean the following.

Definition 4.3 *Let $\theta = (\theta_v)_{v \in V} \in \mathbb{R}^V$ be a given nonzero vector and $\Lambda \subseteq \mathbb{R}^V$ a given nonempty subset. Consider the problem*

$$\text{maximize} \quad \theta'\lambda \quad \text{subject to} \quad \lambda \in \Lambda\,. \tag{4.9}$$

Define $V^+ = \{v \in V \,:\, \theta_v > 0\}$ and $V^- = \{v \in V \,:\, \theta_v < 0\}$. A point $\widehat{\lambda} = (\widehat{\lambda}_v)_{v \in V} \in \Lambda$ is said to be a weak Pareto solution to (4.9) iff there is no $\lambda = (\lambda_v)_{v \in V} \in \Lambda$ such that

$$\lambda_v > \widehat{\lambda}_v \ \ \forall\, v \in V^+ \quad \text{and} \quad \lambda_v < \widehat{\lambda}_v \ \ \forall\, v \in V^-\,.$$

□

Lemma 4.4 *Let $\theta = (\theta_v)_{v \in V} \in \mathbb{R}^V$ be a nonzero vector, and let $c_e \in \mathbb{R} \cup \{-\infty\}$, $d_e \in \mathbb{R} \cup \{+\infty\}$ with $c_e \leq d_e$ for all $e \in E$. Consider the linear program*

$$\text{maximize} \quad \theta'\lambda \quad \text{subject to} \quad c_e \leq y_e \leq d_e\ \forall\, e \in E\,, \quad \text{where } y = A'\lambda\,. \tag{4.10}$$

As in Definition 4.3 *we denote*

$$V^+ = \{v \in V \,:\, \theta_v > 0\}\ ,\quad V^- = \{v \in V \,:\, \theta_v < 0\}\ ,\ \text{and also}\ \ V^0 = \{v \in V \,:\, \theta_v = 0\}\ .$$

Let $\widehat{\lambda}$ *be a feasible point to* (4.10)*, and* $\widehat{y} = A'\widehat{\lambda}$*. Define* $E^= = \{e \in E \,:\, c_e = d_e\}$*, and*

$$E^+(\widehat{\lambda}) \,=\, \big\{e \in E \setminus E^= \,:\, \widehat{y}_e = d_e\big\}\ ,\quad E^-(\widehat{\lambda}) \,=\, \big\{e \in E \setminus E^= \,:\, \widehat{y}_e = c_e\big\}\ ,$$
$$\text{and}\quad E^0(\widehat{\lambda}) \,=\, \big\{\, e \in E \setminus E^= \,:\, c_e < \widehat{y}_e < d_e \big\}\ .$$

Then: $\widehat{\lambda}$ *is a weak Pareto solution to* (4.10) *if and only if there exists a vector* $\sigma = (\sigma_e)_{e\in E}$ *with components* $\sigma_e \in \{-1, 0, +1\}$*, (for all* $e \in E$*), and such that:*

$$\sigma_e \geq 0\ \forall\, e \in E^+(\widehat{\lambda})\ ,\quad \sigma_e \leq 0\ \forall\, e \in E^-(\widehat{\lambda})\ ,\quad \sigma_e = 0\ \forall\, e \in E^0(\widehat{\lambda})\ ;$$
$$\text{the vector}\ \ A\sigma \,=\, a \,=\, (a_v)_{v\in V}\ \text{is nonzero,}\ a_v \in \{-1, 0, +1\}\ \forall\, v \in V,$$
$$\text{and}\quad a_v \geq 0\ \forall\, v \in V^+\ ,\quad a_v \leq 0\ \forall\, v \in V^-\ ,\quad a_v = 0\ \forall\, v \in V^0\ .$$

Proof. The vector $\widehat{\lambda}$ is *not* a weak Pareto solution to (4.10) if and only if there exists a vector $\xi = (\xi_v)_{v\in V}$ such that, denoting $\eta = (\eta_e)_{e\in E} = A'\xi$,

$$\xi_v > 0\ \ \forall\, v \in V^+\ ,\quad \xi_v < 0\ \ \forall\, v \in V^-\ ,$$
$$\eta_e \leq 0\ \ \forall\, e \in E^+(\widehat{\lambda})\ ,\quad \eta_e \geq 0\ \ \forall\, e \in E^-(\widehat{\lambda})\ ,\quad \eta_e = 0\ \ \forall\, e \in E^=\ .$$

This can also be expressed by saying that $\widehat{\lambda}$ is not a weak Pareto solution to (4.10) if and only if the following condition (a) holds.

(a) *There exists a vector*

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \in \mathcal{R}\Big(\begin{bmatrix} I_V \\ A' \end{bmatrix}\Big)$$

such that $\xi_v \in (\,0\,,\,\infty)\ \ \forall\, v \in V^+\ ,\quad \xi_v \in (-\infty\,,\,0\,)\ \ \forall\, v \in V^-\ ,\quad \xi_v \in \mathbb{R}\ \ \forall\, v \in V^0\ ,$
$\eta_e \in (-\infty\,,\,0\,]\ \ \forall\, e \in E^+(\widehat{\lambda})\ ,\quad \eta_e \in [\,0\,,\,\infty)\ \ \forall\, e \in E^-(\widehat{\lambda})\ ,\quad \eta_e \in \{0\}\ \ \forall\, e \in E^=\ ,$
$\eta_e \in \mathbb{R}\ \ \forall\, e \in E^0(\widehat{\lambda})\ .$

So, by Theorem 22.6 in Rockafellar (1972), the vector $\widehat{\lambda}$ is a weak Pareto solution to (4.10) if and only if the alternative condition (b) holds.

(b) *There exists an elementary vector* $\begin{pmatrix} a \\ \omega \end{pmatrix}$ *of* $\mathcal{N}\Big(\big[I_V\,,\,A\big]\Big)$ *, where* $a = (a_v)_{v\in V}$ *and* $\omega = (\omega_e)_{e\in E}$ *, such that*

$$\sum_{v\in V^+} a_v\,(\,0\,,\,\infty)\ +\ \sum_{v\in V^-} a_v\,(-\infty\,,\,0\,)\ +\ \sum_{v\in V^0} a_v\,\mathbb{R}\ +\ \sum_{e\in E^+(\widehat{\lambda})} \omega_e\,(-\infty\,,\,0\,]$$
$$+\ \sum_{e\in E^-(\widehat{\lambda})} \omega_e\,[\,0\,,\,\infty)\ +\ \sum_{e\in E^=} \omega_e\,\{0\}\ +\ \sum_{e\in E^0(\widehat{\lambda})} \omega_e\,\mathbb{R} \qquad >\ 0\ . \tag{4.11}$$

This is converted into the format stated in the assertion. Namely, (4.11) means

$$a_v \geq 0\ \ \forall\, v \in V^+\ ,\quad a_v \leq 0\ \ \forall\, v \in V^-\ ,\quad a_v = 0\ \ \forall\, v \in V^0\ ,$$
$$\omega_e \leq 0\ \ \forall\, e \in E^+(\widehat{\lambda})\ ,\quad \omega_e \geq 0\ \ \forall\, e \in E^-(\widehat{\lambda})\ ,\quad \omega_e = 0\ \ \forall\, e \in E^0(\widehat{\lambda})\ ,$$
$$\text{and the vector}\ a = (a_v)_{v\in V}\ \text{is nonzero.}$$

Since A is totally unimodular, so is the matrix $\big[I_V\,,\,A\big]$, (cf. Schrijver (1999, p. 267)). Hence, by Lemma 2.1, in condition (b) the elementary vector $\begin{pmatrix} a \\ \omega \end{pmatrix}$ of $\mathcal{N}\Big(\big[I_V\,,\,A\big]\Big)$ can be chosen to have all its components in $\{-1, 0, +1\}$. Furthermore, by

$$0 \,=\, \big[I_V\,,\,A\big]\begin{pmatrix} a \\ \omega \end{pmatrix} \,=\, a \,+\, A\omega\ ,$$

and taking $\sigma = -\omega$, we have $a = A\sigma$. Now condition (b) emerges in the required format. $\square$

Remark. Below, we will mostly be concerned with a linear program (4.10) whose maximum value is *finite*, i.e., the feasible region of (4.10) is nonempty and the objective linear function is bounded above on that region. Then, necessarily, $\theta \in \mathcal{R}(A)$. For, suppose $\theta \notin \mathcal{R}(A)$. Then $\theta = \theta^{(1)} + \theta^{(2)}$ with $\theta^{(1)} \in \mathcal{R}(A)$ and $\theta^{(2)} \in \mathcal{N}(A')$, $\theta^{(2)} \neq 0$. Choose any feasible point λ to (4.10). Then, for an arbitrary scalar $t > 0$, the point $\lambda + t\theta^{(2)}$ is again feasible and

$$\theta' \, (\lambda + t\theta^{(2)}) \;=\; \theta'\lambda \,+\, t\,\theta^{(2)'}\theta^{(2)} \longrightarrow \infty \quad \text{for } t \to \infty\,,$$

which is a contradiction.

5 A conceptual dual algorithm

Suppose that we have have an algorithm, we call it an *Oracle Y*, which achieves the following.

Oracle Y

Let a problem (4.10) *be given (with $\theta \neq 0$) such that its maximum value is finite. Let a feasible point λ be given. Then Oracle Y returns a weak Pareto solution $\widehat{\lambda}$ to* (4.10) *and a vector $\sigma = (\sigma_e)_{e \in E}$ according to* Lemma 4.4*.*

By linear programming methods it should be possible to construct an Oracle Y with polynomially (in $\#E + \#V$) bounded running time. For vector and matrix apportionment problems specific Oracles Y will be described in Sections 6 and 7. However, the dual algorithm below (based on an Oracle Y) for solving the dual and the primal problem of Theorem 4.2 will call Oracle Y up to $\delta(b - A\nu^0)$ times, where ν^0 is determined by the starting point λ^0. So the method will benefit from a foregoing heuristics, as the alternating scaling algorithm in case of a matrix apportionment problem (see Section 8), which provides a starting point λ^0 such that the l^1-distance $\delta(b - A\nu^0)$ is small or moderate.

Conceptual dual algorithm, (needs an Oracle Y)

(o) Start with any $\lambda \in \mathbb{R}^V$. Let $y = (y_e)_{e \in E} = A'\lambda$. For each $e \in E$ compute a $\nu_e \in \{0, 1, \ldots, \mu_e\}$ such that $y_e \in I_e(\nu_e)$, and let $\nu = (\nu_e)_{e \in E}$ and $\theta = b - A\nu$.

(i) If $\theta = 0$ then λ and ν are optimal solutions to the dual and the primal problem, resp. Otherwise ($\theta \neq 0$) go to (ii).

(ii) Apply Oracle Y to problem (4.10) with c_e and d_e being the left and the right boundary point, resp., of $I_e(\nu_e)$, ($e \in E$). So we get a weak Pareto solution $\widehat{\lambda}$ to (4.10) and a vector $\sigma = (\sigma_e)_{e \in E}$ according to Lemma 4.4. Set $\widehat{y} = A'\widehat{\lambda}$, $\widehat{\nu} = \nu + \sigma$, and $\widehat{\theta} = b - A\widehat{\nu}$. By the properties of σ we have

$$\widehat{\nu}_e \in \{0, 1, \ldots, \mu_e\}\,, \quad \text{and} \quad \widehat{y}_e \in I_e(\widehat{\nu}_e) \quad \forall\, e \in E\,,$$

and moreover, since $\theta = b - A\nu$ and $\widehat{\theta} = \theta - a$, where $a = (a_v)_{v \in V} = A\sigma$:

$$\begin{aligned} \delta(\widehat{\theta}) \;&=\; \sum_{v \in V} |\widehat{\theta}_v| \;=\; \sum_{v \in V^+} (\theta_v - a_v) \;+\; \sum_{v \in V^-} (a_v - \theta_v) \\ &=\; \sum_{v \in V} |\theta_v| \;-\; \sum_{v \in V} |a_v| \;\leq\; \sum_{v \in V} |\theta_v| \;-\; 1 \;=\; \delta(\theta) - 1\,. \end{aligned}$$

Replace λ by $\widehat{\lambda}$, ν by $\widehat{\nu}$, θ by $\widehat{\theta}$ and go to step (i).

Since $\delta(\theta)$, the l^1-norm of the integer vector θ, is decreased each time by (ii) the algorithm will terminate after finitely many cycles with optimal solutions to the dual and the primal problem.

6 Vector apportionment problems

Let V be a one-point set, $E = \{1, \ldots, p\}$, where $p \geq 2$, and $A = [1, \ldots, 1]$, i.e., the primal problem reads as

$$\text{minimize} \quad F(x) = \sum_{j=1}^{p} f_j(x_j) \tag{6.1}$$

$$\text{subject to} \quad x = (x_1, \ldots, x_p)' \in \mathbb{Z}^p \ , \quad 0 \leq x \leq \mu \ , \quad \sum_{j=1}^{p} x_j = h \ , \tag{6.2}$$

where $\mu = (\mu_1, \ldots, \mu_p)'$ is a given positive integer vector and h, (the *house size*), is a given positive integer such that $\sum_{j=1}^{p} \mu_j \geq h$. Obviously, the elementary sign vectors z of $\mathcal{N}(A)$ are those having exactly one component equal to $+1$, exactly one component equal to -1, and the remaining components equal to zero. So conditions (ii) and (iii) of Theorem 2.3, characterizing the optimality of a feasible point x^*, say the same, namely:

$$\max_{1 \leq i \leq p} \Delta f_i(x_i^*) \ \leq \ \min_{1 \leq j \leq p} \Delta f_j(x_j^* + 1) \ ,$$

cp. Saaty (1970, p. 184). Let $c_j \in \mathbb{R} \cup \{-\infty\}$ and $d_j \in \mathbb{R} \cup \{\infty\}$ with $c_j \leq d_j$, $(1 \leq j \leq p)$, be given. An *Oracle X* which decides between alternatives (a) and (b) of Lemma 2.2 is easily established:

Oracle X

Compute $\max_{1 \leq i \leq p} c_i$ and $\min_{1 \leq j \leq p} d_j$; if the former does not exceed the latter then choose a real λ between the max and the min, and $y = (\lambda, \ldots \lambda)'$ satisfies (a) of Lemma 2.2. Otherwise, find an i_0 and a j_0 such that $c_{i_0} > d_{j_0}$; then the elementary sign vector z of $\mathcal{N}(A)$ with $z_{i_0} = -1$, $z_{j_0} = 1$, and $z_j = 0$ else, satisfies (b) of Lemma 2.2. □

The dual objective function G from Section 4 is a function of a scalar variable $\lambda \in \mathbb{R}$ and (4.7) rewrites as

$$G(\lambda) = \Big(h - \sum_{j=1}^{p} \nu_j\Big) \lambda + F(\nu) \ ,$$
$$\text{if } \lambda \in I_j(\nu_j) \text{ and } \nu_j \in \{0, 1, \ldots, \mu_j\} \, , \ \forall \ j = 1, \ldots, p \, .$$

An *Oracle Y* is simple to establish since θ and λ in (4.10) are scalars, and the linear program (4.10) becomes:

$$\text{maximize} \quad \theta \, \lambda \quad \text{s.t.} \quad c_j \leq \lambda \leq d_j \quad \forall \ j = 1, \ldots, p \, ,$$

where θ is a given nonzero real number and c_j, d_j, $(1 \leq j \leq p)$, are as above. Assume that the maximum value of that linear program is finite, i.e.,

$$\max_{1 \leq i \leq p} c_i \ \leq \ \min_{1 \leq j \leq p} d_j \ ,$$
$$\min_{1 \leq j \leq p} d_j \ < \ +\infty \ \text{ if } \theta > 0 \, , \quad \text{and} \quad \max_{1 \leq i \leq p} c_i \ > \ -\infty \ \text{ if } \theta < 0 \, .$$

Oracle Y

A weak Pareto solution is the same as an optimal solution, which is given by

$$\widehat{\lambda} = \begin{cases} d_{j_0} = \min_j d_j & , \text{if } \theta > 0 \\ c_{i_0} = \max_i c_i & , \text{if } \theta < 0 \end{cases} \quad ,$$

and a vector $\sigma = (\sigma_1, \ldots, \sigma_p)'$ according to Lemma 4.4 is given by

$$\sigma_{j_0} = +1 \text{ and } \sigma_j = 0 \;\forall\; j \neq j_0 \;, \quad \text{in case } \theta > 0\,,$$
$$\sigma_{i_0} = -1 \text{ and } \sigma_j = 0 \;\forall\; j \neq i_0 \;, \quad \text{in case } \theta < 0\,.$$

□

The resulting dual algorithm was studied by Happacher and Pukelsheim (1996, p. 378) and Dorfleitner and Klein (1999), and implemented in the Java program Bazi (`www.uni-augsburg.de/bazi`). A favourable choice of the initial value for λ was suggested by Happacher and Pukelsheim (2000, p. 154).

7 Matrix apportionment problems

Let $V = \{R_1, \ldots, R_k, C_1, \ldots, C_\ell\}$ a set of $k + \ell$ elements, where $k \geq 2$ and $\ell \geq 2$, and let E be a given nonempty subset of the set of all (ordered) pairs (i, j), $(1 \leq i \leq k,\ 1 \leq j \leq \ell)$. That is, (V, E) constitutes a bipartite (undirected) graph. Let $A = (a_{v,e})_{v \in V, e \in E}$ be its vertex-edge incidence matrix, whose entries $a_{v,e}$ are defined by (1.7). Let $b = (r_1, \ldots, r_k, c_1, \ldots, c_\ell)'$ and $\mu = (\mu_{i,j})_{(i,j) \in E}$ be given (column) vectors of positive integers r_i, c_j, and $\mu_{i,j}$, such that the feasible region (1.8) is nonempty, (which implies, of course, that $\sum_{i=1}^k r_i = \sum_{j=1}^\ell c_j = h$, the *house size*). The elementary sign vectors $z = (z_{i,j})_{(i,j) \in E}$ of $\mathcal{N}(A)$ correspond to the elementary cycles in the bipartite graph (V, E), (cf. Rockafellar (1972, p. 204)). Therefore we will call those vectors z *elementary cycle vectors*, the precise definition of which is as follows. A vector $z = (z_{i,j})_{(i,j) \in E}$ is an elementary cycle vector iff there are an integer $n \geq 2$, pairwise distinct $i_0, i_1, \ldots, i_{n-1} \in \{1, \ldots, k\}$, and pairwise distinct $j_1, \ldots, j_n \in \{1, \ldots, \ell\}$ such that, with $i_n := i_0$ and some $s \in \{\pm 1\}$, one has

$$(i_m, j_{m+1}) \in E \;\; (0 \leq m \leq n-1)\;, \quad (i_m, j_m) \in E \;\; (1 \leq m \leq n)\;, \quad \text{and}$$
$$z_{i,j} = \begin{cases} \;\; s & , \text{if } i = i_m,\ j = j_{m+1},\ 0 \leq m \leq n-1 \\ -s & , \text{if } i = i_m,\ j = j_m\ ,\ \ 1 \leq m \leq n \\ \;\; 0 & , \text{else} \end{cases} \qquad \forall\, (i,j) \in E\,. \tag{7.1}$$

Here we write vectors $\lambda \in \mathbb{R}^V$ as

$$\lambda = (\alpha', \beta')'\;, \quad \text{where } \alpha = (\alpha_1, \ldots, \alpha_k)' \in \mathbb{R}^k \text{ and } \beta = (\beta_1, \ldots, \beta_\ell)' \in \mathbb{R}^\ell\;.$$

The linear subspace $\mathcal{R}(A')$ of $\mathbb{R}^E$ consists of all vectors $y = (y_{i,j})_{(i,j) \in E}$ such that

$$y_{i,j} = \alpha_i + \beta_j \;\; \forall\, (i,j) \in E \quad \text{for some } \alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell \in \mathbb{R}.$$

Let $c_{i,j} \in \mathbb{R} \cup \{-\infty\}$ and $d_{i,j} \in \mathbb{R} \cup \{+\infty\}$ with $c_{i,j} \leq d_{i,j}$, for all $(i, j) \in E$, be given. The alternatives (a) and (b*) of Lemma 2.2 rewrite as follows.

(a) There exist real numbers $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_\ell$ such that

$$c_{i,j} \;\leq\; \alpha_i + \beta_j \;\leq\; d_{i,j} \quad \forall\, (i,j) \in E\;.$$

(b*) There exists an elementary cycle vector $z = (z_{i,j})_{(i,j) \in E}$ such that

$$\sum_{(i,j) \in E^-(z)} c_{i,j} \;>\; \sum_{(i,j) \in E^+(z)} d_{i,j}\;,$$

where $E^+(z) = \{(i,j) \in E : z_{i,j} = +1\}$ and $E^-(z) = \{(i,j) \in E : z_{i,j} = -1\}$.

The *Oracle X* described next is an adapted version of the *Compatible Tension Algorithm* from graph theory, (cf. Berge (1991, pp. 94-96)).

Oracle X

Given: $c_{i,j} \in \mathbb{R} \cup \{-\infty\}$ and $d_{i,j} \in \mathbb{R} \cup \{+\infty\}$ with $c_{i,j} \leq d_{i,j}$, for all $(i,j) \in E$.

(o) Start with any $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell \in \mathbb{R}$ such that $\alpha_i + \beta_j \leq d_{i,j} \ \forall\ (i,j) \in E$. Let $y = (\alpha_i + \beta_j)_{(i,j)\in E}$.

(i) Consider the set of *noncompatible components* of y,

$$E_{\text{nc}}(y) = \left\{ (i,j) \in E \, : \, \alpha_i + \beta_j < c_{i,j} \right\} .$$

If $E_{\text{nc}}(y) = \emptyset$ then y satisfies alternative (a).
Otherwise, choose an $(i_0, j_0) \in E_{\text{nc}}(y)$ and go to (ii).

(ii) Apply the following labelling process to the elements of $V = \{R_1, \ldots, R_k, C_1, \ldots, C_\ell\}$, where, after the initial step (L0), the steps (L1) and (L2) are cycled through until R_{i_0} is labelled or no further labelling is possible.

- (L0) Label C_{j_0}.
- (L1) If $(i,j) \in E$ such that C_j is labelled, R_i is unlabelled, and $\alpha_i + \beta_j = d_{i,j}$ then R_i is labelled and gets the label C_j.
- (L2) If $(i,j) \in E$ such that R_i is labelled, C_j is unlabelled, and $\alpha_i + \beta_j \leq c_{i,j}$, then C_j is labelled and gets the label R_i.

Let $I = \{ i \, : \, R_i \text{ is labelled} \}$ and $\overline{I} = \{1, \ldots, k\} \setminus I$;
$J = \{ j \, : \, C_j \text{ is labelled} \}$ and $\overline{J} = \{1, \ldots, \ell\} \setminus J$.

If $i_0 \in I$, i.e., R_{i_0} is labelled, then go to (iii). Otherwise, i.e., R_{i_0} is unlabelled, then go to (iv).

(iii) (if $i_0 \in I$). Backtracking from R_{i_0} to C_{j_0} according to labels yields a finite sequence

$$i_0 \, , \, j_1 \, , \, i_1 \, , \, j_2 \, , \, \ldots \, , \, i_{n-1} \, , \, j_n = j_0 \, ,$$

for some $n \geq 2$, pairwise distinct $i_0, i_1, \ldots, i_{n-1} \in I$, pairwise distinct $j_1, \ldots, j_n \in J$, and such that R_{i_m} is labelled by $C_{j_{m+1}}$ $(0 \leq m \leq n-1)$ and C_{j_m} is labelled by R_{i_m} $(1 \leq m \leq n-1)$. That is, we have $(i_m, j_{m+1}) \in E$ $(0 \leq m \leq n-1)$, $(i_m, j_m) \in E$ $(1 \leq m \leq n-1)$, and

$$\alpha_{i_m} + \beta_{j_{m+1}} = d_{i_m, j_{m+1}} \ \ (0 \leq m \leq n-1) \, , \quad \alpha_{i_m} + \beta_{j_m} \leq c_{i_m, j_m} \ \ (1 \leq m \leq n-1) \, ;$$

we also have $(i_0, j_0) \in E$ and $\alpha_{i_0} + \beta_{j_0} < c_{i_0, j_0}$. Define the elementary cycle vector $z = \big(z_{i,j}\big)_{(i,j)\in E}$ by (7.1) with $s = +1$ (and $i_n := i_0$). Then,

$$\sum_{(i,j)\in E^-(z)} c_{i,j} = \sum_{m=1}^{n} c_{i_m, j_m} > \sum_{m=1}^{n} \big(\alpha_{i_m} + \beta_{j_m} \big) \, , \quad \text{(note: } i_n = i_0, \, j_n = j_0\text{)},$$

$$\sum_{(i,j)\in E^+(z)} d_{i,j} = \sum_{m=0}^{n-1} d_{i_m, j_{m+1}} = \sum_{m=0}^{n-1} \big(\alpha_{i_m} + \beta_{j_{m+1}} \big) = \sum_{m=1}^{n} \big(\alpha_{i_m} + \beta_{j_m} \big) \, ,$$

and hence

$$\sum_{(i,j)\in E^-(z)} c_{i,j} > \sum_{(i,j)\in E^+(z)} d_{i,j} \, .$$

So the elementary cycle vector z satisfies alternative (b*).

(iv) (if $i_0 \in \overline{I}$). By (L1) and (L2) of the labelling process from (ii) we have, for $(i,j) \in E$,

$$\alpha_i + \beta_j < d_{i,j} \ \text{ if } i \in \overline{I} \, , \ j \in J \, , \quad \text{and} \quad \alpha_i + \beta_j > c_{i,j} \ \text{ if } i \in I \, , \ j \in \overline{J} \, .$$

Define $\varepsilon_1 = \min\{ d_{i,j} - (\alpha_i + \beta_j) : (i,j) \in E,\ i \in \overline{I},\ j \in J\}$ and $\varepsilon_2 = \min\{ \alpha_i + \beta_j - c_{i,j} : (i,j) \in E,\ i \in I,\ j \in \overline{J}\}$, with the usual conventions $+\infty - \gamma = +\infty$, $\gamma - (-\infty) = +\infty$, (for any real number γ) and $\min \emptyset = +\infty$. Clearly, $0 < \varepsilon_1, \varepsilon_2 \le +\infty$. If $\varepsilon_1 < +\infty$ or $\varepsilon_2 < +\infty$ then define $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, Otherwise (if $\varepsilon_1 = \varepsilon_2 = +\infty$) define $\varepsilon = c_{i_0,j_0} - (\alpha_{i_0} + \beta_{j_0})$. Define for all $1 \le i \le k$, $1 \le j \le \ell$,

$$\widetilde{\alpha}_i = \begin{cases} \alpha_i - \varepsilon & \text{, if } i \in I \\ \alpha_i & \text{, if } i \in \overline{I} \end{cases}, \quad \widetilde{\beta}_j = \begin{cases} \beta_j + \varepsilon & \text{, if } j \in J \\ \beta_j & \text{, if } j \in \overline{J} \end{cases},$$

and $\widetilde{y} = (\widetilde{\alpha}_i + \widetilde{\beta}_j)_{(i,j)\in E}$. Then, for all $(i,j) \in E$,

$$\widetilde{\alpha}_i + \widetilde{\beta}_j = \begin{cases} \alpha_i + \beta_j - \varepsilon & \text{, if } (i,j) \in I \times \overline{J} \\ \alpha_i + \beta_j + \varepsilon & \text{, if } (i,j) \in \overline{I} \times J \\ \alpha_i + \beta_j & \text{, else} \end{cases}.$$

So, by the choice of ε, the vector $\widetilde{y} = (\widetilde{\alpha}_i + \widetilde{\beta}_j)_{(i,j)\in E}$ again satisfies $\widetilde{\alpha}_i + \widetilde{\beta}_j \le d_{i,j}\ \forall\ (i,j) \in E$ and, moreover,

$$E_{\text{nc}}(\widetilde{y}) = \{(i,j) \in E : \widetilde{\alpha}_i + \widetilde{\beta}_j < c_{i,j}\} \subseteq E_{\text{nc}}(y) .$$

If $(i_0, j_0) \notin E_{\text{nc}}(\widetilde{y})$ then replace the α_i, the β_j, and y by the $\widetilde{\alpha}_i$, the $\widetilde{\beta}_j$, and $\widetilde{y}$, resp., and return to step (i). Otherwise (if $(i_0, j_0) \in E_{\text{nc}}(\widetilde{y})$) then replace the α_i, the β_j, and y by the $\widetilde{\alpha}_i$, the $\widetilde{\beta}_j$, and $\widetilde{y}$, resp., and return to step (ii), (note that the labelling process in (ii) needs not be started afresh, but the labelling obtained previously may be kept and additional labelling occurs due to the construction of the new point $\widetilde{y}$). □

Note: A rough analysis shows that Oracle X has running time $O\Big((k + \ell)\, (\#E)^2 \Big)$. □

Let us consider the dual problem. The objective function G from (4.7) turns into

$$G(\alpha, \beta) = \sum_{i=1}^{k} (r_i - \nu_{i,+})\alpha_i + \sum_{j=1}^{\ell} (c_j - \nu_{+,j})\beta_j + F(\nu) ,$$
$$\text{if } \alpha_i + \beta_j \in I_{i,j}(\nu_{i,j}) \text{ and } \nu_{i,j} \in \{0, 1, \dots, \mu_{i,j}\}\ \forall\ (i,j) \in E , \tag{7.2}$$

where for $\nu = (\nu_{i,j})_{(i,j)\in E}$ we have denoted $\nu_{i,+} = \sum_{j\,:\,(i,j)\in E} \nu_{i,j}$ and $\nu_{+,j} = \sum_{i\,:\,(i,j)\in E} \nu_{i,j}$, and the intervals $I_{i,j}(\nu_{i,j})$ are from (4.5). For establishing an *Oracle Y*, we firstly describe the weak Pareto solutions to a linear program (4.10) for the present situation.

Lemma 7.1 *Let $\theta = (\phi', \psi')' \in \mathbb{R}^{k+\ell}$, $\theta \ne 0$, where $\phi = (\phi_1, \dots, \phi_k)' \in \mathbb{R}^k$ and $\psi = (\psi_1, \dots, \psi_\ell)' \in \mathbb{R}^\ell$, and let $c_{i,j} \in \mathbb{R} \cup \{-\infty\}$ and $d_{i,j} \in \mathbb{R} \cup \{+\infty\}$ with $c_{i,j} \le d_{i,j}$ for all $(i,j) \in E$. Consider the linear program in the variable $\lambda = (\alpha', \beta')' \in \mathbb{R}^{k+\ell}$,*

$$\text{maximize} \quad \theta'\lambda = \phi'\alpha + \psi'\beta \quad \text{subject to} \quad c_{i,j} \le \alpha_i + \beta_j \le d_{i,j}\ \forall\ (i,j) \in E .$$

Define $I^+ = \{i : \phi_i > 0\}$, $I^- = \{i : \phi_i < 0\}$, $J^+ = \{j : \psi_j > 0\}$, and $J^- = \{j : \psi_j < 0\}$. Let $\widehat{\lambda} = (\widehat{\alpha}', \widehat{\beta}')'$ be a feasible point to the linear program. Define a directed graph $\mathcal{D}(\widehat{\lambda})$ with vertex set $V = \{R_1, \dots, R_k, C_1, \dots, C_\ell\}$ and whose arcs are given as follows:
There is an arc with initial point R_i and end point C_j iff $(i,j) \in E$ and $\widehat{\alpha}_i + \widehat{\beta}_j = d_{i,j}$;
there is an arc with initial point C_j and end point R_i iff $(i,j) \in E$ and $\widehat{\alpha}_i + \widehat{\beta}_j = c_{i,j}$.
Then: $\widehat{\lambda}$ is a weak Pareto solution to the linear program if and only if in $\mathcal{D}(\widehat{\lambda})$ there is a directed path from some vertex of $\{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ to some vertex of $\{R_i : i \in I^-\} \cup \{C_j : j \in J^+\}$.

Proof. **1.** Assume that $\widehat{\lambda}$ is a weak Pareto solution. Suppose that there does not exist a pair v, w of vertices $v \in \{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ and $w \in \{R_i : i \in I^-\} \cup \{C_j : j \in J^+\}$ such that there is a directed path in $\mathcal{D}(\widehat{\lambda})$ from v to w. Consider the subset V_1 of all vertices $w \in V$ such that $w \in \{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ <u>or</u> there exists a directed path in $\mathcal{D}(\widehat{\lambda})$ from some vertex $v \in \{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ to w. So, in particular, $R_i \notin V_1$ for all $i \in I^-$ and $C_j \notin V_1$ for all $j \in J^+$. Moreover, we have:

$$\begin{aligned}&\text{If } (i,j) \in E,\ R_i \in V_1, \text{ and } C_j \notin V_1 \text{ then } \widehat{\alpha}_i + \widehat{\beta}_j < d_{i,j};\\ &\text{if } (i,j) \in E,\ R_i \notin V_1, \text{ and } C_j \in V_1 \text{ then } \widehat{\alpha}_i + \widehat{\beta}_j > c_{i,j}.\end{aligned}$$

So we can choose a positive real ε such that

$$\begin{aligned}\varepsilon &\le d_{i,j} - (\widehat{\alpha}_i + \widehat{\beta}_j) \quad \text{for all } (i,j) \in E \text{ with } R_i \in V_1 \text{ and } C_j \notin V_1,\\ \varepsilon &\le \widehat{\alpha}_i + \widehat{\beta}_j - c_{i,j} \quad \text{for all } (i,j) \in E \text{ with } R_i \notin V_1 \text{ and } C_j \in V_1.\end{aligned}$$

Define a new point $\lambda = (\alpha', \beta')' \in \mathbb{R}^{k+\ell}$ by

$$\alpha_i = \begin{cases} \widehat{\alpha}_i + \varepsilon/2 & \text{, if } R_i \in V_1 \\ \widehat{\alpha}_i - \varepsilon/2 & \text{, else} \end{cases} \quad , \quad \beta_j = \begin{cases} \widehat{\beta}_j - \varepsilon/2 & \text{, if } C_j \in V_1 \\ \widehat{\beta}_j + \varepsilon/2 & \text{, else} \end{cases} .$$

Then, by the choice of ε, the point λ is feasible to the linear program. Moreover, consider the *positive* components of the coefficient vector θ which are ϕ_i for $i \in I^+$ and ψ_j for $j \in J^+$, and consider the *negative* components of θ which are ϕ_i for $i \in I^-$ and ψ_j for $j \in J^-$. If $i \in I^+$ then $R_i \in V_1$ and hence $\alpha_i = \widehat{\alpha}_i + \varepsilon/2 > \widehat{\alpha}_i$; if $j \in J^+$ then $C_j \notin V_1$ and hence $\beta_j = \widehat{\beta}_j + \varepsilon/2 > \widehat{\beta}_j$; if $i \in I^-$ then $R_i \notin V_1$ and hence $\alpha_i = \widehat{\alpha}_i - \varepsilon/2 < \widehat{\alpha}_i$; if $j \in J^-$ then $C_j \in V_1$ and hence $\beta_j = \widetilde{\beta}_j - \varepsilon/2 < \widehat{\beta}_j$. This shows that the point $\widehat{\lambda}$ is *not* a weak Pareto solution, contradicting the assumption.

2. Assume that there exist $v \in \{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ and $w \in \{R_i : i \in I^-\} \cup \{C_j : j \in J^+\}$ and a directed path in $\mathcal{D}(\widehat{\lambda})$ from v to w. We distinguish the four cases:
(i) $v = R_p$ and $w = R_q$ for some $p \in I^+$ and $q \in I^-$;
(ii) $v = R_p$ and $w = C_q$ for some $p \in I^+$ and $q \in J^+$;
(iii) $v = C_p$ and $w = R_q$ for some $p \in J^-$ and $q \in I^-$;
(iv) $v = C_p$ and $w = C_q$ for some $p \in J^-$ and $q \in J^+$.
In either cases we can conclude that $\widehat{\lambda}$ is a weak Pareto solution; examplarily we show this for case (i), while the other three cases are handled analogously.
<u>Case (i):</u> There is a finite sequence $R_{i_1}, C_{j_1}, R_{i_2}, \ldots, C_{j_{n-1}}, R_{i_n}$, where $n \geq 2$, such that $i_1 = p$, $i_n = q$, and there is an arc in $\mathcal{D}(\widehat{\lambda})$ from each vertex of the sequence (except the last) to its successor. That is,

$$\begin{aligned}&(i_m, j_m) \in E \quad \text{and} \quad \widehat{\alpha}_{i_m} + \widehat{\beta}_{j_m} = d_{i_m, j_m} \ , \quad 1 \le m \le n-1 \, ,\\ &(i_{m+1}, j_m) \in E \quad \text{and} \quad \widehat{\alpha}_{i_{m+1}} + \widehat{\beta}_{j_m} = c_{i_{m+1}, j_m} \ , \quad 1 \le m \le n-1 \, .\end{aligned}$$

For any point $\lambda = (\alpha', \beta')'$ feasible to the linear program we have thus

$$\alpha_p - \alpha_q = \sum_{m=1}^{n-1} (\alpha_{i_m} + \beta_{j_m}) - \sum_{m=1}^{n-1} (\alpha_{i_{m+1}} + \beta_{j_m}) \le \sum_{m=1}^{n-1} d_{i_m, j_m} - \sum_{m=1}^{n-1} c_{i_{m+1}, j_m} \ ,$$

and equality holds for $\lambda = \widehat{\lambda}$. So there cannot exist a feasible point λ such that $\alpha_p > \widehat{\alpha}_p$ and $\alpha_q < \widehat{\alpha_q}$, and therefore $\widehat{\lambda}$ is a weak Pareto solution, (recall that ϕ_p is a positive component of

θ and ϕ_q is a negative component of θ). □

The *Oracle Y* given next achieves the following.
Given a linear program as in Lemma 7.1 *which is assumed to have a finite maximum value, and given a feasible point* $\lambda = (\alpha', \beta')'$ *to that linear program. Then, a weak Pareto solution* $\widehat{\lambda} = (\widehat{\alpha}', \widehat{\beta}')'$ *to the linear program* and *a directed path in* $\mathcal{D}(\widehat{\lambda})$ *according to* Lemma 7.1 *is found. From this a vector* $\sigma = (\sigma_{i,j})_{(i,j)\in E}$ *according to* Lemma 4.4. *is obtained by*

$$\sigma_{i,j} = \begin{cases} +1 & \text{, if the path contains an arc from } R_i \text{ to } C_j \\ -1 & \text{, if the path contains an arc from } C_j \text{ to } R_i \\ 0 & \text{, else} \end{cases} \quad \forall\ (i,j) \in E\,. \tag{7.3}$$

Remark
If the maximum value of the linear program from Lemma 7.1 is finite, then:

$$I^+ \cup J^- \neq \emptyset \quad \text{and} \quad I^- \cup J^+ \neq \emptyset\ , \tag{7.4}$$

which can be seen as follows. By the final remark in Section 4, $\theta = (\phi', \psi')' \in \mathcal{R}(A)$, and hence $\sum_{i=1}^k \phi_i = \sum_{j=1}^\ell \psi_j$, which we can rewrite as

$$\sum_{I^+} \phi_i - \sum_{J^-} \psi_j = \sum_{J^+} \psi_j - \sum_{I^-} \phi_i\ ,$$

and that value is positive since $\theta \neq 0$. Hence (7.4) follows. □

Oracle Y

Given the linear program from Lemma 7.1 which is assumed to have a finite maximum value, and given a feasible point $\lambda = (\alpha', \beta')'$ to that program. Let I^+, I^-, J^+, and J^- be defined as in Lemma 7.1

(i) Apply the following labelling process to the elements of $V = \{R_1, \ldots, R_k, C_1, \ldots, C_\ell\}$, where, after the initial step (L0), the steps (L1) and (L2) are cycled through until some R_{i^*} with $i^* \in I^-$ is labelled, or some C_{j^*} with $j^* \in J^+$ is labelled, or no further labelling is possible.

- (L0) Label all R_i for $i \in I^+$ and label all C_j for $j \in J^-$.
- (L1) If $(i,j) \in E$ is such that R_i is labelled, C_j is unlabelled, and $\alpha_i + \beta_j = d_{i,j}$ then C_j is labelled and gets the label R_i.
- (L2) If $(i,j) \in E$ is such that C_j is labelled, R_i is unlabelled, and $\alpha_i + \beta_j = c_{i,j}$ then R_i is labelled and gets the label C_j.

Let $I = \{i : R_i \text{ is labelled}\}$ and $\overline{I} = \{1, \ldots, k\} \setminus I$;
$J = \{j : C_j \text{ is labelled}\}$ and $\overline{J} = \{1, \ldots, \ell\} \setminus J$.
If $I \cap I^- \neq \emptyset$ or $J \cap J^+ \neq \emptyset$ then go to (ii). Otherwise go to (iii).

(ii) (if $I \cap I^- \neq \emptyset$ or $J \cap J^+ \neq \emptyset$)
Backtracking from some R_{i^*} with $i^* \in I \cap I^-$ or from some C_{j^*} with $j^* \in J \cap J^+$ according to labels yields a directed path in $\mathcal{D}(\lambda)$ from some vertex $v \in \{R_i : i \in I^+\} \cup \{C_j : j \in J^-\}$ to that vertex $w = R_{i^*}$ or $w = C_{j^*}$. By Lemma 7.1, λ is a weak Pareto solution. Choose σ by (7.3).

(iii) (if $I \cap I^- = J \cap J^+ = \emptyset$). By (L1) and (L2) from (ii) we have:

$$\text{If } (i,j) \in E,\ i \in I,\ j \in \overline{J} \text{ then } \alpha_i + \beta_j < d_{i.j}\,;$$
$$\text{if } (i,j) \in E,\ i \in \overline{I},\ j \in J \text{ then } \alpha_i + \beta_j > c_{i.j}\,.$$

Let $\varepsilon_1 = \min\{ d_{i,j} - (\alpha_i + \beta_j) : (i,j) \in E, \ i \in I, \ j \in \overline{J} \}$ and
$\varepsilon_2 = \min\{ \alpha_i + \beta_j - c_{i,j} : (i,j) \in E, \ i \in \overline{I}, \ j \in J \}$,
where the usual conventions $+\infty - t = +\infty$, $t - (-\infty) = +\infty$ (for a real t), and $\min \emptyset = +\infty$ are used. Clearly, $0 < \varepsilon_1 \leq +\infty$ and $0 < \varepsilon_2 \leq +\infty$. Not both of them are equal to $+\infty$ which can be seen as follows. Suppose that $\varepsilon_1 = \varepsilon_2 = +\infty$. For an arbitrary real $\varepsilon > 0$ define $\widetilde{\alpha} = (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_k)$ and $\widetilde{\beta} = (\widetilde{\beta}_1, \ldots, \widetilde{\beta}_\ell)$ by

$$\widetilde{\alpha}_i = \begin{cases} \alpha_i + \varepsilon & \text{, if } i \in I \\ \alpha_i & \text{, if } i \in \overline{I} \end{cases} , \quad \text{and} \quad \widetilde{\beta}_j = \begin{cases} \beta_j - \varepsilon & \text{, if } j \in J \\ \beta_j & \text{, if } j \in \overline{J} \end{cases} . \tag{7.5}$$

Then

$$\widetilde{\alpha}_i + \widetilde{\beta}_j = \begin{cases} \alpha_i + \beta_j + \varepsilon & \text{, if } (i,j) \in I \times \overline{J} \\ \alpha_i + \beta_j - \varepsilon & \text{, if } (i,j) \in \overline{I} \times J \\ \alpha_i + \beta_j & \text{, else} \end{cases} \quad \forall\ (i,j) \in E , \tag{7.6}$$

and thus $\widetilde{\lambda} = (\widetilde{\alpha}', \widetilde{\beta}')'$ is again feasible to the linear program. Now,

$$\theta' \widetilde{\lambda} = \phi' \widetilde{\alpha} + \psi' \widetilde{\beta} = \theta' \lambda + \varepsilon \Big(\sum_I \phi_i - \sum_J \psi_j \Big) .$$

Since $I \cap I^- = \emptyset$ and $J \cap J^+ = \emptyset$, (and $I^+ \subseteq I$, $J^- \subseteq J$), we have

$$\sum_I \phi_i - \sum_J \psi_j = \sum_{I^+} \phi_i - \sum_{J^-} \psi_j ,$$

and that value is positive by (7.4). So $\theta' \widetilde{\lambda}$ gets arbitrarily large by choosing ε arbitrarily large, which is a contradiction. Thus, $\varepsilon_1 < +\infty$ or $\varepsilon_2 < +\infty$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, and again define $\widetilde{\alpha}$ and $\widetilde{\beta}$ by (7.5) which entails (7.6). By the choice of ε the point $\widetilde{\lambda} = (\widetilde{\alpha}', \widetilde{\beta}')'$ is again feasible to the linear program and, moreover, there is an $(i,j) \in E$ with $i \in I$, $j \in \overline{J}$, and $\widetilde{\alpha}_i + \widetilde{\beta}_j = d_{i,j}$, <u>or</u> there is an $(i,j) \in E$ with $i \in \overline{I}$, $j \in J$, and $\widetilde{\alpha}_i + \widetilde{\beta}_j = c_{i,j}$. Replace α and β by $\widetilde{\alpha}$ and $\widetilde{\beta}$, resp., and return to step (i), (note that the labelling process needs not be started afresh, but the previously obtained labelling may be kept and additional labelling occurs). □

Note: A rough analysis shows that Oracle Y has running time $O\Big((k + \ell)\, \#E \Big)$. □

The Oracle Y and the resulting dual algorithm include (and generalize) the method for biproportional rounding of matrices of Balinski and Demange (1989, pp. 205ff.), see also Balinski and Rachev (1997, pp. 20ff.), and Rote and Zachariasen (2007).

8 Dual alternating scaling algorithm

Let us consider still another approach for matrix apportionment problems, to maximize the dual objective function $G(\alpha, \beta)$ from (7.2) over $(\alpha, \beta) \in \mathbb{R}^{k+\ell}$. The approach is simple as well as tempting: Use the alternating maximization procedure, i.e., maximize first over α for a fixed β, then maximize over β while keeping the before obtained α fixed, and so on. The name "alternating *scaling* algorithm" comes from biproportional rounding in its original *multiplicative* formulation, (cf. Gaffke and Pukelsheim (2007)), which includes the variables α_i and β_j via *multipliers* $\rho_i = \exp(\alpha_i)$ and $\gamma_j = \exp(\beta_j)$.

As we will show next, each maximization "half-step" consists in solving k or ℓ, resp., *vector* apportionment problems and their duals as discussed in Section 6. However, the function G is nondifferentiable, and thus the sequence of points (α, β) generated might not converge to a maximizer of G, (cf. Bazaraa, Sherali, Shetty (1993, pp. 285-287)). In fact, we shall demonstrate by example that the alternating maximization procedure may stall at a nonoptimal point $(\alpha^{(0)}, \beta^{(0)})$. Despite this deficiency, the method can be used as a first optimization part to approach the optimum, then followed by the dual algorithm from Sections 5 and 7.

Let us examine the half-steps of the alternating procedure in detail. We restrict attention to a first half-step, a second half-step is analogous. Let $\beta = (\beta_1, \ldots, \beta_\ell)' \in \mathbb{R}^\ell$ be considered fixed and consider $G(\alpha, \beta)$ from (7.2) as a function of $\alpha = (\alpha_1, \ldots, \alpha_k)' \in \mathbb{R}^k$. For each $i \in \{1, \ldots, k\}$, we denote $E(i) = \{j : (i,j) \in E\}$ which is nonempty since the feasible region (1.8) is assumed to be nonempty. Writing

$$\sum_{j=1}^{\ell} (c_j - \nu_{+,j})\beta_j \;=\; c'\beta \;-\; \sum_{i=1}^{k} \sum_{j \in E(i)} \nu_{i,j}\beta_j \;,$$

and observing the definition (4.5) of the intervals $I_{i,j}(\nu_{i,j})$, we can rewrite (7.2) as

$$G(\alpha, \beta) \;=\; c'\beta \;+\; \sum_{i=1}^{k} \Big[\Big(r_i - \sum_{j \in E(i)} \nu_{i,j} \Big) \alpha_i \;+\; \sum_{j \in E(i)} \big(f_{i,j}(\nu_{i,j}) - \nu_{i,j}\beta_j \big) \Big] \;,$$
$$\text{if} \;\; \Delta f_{i,j}(\nu_{i,j}) - \beta_j \le \alpha_i \le \Delta f_{i,j}(\nu_{i,j} + 1) - \beta_j \;\; \text{and} \;\; \nu_{i,j} \in \{0, 1, \ldots, \mu_{i,j}\} \;\; \forall\, (i,j) \in E \,.$$

Introducing the functions

$$f_{i,j,\beta}(t) \;=\; f_{i,j}(t) \;-\; \beta_j t \;, \quad t \in [\,0\,,\, \mu_{i,j}] \;,\; (i,j) \in E \,,$$

we have $\Delta f_{i,j,\beta}(n) \;=\; \Delta f_{i,j}(n) - \beta_j$ for all $n = 0, 1, \ldots, \mu_{i,j} + 1$, and thus

$$G(\alpha, \beta) \;=\; c'\beta \;+\; \sum_{i=1}^{k} G_{i,\beta}(\alpha_i) \;, \quad \text{where for each } i = 1, \ldots, k \;:$$
$$G_{i,\beta}(\alpha_i) \;=\; \Big(r_i - \sum_{j \in E(i)} \nu_{i,j} \Big) \alpha_i \;+\; \sum_{j \in E(i)} f_{i,j,\beta}(\nu_{i,j}) \;,$$
$$\text{if} \;\; \Delta f_{i,j,\beta}(\nu_{i,j}) \le \alpha_i \le \Delta f_{i,j,\beta}(\nu_{i,j} + 1) \,,\; \nu_{i,j} \in \{0, 1, \ldots, \mu_{i,j}\} \quad \forall\, j \in E(i).$$

We see, firstly, that maximizing $G(\alpha, \beta)$ over $\alpha \in \mathbb{R}^k$ can be done by maximizing separately for each $i = 1, \ldots, k$ the function $G_{i,\beta}(\alpha_i)$ over $\alpha_i \in \mathbb{R}$, and secondly, in view of Section 6, that for each i the function $G_{i,\beta}$ is just the dual objective function to the vector apportionment problem,

$$\text{minimize} \quad F_{i,\beta}(x_i) \;=\; \sum_{j \in E(i)} f_{i,j,\beta}(x_{i,j})$$
$$\text{subject to} \quad x_i = (x_{i,j})_{j \in E(i)} \in \mathbb{Z}^{E(i)} \;,\quad 0 \le x_{i,j} \le \mu_{i,j} \;\forall\, j \in E(i) \;,\; \sum_{j \in E(i)} x_{i,j} \;=\; r_i \;,$$

(note that i is considered fixed). So, solving each of the k vector apportionment problems, yields a maximizer α of $G(\,\cdot\,, \beta)$ along with an $x = (x_{i,j})_{(i,j) \in E} \in \mathbb{Z}^E$ satisfying $0 \le x \le \mu$, one half of the equality restrictions, i.e., $x_{i,+} = r_i$ for all i, and

$$\Delta f_{i,j}(x_{i,j}) \;\le\; \alpha_i + \beta_j \;\le \Delta f_{i,j}(x_{i,j} + 1) \quad \forall\, (i,j) \in E \,. \tag{8.1}$$

Analogously, a second half step of maximizing $G(\alpha, \beta)$ over $\beta \in \mathbb{R}^\ell$ for a fixed α (obtained from the foregoing first half-step) means to solve ℓ vector apportionment problems. This yields a maximizing β and (another) integer point $x = (x_{i,j})_{i,j \in E} \in \mathbb{Z}^E$ satisfying $0 \le x \le \mu$, the other half of the equality restrictions, i.e., $x_{+,j} = c_j$ for all j, and (8.1). If it happens that the point x obtained in a half-step satisfies all the equality restrictions, $x_{i,+} = r_i$ for all i and $x_{+,j} = c_j$ for all j, then by Theorem 4.2 the point x and the point (α, β) at hand are optimal solutions to the primal and the dual problem, resp. However, as remarked above, that occurrence cannot be guaranteed in general. Next we will give a negative example, built by an artificial instance of biproportional rounding. For biproportional rounding of a *positive* matrix $W = (w_{i,j})_{\substack{1\le i\le k\\ 1\le j\le \ell}}$ the functions $f_{i.j}$, $1 \le i \le k$, $1 \le j \le \ell$, are such that

$$f_{i,j}(0) = 0\,, \quad \text{and} \quad \Delta f_{i,j}(n) = \log \frac{s(n)}{w_{i,j}}\,, \quad (n = 1, \ldots, \mu_{i,j})\,,$$

where $0 < s(1) < s(2) < s(3) < \ldots$ is a given sign-post sequence. The primal problem is to

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{i,j}(x_{i,j}) \\ &\text{subject to} \quad x_{i,j} \in \mathbb{Z}\,,\ x_{i,j} \ge 0 \ \ \forall\, i,j\,, \quad x_{i,+} = r_i \ \forall\, i\,, \quad x_{+,j} = c_j \ \forall j\,, \end{aligned}$$

where $r_1, \ldots, r_k$ and $c_1, \ldots, c_\ell$ are given positive integers. Note that here no upper bound μ occurs, i.e., μ may be any integer vector whose components are large enough to define *redundant* upper bounds.

Example 8.1
Let $k = \ell = 5$, $r_i = 1$ $(1 \le i \le 5)$, $c_j = 1$ $(1 \le j \le 5)$, and

$$W = \begin{pmatrix} s(1) & s(1) & \varepsilon & \varepsilon & \varepsilon \\ s(1) & s(1) & \varepsilon & \varepsilon & \varepsilon \\ s(1) & s(1) & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & s(1) & s(1) & s(1) \\ \varepsilon & \varepsilon & s(1) & s(1) & s(1) \end{pmatrix}\,, \quad \text{for some } 0 < \varepsilon < s(1)\,.$$

We start the dual alternating method with initial points $\alpha^{(0)} = \beta^{(0)} = 0 = (0,0,0,0,0)$. The first half-step of maximizing $G(\alpha, 0)$ over α has solutions $\alpha^{(1)}$ characterized by (8.1) with some nonnegative integer point $x = (x_{i,j})_{1\le i,j\le 5}$ such that $x_{i,+} = 1$ for all $i = 1, \ldots, 5$, i.e. x is a 0-1-matrix with precisely one 1 in each row. Now, (8.1) rewrites as

$$\max_{1\le j\le 5} \log \frac{s(x_{i,j})}{w_{i,j}} \le \alpha_i^{(1)} \le \min_{1\le j\le 5} \log \frac{s(x_{i,j}+1)}{w_{i,j}} \qquad \text{for all } i = 1, \ldots, 5\,,$$

where we define $s(0) = 0$ and $\log(0) = -\infty$. We conclude that $\alpha^{(1)} = 0$ (uniquely), and x is such that

$$x = \begin{pmatrix} B_1 & 0_{3\times 3} \\ 0_{2\times 2} & B_2 \end{pmatrix}\,, \tag{8.2}$$

with any 0-1-matrices B_1 (3×2) and B_2 (2×3) which have precisely one 1 entry in each row. The second half-step is thus to maximize $G(0, \beta)$ over β. The solutions $\beta^{(1)}$ are characterized by (8.1) with some nonnegative integer point $x = (x_{i,j})_{1\le i,j\le 5}$ such that $x_{+,j} = 1$ for all $j = 1, \ldots, 5$, i.e. x is a 0-1-matrix with precisely one 1 in each column. Since (8.1) rewrites as

$$\max_{1\le i\le 5} \log \frac{s(x_{i,j})}{w_{i,j}} \le \beta_j^{(1)} \le \min_{1\le i\le 5} \log \frac{s(x_{i,j}+1)}{w_{i,j}} \qquad \text{for all } j = 1, \ldots, 5\,,$$

we conclude that $\beta^{(1)} = 0$ (uniquely), and x is such that

$$x = \begin{pmatrix} C_1 & 0_{3\times 3} \\ 0_{2\times 2} & C_2 \end{pmatrix}\,, \tag{8.3}$$

with any 0-1-matrices C_1 (3×2) and C_2 (2×3) with precisely one 1 entry in each column.
So the procedure stalls at the point $(\alpha, \beta) = (0,0)$, which is nonoptimal: For any possible choices of B_1, B_2 the matrix x from (8.2) does not satisfy the column sums equations, and for any possible choice of C_1, C_2 the matrix x from (8.3). does not satisfy the row sums equations. So there is no feasible point x^* to the primal problem such that (8.1) holds for $(\alpha, \beta) = (0,0)$. Thus, by Theorem 4.2, the point $(0,0)$ is nonoptimal.
For example, if $s(1) = 0.5$ and $\varepsilon = 0.2$ (and $1 \le s(2) \le 2$), then an optimal dual solution (α^*, β^*) is given by

$$\alpha_1^* = \alpha_2^* = \alpha_3^* = \log(2.5)\ , \quad \alpha_4^* = \alpha_5^* = 0\ ,$$
$$\beta_1^* = \beta_2^* = \log(0.4)\ , \quad \beta_3^* = \beta_4^* = \beta_5^* = 0\ ,$$

and one optimal primal solution (among a total of 33 optimal solutions) is given by

$$x^* = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

which can easily be verified by checking the optimality condition (8.1) for the (feasible) pair x^* and (α^*, β^*). □

References

Balinski, M.L.; Demange, G. (1989): Algorithms for proportional matrices in reals and integers. *Mathematical Programming* **45**, 193-210.

Balinski, M.L.; Rachev, S.T. (1997): Rounding proportions: Methods of rounding. *Math. Scientist* **22**, 1-26.

Balinski, M. (2006): Apportionment: Uni- and bi-dimensional. In: Simeone, B.; Pukelsheim, F. (eds): *Mathematics and Democracy – Recent Advances in Voting Systems and Collectice Choice*, Berlin, 2006, pp. 43-54.

Bazaraa, M.S.; Sherali, H.D.; Shetty, C.M. (1993): *Nonlinear Programming – Theory and Algorithms*. Second Edition. Wiley, New York.

Berge, C. (1991): *Graphs*. Third revised Edition. Elsevier Inc., New York.

Dorfleitner, G.; Klein, T. (1999): Rounding with multiplier methods: An efficient algorithm and applications in statistics. *Statistical Papers* **40**, 143-157.

Fourer, R. (1985): A simplex algorithm for piecewise-linear programming I: Derivation and proof. *Mathematical Programming* **33**, 204-233.

Gaffke, N.; Pukelsheim, F. (2007): Divisor methods for proportional representation systems: An optimization approach to vector and matrix apportionment problems. (Submitted).

Graver, J.E. (1975): On the foundations of linear and integer linear programming. *Mathematical Programming* **8**, 207-226.

Happacher, M.; Pukelsheim, F. (1996): Rounding probabilities: Unbiased multipliers. *Statistics & Decisions* **14**, 373-382.

Happacher, M.; Pukelsheim, F. (2000): Rounding probabilities: Maximum probability and minimum complexity multipliers. *Journal of Statistical Planning and Inference* **85**, 145-158.

Hemmecke, R. (2003): Test sets for integer programs with $\mathbb{Z}$-convex objective. Typescript, UC Davis. e-print available from `http://front.math.ucdavis.edu/math.CO/0309154, 2003`

Hochbaum, D.S. (2005): Complexity and algorithms for convex network optimization and other nonlinear problems. *4OR* **3**, 171-216.

Hochbaum, D.S.; Shanthikumar, J.G. (1990): Convex separable optimization is not much harder that linear optimization. *Journal of the Association for Computing Machinery* **37**, 843-862.

Murota, K.; Saito, H.; Weismantel, R. (2004): Optimality criterion for a class of nonlinear integer programs. *Operations Research Letters* **32**, 468-472.

Rockafellar, R.T. (1972): *Convex Analysis.* Princeton Univ. Press, Princeton.

Rote, G.; Zachariasen, M. (2007): Matrix scaling by network flow. In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Proceedings in Applied Mathematics, Volume 125*, Philadelphia PA, 2007, pp. 848-854.

Saaty, T.L. (1970): *Optimization in Integers and Related Extremal Problems.* McGraw-Hill, New York.

Schrijver, A. (1999): *Theory of Linear and Integer Programming.* Wiley, New York.

Sun, J; Tsai, K.-H.; Qi, L. (1993): A simplex method for network programs with convex separable piecewise linear costs and its applications to stochastic transshipment problems. In: D.Z. Du and P.M. Pardalos (eds): *Network Optimization Problems.* World Scientific Publ.Co.

Te Riele, H.J.J. (1978): The proportional representation problem in the Second Chamber: an approach via minimal distances. *Statistica Neerlandica* **32**, 163-179.

Thépot, J. (1986): Scrutin Proportionnel et Optimization Mathématique. (In French). *R.A.I.R.O.* **20**, 123-138.