UNIVERSITÄT AUGSBURG
Universitätsbibliothek

OPEN ACCESS
OPUS AUGSBURG

Divisor methods for proportional representation systems: an optimization approach to vector and matrix apportionment problems

Norbert Gaffke, Friedrich Pukelsheim

Angaben zur Veröffentlichung / Publication details:

Gaffke, Norbert, and Friedrich Pukelsheim. 2008. "Divisor methods for proportional representation systems: an optimization approach to vector and matrix apportionment problems." *Mathematical Social Sciences* 56 (2): 166–84. https://doi.org/10.1016/j.mathsocsci.2008.01.004.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
Deutsches Urheberrecht
Weitere Informationen finden Sie unter: / For more information see:
https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/



Divisor methods for proportional representation systems: An optimization approach to vector and matrix apportionment problems

Norbert Gaffke[a], Friedrich Pukelsheim[b,*]

[a] *Institut für Mathematische Stochastik, Otto-von-Guericke-Universität, D-39016 Magdeburg, Germany*
[b] *Institut für Mathematik, Universität Augsburg, D-86135 Augsburg, Germany*

1. Introduction

Apportionment methods are procedures to allocate a fixed number of seats of a political body proportionally to a given set of weights, such as vote counts, population data, faction sizes, or the like. The total number of seats, the house size, is prespecified, and the point is that it must be dealt

* Corresponding author.
E-mail addresses: Gaffke@Mathematik.Uni-Magdeburg.De (N. Gaffke), Pukelsheim@Math.Uni-Augsburg.De (F. Pukelsheim).

out precisely. The seemingly unimportant issue of whether, due to a rounding effect of some sort or other, the last units are allocated here or there is, time and again, a matter of fierce political debate. After all, the units are seats, and a seat more or less may be decisive for a political group to command a majority or not.

The traditional instances of the problem may be called vector problems, since the input enters in the form of a weight vector. The first half of the seminal monograph of Balinski and Young (2001) tells the tale of two centuries of US apportionment history. The second half draws on the empirical evidence provided, and establishes a Theory of Apportionment analyzing the methods on the ground of their axiomatic properties.

Recent applications give rise to matrix problems, with input data in the form of a weight matrix v. Its entries v_{ij} signify the number of votes cast in district i for party j. The rows reflect a first, geographical subdivision of the population into electoral districts. The columns accommodate the parties and mirror a second, political (or economical, or social) partitioning. The house size is assumed to be broken down either way, providing district magnitudes as row marginals, and overall party seats as column marginals. The task is to find a two-way apportionment, with seat allocations x_{ij} summing row- and column-wise to the prespecified row and column marginals, and, as a whole, achieving some sort of proportionality relative to the input weights v_{ij}.

Appropriate procedures for solving a matrix apportionment problem are *biproportional scaling methods*. They operate in two steps. Firstly, the weights v_{ij} are scaled into $\rho_i v_{ij} \gamma_j$, so that the total closes in on the prespecified marginals. However, the scaling operation yields quantities that are continuous, not discrete. Therefore, secondly, the scaled weights are discretized by rounding them to the seat allocations x_{ij}. In a nutshell, the present paper identifies the row multipliers ρ_i and the column multipliers γ_j as dual variables, of a suitable primal–dual pairing of optimization problems.

In practical applications, the weights v_{ij} are often large and must be considerably scaled down. Since quoting tiny multipliers looks awkward, common sense suggests to switch to the inverses, divisors. For this reason much of the apportionment literature speaks of *divisor methods*, rather than *multiplier methods*. Both views have their merits, and are subsumed under our neutral heading "scaling methods". In the theoretical development, we prefer to work with multipliers. In the empirical examples, we quote divisors.

Our paper is organized as follows. Since the discretization step heavily relies on rounding rules, these are carefully introduced in Section 2. The intervals $[n-1, n]$ are equipped with *signposts* $s(n)$ such that a real value t below $s(n)$ is rounded down to $n-1$, above $s(n)$, up to n. We distinguish between pervious rounding rules, where small values may be rounded to zero, and impervious rounding rules, where this never happens. Section 3 outlines the vector apportionment problem. In Section 4, its solutions are characterized through a familiar max–min inequality. Section 5 proposes a novel optimization viewpoint, the primal goal function being the product of consecutive signposts. The variable of the dual problem is the multiplier that is instrumental for carrying out the rounding rule.

Section 6 sets the stage for matrix problems. In Section 7 we motivate the goal function through heuristically more accessible double-quotients. The max–min inequality of vector problems is replaced by a system of critical inequalities, each along a cycle with positive weights. Our main result is Theorem 7.1, linking the validity of all critical inequalities to the existence of row and column multipliers. The proof is based on a theorem of the alternative in Rockafellar (1970), and also follows from Theorem 4.1 in Golitschek et al. (1983). Section 8 verifies that the primal–dual relationship for vector problems carries over to matrix problems almost *verbatim*. Section 9 concludes the paper with a re-view of what has been achieved.

Biproportional divisor methods for matrix problems were first introduced by Balinski and Demange (1989a,b), see also Balinski and Rachev (1997). Pukelsheim and Schuhmacher (2004) report on how the *biproportional divisor method with standard rounding* became part of the 2003 Zurich electoral law. The practical use spurred interest in biproportional methods, see Balinski (2004, 2006), Bochsler (2005), Zachariassen and Zachariassen (2006), Balinski and Pukelsheim (2006, 2007) and Simeone and Pukelsheim (2006).

A quick two-step algorithm to calculate vector apportionments was proposed by Happacher and Pukelsheim (1996, 2000) and Dorfleitner and Klein (1999), and is implemented in the Java program Bazi (Pukelsheim, 2004). The two-step algorithm was extended to the *Alternating Scaling* (AS) algorithm to solve matrix problems, by treating rows and columns as sets of vector problems. Though for all practical purposes the AS algorithm works fine, there are examples where it stalls (Gaffke and Pukelsheim, 2008). Balinski and Demange (1989b) put forward a transportation-type algorithm, always converging, which the Bazi software names the *Tie and Transfer* (TT) algorithm. The complexity of the TT algorithm is studied by Rote and Zachariassen (2007). Extensive simulation experiments of Maier (2006) suggest that a hybrid version, with AS starting and TT finishing, is the fastest.

The AS algorithm is a discrete variant of the well-known (continuous) *Iterative Proportional Fitting* (IPF) procedure, also known as matrix raking, diagonal matrix scaling, iterative scaling procedure, or RAS method (Balinski and Pukelsheim, 2006). The complexity of the IPF procedure is investigated by Kalantari et al. (2008), who also provide an overview of prior complexity references. The IPF procedure has found a wealth of applications, such as contingency table analysis in statistics (Deming and Stephan, 1940; Stephan, 1942; Bishop et al., 1975), input–output analysis in economics (Bacharach, 1965, 1970), or calculation of doubly stochastic matrices (Marshall and Olkin, 1968).

2. Rounding rules

Scaling methods of apportionment are based on signpost sequences which, in turn, induce rounding rules. The idea is to place in an integer interval $[n-1, n]$ a dividing point $s(n)$ below which we round down to $n-1$, and above which we round up to n.

By definition, a sequence $s(1), s(2), \ldots$ is called a *signpost sequence* when the *signposts* $s(n)$ come to lie between the bounds $n-1$ and n,

$$s(n) \in [n-1, n] \quad \text{for all } n = 1, 2, \ldots,$$

such that it does not happen that out of two signposts one hits its upper bound and the other its lower bound: There are no two subscripts $n = 1, 2, \ldots$ and $m = 2, 3, \ldots$ fulfilling

$$s(n) = n \quad \text{and} \quad s(m) = m - 1.$$

A signpost sequence is called *pervious* when the first signpost is positive, $s(1) > 0$, and *impervious* when the first signpost is zero, $s(1) = 0$. All signpost sequences are taken to be preceded by $s(0) = 0$ (Balinski and Young, 2001, p. 120).

Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ denote the set of natural numbers including zero. Signpost sequences decompose the half-line $[0, \infty)$ into the intervals $[s(n), s(n+1))$, with $n \in \mathbb{N}$. For impervious signpost sequences the initial interval $[s(0), s(1)) = [0, 0)$ is empty and dispensible, yet it is included for notational convenience. The name "signpost" alludes to the telling figures in Balinski and Young (2001, pp. 63–65), the notation $s(\cdot)$ nicely going along with the German "Sprungstelle" and the French "seuil". The definition rules out that there are two signposts with

one hitting its upper bound, the other one, its lower bound. This guarantees that the ensuing apportionment method will be exact, see the Remark after Corollary 5.3. It also implies that every such sequence is strictly monotonic.

By definition, a signpost sequence induces a *rounding rule* $[\![\cdot]\!]$ from the nonnegative half-line $[0, \infty)$ into one-element or two-element subsets of integers through $[\![0]\!] := \{0\}$ and

$$[\![t]\!] := \{n \in \mathbb{N} \mid s(n) \le t \le s(n+1)\} \quad \text{for all } t > 0.$$

A rounding rule is called *pervious* or *impervious* accordingly as its generating signpost sequence is pervious or impervious.

Some remarks may elucidate the definition of rounding rules. For positive arguments $t > 0$, a rounding rule returns the singleton $\{n\}$ as long as t strictly lies between its neighboring signposts, $s(n) < t < s(n+1) \Rightarrow [\![t]\!] = \{n\}$. This is unambiguously expressed by saying that "t is rounded to n".

Ties, of the sort $t = s(n) > 0$, need more attention. In such cases t lies in two intervals, in $[s(n-1), s(n)]$, as well as in $[s(n), s(n+1)]$. Here the rounding rule yields a two-element set, $[\![t]\!] = \{n-1, n\}$, leaving it undecided whether to round down to $n-1$, or up to n.

The boundary value $t = 0$ is special, in always being rounded to the singleton $\{0\}$. The motto is "nil input, nil output". A pervious rounding rule rounds all small positive quantities below $s(1) > 0$ to zero, $[\![t]\!] = \{0\} \Leftrightarrow t < s(1)$. An impervious rounding rule returns zero if and only if the input is zero: $[\![t]\!] = \{0\} \Leftrightarrow t = 0$.

The most prominent example is *standard rounding* that places its signposts in the middle of each integer interval, $s(n) = n - 1/2$. Standard rounding is pervious, in that all values below $s(1) = 1/2$ are rounded to zero. Another pervious example is the identity sequence $s(n) = n$, for which the corresponding rounding rule truncates non-integer reals to their integer part, $n < t < n+1 \Rightarrow [\![t]\!] = \{n\}$. A third example is *geometric rounding* that is based on the signposts $s(n) = \sqrt{(n-1) \times n}$, the geometric mean of the bounds $n-1$ and n. Geometric rounding is impervious, since $s(1) = \sqrt{0 \times 1} = 0$.

3. Vector problems

The vector apportionment problem features ℓ "parties" $j = 1, \ldots \ell$, each coming with a nonnegative *weight* $v_j \ge 0$. The "parties" may be political parties weighted by their vote counts of an election, or they may be states of a union weighted by their populations.

In view of the matrix apportionment problem of Section 6 we admit the possibility of some weights being zero. Of course, it is assumed that the weights do not vanish simultaneously, $v := (v_1, \ldots, v_\ell) \ne 0$. We define the *support* of a vector as the set of indices where the components are nonzero, $\mathrm{supp}(v) := \{j | v_j \ne 0\}$. That is, the indices $j \le \ell$ with $v_j > 0$, referred to frequently in the following, are conveniently quoted as $j \in \mathrm{supp}(v)$.

An apportionment method maps a weight v_j into some integer $x_j \in \mathbb{N}$; we say that party j is apportioned x_j "seats". Parties with weight zero cannot possibly win a seat. Hence all apportionment results must satisfy the *nil rule*: $v_j = 0 \Rightarrow x_j = 0$, that is, $x_j \ge 1 \Rightarrow v_j > 0$. In terms of the support sets of the weight vector v and the apportionment vector $x := (x_1, \ldots, x_\ell) \in \mathbb{N}^\ell$ this means $\mathrm{supp}(x) \subseteq \mathrm{supp}(v)$. With an impervious rounding rule, parties with positive weights are in fact guaranteed representation, $v_j = 0 \Leftrightarrow x_j = 0$, that is, $x_j \ge 1 \Leftrightarrow v_j > 0$. This *representation warranty* is expressed through $\mathrm{supp}(x) = \mathrm{supp}(v)$. We code the two cases into

the definition

$$\mathbb{N}_v^\ell := \begin{cases} \left\{x \in \mathbb{N}^\ell \mid \operatorname{supp}(x) \subseteq \operatorname{supp}(v)\right\} & \text{in case } s(1) > 0, \\ \left\{x \in \mathbb{N}^\ell \mid \operatorname{supp}(x) = \operatorname{supp}(v)\right\} & \text{in case } s(1) = 0. \end{cases}$$

Let $h \in \mathbb{N}$ be a prespecified *house size* (size of a parliament, committee size). All seats together must exhaust the house size, $x_+ = h$, where we use the generic notation $x_+ := \sum_{j \le \ell} x_j$. A vector $x \in \mathbb{N}_v^\ell$ is called *feasible* when its components sum to the house size, $x_+ = h$. In the pervious case, $s(1) > 0$, feasible vectors exist always. In the impervious case, $s(1) = 0$, a feasible vector exists whenever the house size is large enough to provide every positive weight with at least one seat. Therefore we make the *house size assumption* that there are enough seats to go around, $h \ge h_{\min}$, setting $h_{\min} := 0$ in case $s(1) > 0$, and $h_{\min} := \#\operatorname{supp}(v)$ in case $s(1) = 0$. This ensures the existence of feasible apportionment vectors, for all house sizes $h \ge h_{\min}$.

The task is to identify feasible apportionment vectors x with components that are "in the same proportions" as are the components of the weight vector v. The task needs to be specified further before it is well-defined.

4. One-way proportionality

What is the meaning of a feasible apportionment vector x being proportional to the weight vector v? When two-positive weights $v_{j_1} > 0$ and $v_{j_2} > 0$ are mapped into seat numbers x_{j_1} and x_{j_2}, we express proportionality through the approximate equality

$$\frac{x_{j_1}}{v_{j_1}} \approx \frac{x_{j_2}}{v_{j_2}}.$$

In the decisions of the German Federal Constitutional Court the seat-to-vote ratio x_j/v_j signifies the *success value of a voter's ballot* cast for party j (Pukelsheim, 2000). The approximate equality demands that the success values of the ballots of any two voters ought to be equal.

Since all reasonable methods are homogeneous (Balinski and Young, 2001, p. 97), the apportionment results do not depend on whether the input data are normalized or not. The normalized quantities $\widehat{w}_j := x_j/x_+$ and $w_j := v_j/v_+$ comply with standard statistical terminology, in that the likelihood ratios $\widehat{w}_j/w_j$ view the distribution to be fitted, $\widehat{w}_j$, in multiples of the distribution that is given, w_j.

Now the signposts $s(n)$ and the induced rounding rule $[\![\cdot]\!]$ enter the scene. They provide the basic relations $s(x_j) \le x_j \le s(x_j + 1)$. We bound x_{j_1} from below by $s(x_{j_1})$, and x_{j_2} from above by $s(x_{j_2} + 1)$, thus delimiting the range within which we tolerate the approximate equality $x_{j_1}/v_{j_1} \approx x_{j_2}/v_{j_2}$ to vary:

$$\frac{s(x_{j_1})}{v_{j_1}} \le \frac{s(x_{j_2} + 1)}{v_{j_2}} \quad \text{for all } j_1, j_2 \in \operatorname{supp}(v).$$

These *critical inequalities* turn out to be specific enough to be operational. They are condensed into the important *max–min inequality*

$$\max_{j \in \operatorname{supp}(v)} \frac{s(x_j)}{v_j} \le \min_{j \in \operatorname{supp}(v)} \frac{s(x_j + 1)}{v_j}.$$

The right-hand side minimum of the max–min inequality is always positive, provided the apportionment vector x is feasible. The left maximum becomes positive as soon as the house size is strictly above the minimum, $h > h_{\min}$.

The two sides of the max–min inequality give rise to the *multiplier interval*

$$\left[\max_{j \in \operatorname{supp}(v)} \frac{s(x_j)}{v_j}, \min_{j \in \operatorname{supp}(v)} \frac{s(x_j+1)}{v_j} \right].$$

Every *multiplier* $\mu > 0$ in this interval satisfies $s(x_j) \le \mu v_j \le s(x_j+1)$. That is, we have

$$x_j \in [\![\mu v_j]\!] \quad \text{for all } j \le \ell.$$

In summary, the weights v_j are converted into the seats x_j in two steps. First, the common multiplier μ is used to scale the weights, $v_j \mapsto \mu v_j$. Second, the rounding rule is applied to map the scaled weights into integers, $\mu v_j \mapsto x_j \in [\![\mu v_j]\!]$. If the multiplier $\mu > 0$ is determined so as to achieve $x_+ = h$, then the resulting apportionment vector is feasible.

Since the rounding rule may result in two-element sets of the form $[\![\mu v_{j_1}]\!] = \{x_{j_1} - 1, x_{j_1}\}$ or $[\![\mu v_{j_2}]\!] = \{x_{j_2}, x_{j_2} + 1\}$, it is possible that there are several feasible vectors that obey the max–min inequality. The *apportionment set* they form is designated by

$$A_v(h) := \left\{ x \in \mathbb{N}^\ell \mid x_+ = h \text{ and } \exists \mu > 0 \; \forall j \le \ell : \; x_j \in [\![\mu v_j]\!] \right\}.$$

More generally, A is called the *apportionment method* based on the rounding rule $[\![\cdot]\!]$ (Balinski and Young, 2001, p. 99). An apportionment method is termed *pervious* when its underlying rounding rule is pervious, and *impervious* when the rounding rule is impervious.

Example. As an illustration we consider an example from the German Bundestag (Schindler, 1999, p. 2085). There are four parliamentary factions of sizes 234:193:48:44 who want to apportion a committee of size 13, using the apportionment method with standard rounding (Webster/Sainte-Laguë). With the Rule-of-Three divisor $\delta_0 = 519/13$ the weights are scaled into $234/\delta_0 = 5.861\cdots$ etc., to obtain the fair shares 5.86:4.83:1.20:1.10. Then standard rounding yields the apportionment vector 6:5:1:1 which is actually feasible for a committee of size 13. By Corollary 5.3, this is the only solution. With $v = (234, 193, 48, 44)$, we may write $A_v(13) = \{(6, 5, 1, 1)\}$.

In the example the specific multiplier $\mu_0 = h/v_+ = 1/\delta_0$ produces an apportionment vector x with component sum equal to h. This need not always be so, and further computational steps may be necessary to determine a suitable multiplier. In fact, our discussion leaves open whether such multipliers exist or not, and whether the apportionment set $A_v(h)$ is non-empty or empty. Existence of such solutions is an immediate consequence when turning the vector apportionment problem into an integer optimization problem.

5. The vector optimization problem

There is a long tradition of characterizing apportionment methods through optimality criteria, with the intention to justify the particular merits of the resulting method (Balinski and Young, 2001, pp. 102–105; Balinski and Ramírez Gonzales, 1999). Here our aim is complementary, to exhibit the structure that is common to all scaling methods.

Proportional representation poses a problem of scale, not of location, whence we rely on products, not on sums. The goal function $F_v : \mathbb{N}_v^\ell \to (0, \infty)$ is defined by

$$F_v(x) := \prod_{j \in \mathrm{supp}(v)} \prod_{n \le x_j, s(n) > 0} \frac{s(n)}{v_j}.$$

When $x_j = 0$ (and, in the impervious case, when $x_j = 1$), the inner set of indices $\{n \le x_j | s(n) > 0\}$ becomes empty; as usual, an empty product is defined to be unity. We now introduce the vector optimization problem (V) through

$$\begin{aligned} &\text{Minimize} \quad F_v(x), && \text{(V1)}\\ &\text{subject to} \quad x_+ = h, && \text{(V2)}\\ &\text{over the set } x \in \mathbb{N}_v^\ell. && \text{(V3)} \end{aligned}$$

Feasible vectors x are such that they fulfill (V2) and (V3). A feasible vector for which F_v attains the minimum is called *optimal*.

The identity signposts $s(n) = n$ generate the *divisor method with rounding down* associated with the names of Jefferson, D'Hondt, and Hagenbach–Bischoff. They constitute a singular, special case, in that the inverse of the goal function turns out to be equal, up to a constant, to the density function of a multinomial distribution with sample size h and cell probabilities $w_j = v_j / v_+$. That is, $1/F_v(x) \propto \prod_{j \le \ell} w_j^{x_j} / (x_j!)$. Carnal (1993) proved that the density is maximized by the Hagenbach–Bischoff apportionment. We show that the idea carries over to *all* divisor methods, when the approach is generalized as above.

In order to derive necessary and sufficient conditions for optimality, we consider a differential move away from a feasible vector x. The smallest possible transfer is just one unit, from some party j_1 to some other party j_2. The induced mapping $E[j_1, j_2]$ is called the *elementary transfer from j_1 to j_2*. Since the corresponding step in the matrix problem is a bit more involved, we take the space to present a formal definition.

To this end we introduce the $\ell \times 1$ *elementary contrast vector* $e[j_1, j_2]$, with entry j_1 equal to 1, entry j_2 equal to -1, and zeros elsewhere. Thus the component sum is zero. The mapping $E[j_1, j_2] : \mathbb{N}_v^\ell \to \mathbb{N}_v^\ell$ is now defined through

$$E[j_1, j_2](x) := x - e[j_1, j_2]$$

provided $v_{j_2} > 0$ (party j_2 is eligible to receive a seat), and either $x_{j_1} \ge 1$ and $s(1) > 0$ (party j_1 has enough seats to give one away, when the rounding rule is pervious), or else $x_{j_1} \ge 2$ and $s(1) = 0$ (same when the rounding rule is impervious). Otherwise we set $E[j_1, j_2](x) := x$. If x is feasible, so is its transform $E[j_1, j_2](x)$. An analysis of elementary transfers establishes the following.

Theorem 5.1 (*Optimality*). *The following three statements are equivalent, for every apportionment vector $x \in \mathbb{N}_v^\ell$ that is feasible:*

(1) *x is a member of the apportionment set $A_v(h)$.*
(2) *x satisfies the max–min inequality.*
(3) *x is an optimal solution of problem (V).*

Proof. The equivalence of (1) and (2) is an immediate consequence of the definition of the apportionment set $A_v(h)$. It remains to prove the equivalence of (2) and (3). Assuming (2), we

choose a multiplier $\mu > 0$ from the non-empty multiplier interval

$$\left[\max_{j \in \mathrm{supp}(v)} \frac{s(x_j)}{v_j}, \ \min_{j \in \mathrm{supp}(v)} \frac{s(x_j+1)}{v_j} \right].$$

For every vector $y \in \mathbb{N}_v^\ell$ and for every index $j \in \mathrm{supp}(v)$, signpost sequence monotonicity yields two implications:

$$x_j > y_j \Longrightarrow \frac{\prod\limits_{n \le x_j, s(n)>0} s(n)/v_j}{\prod\limits_{n \le y_j, s(n)>0} s(n)/v_j} = \prod_{n=y_j+1}^{x_j} \frac{s(n)}{v_j} \le \left(\frac{s(x_j)}{v_j} \right)^{x_j - y_j} \le \mu^{x_j - y_j},$$

$$x_j < y_j \Longrightarrow \frac{\prod\limits_{n \le x_j, s(n)>0} s(n)/v_j}{\prod\limits_{n \le y_j, s(n)>0} s(n)/v_j} = \frac{1}{\prod\limits_{n=x_j+1}^{y_j} \frac{s(n)}{v_j}} \le \frac{1}{\left(\frac{s(x_j+1)}{v_j} \right)^{y_j - x_j}} \le \mu^{x_j - y_j}.$$

[Note that imperviousness, forcing $y_j \ge 1$, evades the constellation $x_j = 1 > y_j = 0$, for which in the first equality the quotient of the two empty products equals unity and would *not* coincide with the right-hand side, which is $s(1)/v_j = 0$.] Thus $x_j \ne y_j$ permits the estimate $\prod_{n \le x_j, s(n)>0} \big(s(n)/v_j\big) \le \mu^{x_j - y_j} \prod_{n \le y_j, s(n)>0} \big(s(n)/v_j\big)$. Here equality obtains if and only if $x_j = y_j + 1$ and $\mu = s(x_j)/v_j$, or $x_j + 1 = y_j$ and $\mu = s(x_j+1)/v_j$. The estimate evidently holds with equality when $x_j = y_j$. If, in addition, y is feasible, then $x_+ = y_+ = h$, whence the product over all parties with positive weights yields $F_v(x) \le \mu^{x_+ - y_+} F_v(y) = F_v(y)$. Hence x is optimal, (3).

Finally, assume (3). Consider a party j_1 with $x_{j_1} \ge 1$ when the rounding rule is pervious, and with $x_{j_1} \ge 2$ when it is impervious, either option necessitating $v_{j_1} > 0$. Let j_2 be another party with $v_{j_2} > 0$. Optimality of x yields

$$F_v(x) \le F_v\left(E[j_1, j_2](x)\right) = \frac{s(x_{j_2}+1)/v_{j_2}}{s(x_{j_1})/v_{j_1}} F_v(x).$$

This leaves $s(x_{j_1})/v_{j_1} \le s(x_{j_2}+1)/v_{j_2}$. The inequality is trivially true for $x_{j_1} = 0$, as well as for $x_{j_1} = 1$ when the rounding rule is impervious. Hence (2) follows. □

The equivalence of (1) and (3) in Theorem 5.1 means that the apportionment set $A_v(h)$ of Section 4, and the set of optimal solutions of problem (V) coincide:

$$A_v(h) = \left\{ x \in \mathbb{N}_v^\ell \mid x_+ = h \text{ and } F_v(x) = \min_{y \in \mathbb{N}_v^\ell, y_+ = h} F_v(y) \right\}.$$

This proves that the optimality set $A_v(h)$ is always non-empty, under our house size assumption. The proof is appealing in its simplemindedness, just referring to the obvious that a finite set of numbers always contains a minimum. The question of whether there are multiple solutions is answered easily from the max–min inequality.

Corollary 5.2 (*Multiple Solutions*). *The following two statements are equivalent, for every optimal apportionment vector $x \in A_v(h)$:*

(1) *There exists a vector $y \in A_v(h)$ different from x.*
(2) *The max–min inequality for x holds with equality.*

Proof. Assume (1). As $x \neq y$, there exist indices j_1 and j_2 with $x_{j_1} > y_{j_1}$ and $x_{j_2} < y_{j_2}$. The equality $F_v(x) = F_v(y)$ implies $\mu = s(x_{j_1})/v_{j_1}$, as well as $\mu = s(x_{j_2}+1)/v_{j_2}$, as in the proof of Theorem 5.1. Hence we get

$$\min_{j\in\mathrm{supp}(v)} \frac{s(x_j+1)}{v_j} \leq \frac{s(x_{j_2}+1)}{v_{j_2}} = \mu = \frac{s(x_{j_1})}{v_{j_1}} \leq \max_{j\in\mathrm{supp}(v)} \frac{s(x_j)}{v_j}.$$

This yields equality in the max–min inequality, (2).

Conversely, assume (2), that is, $\max_{j\in\mathrm{supp}(v)} s(x_j)/v_j = \min_{j\in\mathrm{supp}(v)} s(x_j+1)/v_j$. We choose two indices $j_1, j_2 \in \mathrm{supp}(v)$ with $s(x_{j_1})/v_{j_1} = s(x_{j_2}+1)/v_{j_2}$. This necessitates $x_{j_1} \geq 1$ when the rounding rule is pervious, and $x_{j_1} \geq 2$ when it is impervious. Thus $y = E[j_1, j_2](x) \neq x$ is feasible, and the proof of Theorem 5.1 yields $F_v(y) = F_v(x)$. Therefore y is optimal as well, (1). □

Corollary 5.3 (*Uniqueness*). *The following two statements are equivalent, for every optimal apportionment vector $x \in A_v(h)$:*

(1) *The set $A_v(h)$ is a singleton, $A_v(h) = \{x\}$.*
(2) *The max–min inequality holds with strict inequality.*

Proof. (1) and (2) are the negations of the assertions in Corollary 5.2. □

Remark. An input weight vector v is called *exact* if it is a positive multiple of some feasible apportionment vector $x \in \mathbb{N}_v^\ell$, say $v = \epsilon x$ with $\epsilon > 0$. Such input cannot lead to equality in the max–min inequality. Otherwise, there are subscripts $j_1 \neq j_2$ with $s(x_{j_1})/(\epsilon x_{j_1}) = s(x_{j_2}+1)/(\epsilon x_{j_2})$, forcing equality in $1 \leq s(x_{j_2}+1)/x_{j_2} = s(x_{j_1})/x_{j_1} \leq 1$. That is, the signpost $s(x_{j_1})$ hits its upper bound x_{j_1} and $s(x_{j_2}+1)$ coincides with its lower bound x_{j_2}, thus violating the definition of signpost sequences. Since the max–min inequality is strict, the optimal apportionment vector x is unique, $A_{\epsilon x}(x_+) = \{x\}$.

The primal minimization problem (V) comes with a dual maximization problem (W), to which we turn next. Since the primal goal function F_v is a product, the Lagrange multiplier λ stemming from the restriction $x_+ = h$ enters in a multiplicative form. For the *dual variable* $\lambda > 0$ and for all $x \in \mathbb{N}_v^\ell$ with $x_+ = h$, we have $\lambda^{h-x_+} = 1$ and

$$F_v(x) = \lambda^{h-x_+} F_v(x) \geq \inf_{y\in\mathbb{N}_v^\ell} \lambda^{h-y_+} F_v(y) =: G_v(\lambda).$$

The infimum in the definition of $G_v(\lambda)$ is evaluated as follows.

Lemma 5.4 (*Dual Goal Function*). *For all $\lambda > 0$, we have $G_v(\lambda) = \lambda^{h-z_+} F_v(z)$ whenever the components of the vector $z = (z_1, \ldots, z_\ell)$ fulfill $z_j \in [\![\lambda v_j]\!]$ for all $j \leq \ell$.*

Proof. The expression $\lambda^{h-z_+} F_v(z)$ involves the products $\prod_{n\leq z_j, s(n)>0} s(n)/(\lambda v_j)$ which are well-defined even when z_j is tied. A tie $\lambda v_j = s(z_j) > 0$ offers two rounding options, $[\![\lambda v_j]\!] = \{z_j - 1, z_j\}$. But because of $s(z_j)/(\lambda v_j) = 1$ it does not matter whether the product extends to z_j, or stops at $z_j - 1$.

To calculate $G_v(\lambda)$, the infimum over the vectors y is evaluated via their components y_j. For pervious methods we have $\mathrm{supp}(y) \subseteq \mathrm{supp}(v)$, and get

$$G_v(\lambda) = \lambda^h \inf_{y\in\mathbb{N}_v^\ell} \left(\lambda^{-y_+} \prod_{j\in\mathrm{supp}(v)} \prod_{n=1}^{y_j} \frac{s(n)}{v_j} \right) = \lambda^h \prod_{j\in\mathrm{supp}(v)} \left(\inf_{y_j\in\mathbb{N}} \prod_{n=1}^{y_j} \frac{s(n)}{\lambda v_j} \right).$$

For impervious methods, $\operatorname{supp}(y) = \operatorname{supp}(v)$ implies

$$G_v(\lambda) = \lambda^h \inf_{y \in \mathbb{N}_v^\ell} \left(\lambda^{-y_+} \prod_{j \in \operatorname{supp}(v)} \prod_{n=2}^{y_j} \frac{s(n)}{v_j} \right) = \lambda^{h - \#\operatorname{supp}(v)} \prod_{j \in \operatorname{supp}(v)} \left(\inf_{0 \neq y_j \in \mathbb{N}} \prod_{n=2}^{y_j} \frac{s(n)}{\lambda v_j} \right).$$

Either case permits the estimate $\prod_{n \le z_j, s(n)>0} \big(s(n)/v_j\big) \le \lambda^{z_j - y_j} \prod_{n \le y_j, s(n)>0} \big(s(n)/v_j\big)$, as in the proof of Theorem 5.1. Hence the minimum of $\prod_{n \le y_j, s(n)>0} \big(s(n)/(\lambda v_j)\big)$ over y_j is given by $\prod_{n \le z_j, s(n)>0} \big(s(n)/(\lambda v_j)\big)$. Thus $G_v(\lambda) = \lambda^{h - z_+} F_v(z)$ finishes the proof. □

We are now in a position to introduce the dual problem (W):

$$\text{Maximize } G_v(\lambda), \tag{W1}$$

$$\text{subject to } \lambda \in (0, \infty). \tag{W2}$$

Primal and dual problems are linked together through a Duality Theorem.

Theorem 5.5 (*Duality*). *Problems (V) and (W) both have optimal solutions, and share the same optimal value. Moreover, for all apportionment vectors $x \in \mathbb{N}_v^\ell$ that are feasible and for all multipliers $\mu > 0$, the following three statements are equivalent:*

(1) *x is an optimal solution of (V), and μ is an optimal solution of (W).*
(2) *x satisfies the max–min inequality, and μ lies in the associated multiplier interval.*
(3) *For all $j \le \ell$ we have $x_j \in [\![\mu v_j]\!]$.*

Proof. The optimal values satisfy $\min_{y \in \mathbb{N}_v^\ell, y_+ = h} F_v(y) \ge \sup_{\lambda > 0} G_v(\lambda)$, by definition of G_v. Let $x \in A_v(h)$. Then there exists some $\mu > 0$ with $x_j \in [\![\mu v_j]\!]$, for all $j \le \ell$. Lemma 5.4 yields $F_v(x) = G_v(\mu)$. With Theorem 5.1, the assertions follow. □

6. Matrix problems

Now we start afresh, to round a two-way array of weights in the presence of two-way side conditions, to be called a *matrix problem*. In the language of electoral systems, the electoral region is subdivided into k electoral districts, where ℓ parties are competing for seats. In practice we are given integer weights $v_{ij} \in \mathbb{N}$, designating the counts of votes cast in district i for party j. A vanishing weight $v_{ij} = 0$ may occur when party j does not run in district i, which is quite legitimate.

For our theoretical development we more generally assume the weights $v_{ij} \ge 0$ to be nonnegative reals, arranged in a *weight matrix* $v \in [0, \infty)^{k \times \ell}$. Each row is assumed to contain at least one positive weight, and so is each column. The positive entries in v are referred to through the support set, $\operatorname{supp}(v) := \{(i, j) | v_{ij} \neq 0\}$.

The number of seats to be apportioned in district i to party j is denoted by x_{ij}, forming the *apportionment matrix* $x \in \mathbb{N}^{k \times \ell}$. Again we impose the *nil rule* $\operatorname{supp}(x) \subseteq \operatorname{supp}(v)$, and for impervious methods, the *representation warranty* $\operatorname{supp}(x) = \operatorname{supp}(v)$:

$$\mathbb{N}_v^{k \times \ell} := \begin{cases} \left\{ x \in \mathbb{N}^{k \times \ell} \mid \operatorname{supp}(x) \subseteq \operatorname{supp}(v) \right\} & \text{in case } s(1) > 0, \\ \left\{ x \in \mathbb{N}^{k \times \ell} \mid \operatorname{supp}(x) = \operatorname{supp}(v) \right\} & \text{in case } s(1) = 0, \end{cases}$$

The side conditions stipulate that district i has $r_i \geq 1$ seats to fill (row marginals), and that party j is apportioned $c_j \geq 1$ seats (column marginals). Of course, the marginals must be consistent in summing to the same house size, $r_+ = c_+ =: h$.

We focus on apportionment matrices $x \in \mathbb{N}_v^{k\times\ell}$ that are *feasible*, that is, for which row sums meet row marginals, $x_{i+} = r_i$, and column sums, column marginals, $x_{+j} = c_j$. This requirement is succinctly expressed through $x1_\ell = r$ and $x'1_k = c$, where the unity vector $1_\ell \in \mathbb{R}^\ell$ has all components equal to unity, the vectors $r := (r_1, \ldots, r_k)'$ and $c := (c_1, \ldots, c_\ell)'$ assemble the row and column marginals, and a prime denotes transposition.

A max-flow min-cut algorithm may be employed to efficiently test whether a problem admits feasible apportionment matrices (Joas, 2005). Then the task is one of finding feasible matrices x with entries reasonably proportional to the entries of the weight matrix v.

7. Two-way proportionality

Equal proportions between input weights and output seats are harder to achieve when proportionality must be observed in two directions. We briefly digress and present some plausibility arguments, using double-quotients, to motivate the critical inequalities to be introduced thereafter. In the context of electoral systems, double-quotients were first employed in the 1980 Bern electoral law (Carnal and Riedwyl, 1982, 1984).

Let x be a feasible apportionment matrix. Our starting point is to contemplate the meaning of the approximate equality of two double-quotients,

$$\frac{\frac{x_{i_1 j_1}}{v_{i_1 j_1}}}{\frac{x_{i_1 j_2}}{v_{i_1 j_2}}} \approx \frac{\frac{x_{i_2 j_1}}{v_{i_2 j_1}}}{\frac{x_{i_2 j_2}}{v_{i_2 j_2}}}.$$

This means that the success values of any voter's ballots cast for parties j_1 or j_2, relate to each other in approximately equal proportions in district i_1 (on the left-hand side), or in district i_2 (on the right-hand side).

That this is a promising view is confirmed by its symmetry in rows and columns. Straightforward cross-multiplication yields

$$\frac{\frac{x_{i_1 j_1}}{v_{i_1 j_1}}}{\frac{x_{i_2 j_1}}{v_{i_2 j_1}}} \approx \frac{\frac{x_{i_1 j_2}}{v_{i_1 j_2}}}{\frac{x_{i_2 j_2}}{v_{i_2 j_2}}}.$$

Thus the success values of any voter's ballots cast in districts i_1 or i_2, relate to each other in approximately equal proportions whether they favor party j_1 (left), or party j_2 (right).

Rather than dealing with double-quotients of success values, we cross-multiply and demand approximate equality of the resulting products,

$$\frac{x_{i_1 j_1}}{v_{i_1 j_1}} \frac{x_{i_2 j_2}}{v_{i_2 j_2}} \approx \frac{x_{i_1 j_2}}{v_{i_1 j_2}} \frac{x_{i_2 j_1}}{v_{i_2 j_1}}.$$

The products make sense provided the denominator weights are positive. It is insufficient, though, to only compare rectangular configurations originating at the intersection of two rows i_1 and i_2 and two columns j_1 and j_2.

Example. To illustrate the deficiency, consider a 3×3 weight matrix v with diagonal weights zero and off-diagonal weights positive,

$$v = \begin{pmatrix} 0 & v_{12} & v_{13} \\ v_{21} & 0 & v_{23} \\ v_{31} & v_{32} & 0 \end{pmatrix}.$$

Certainly there is a need for an appraisal of proportionality between the seat apportionments and the weights. Yet no rectangular configuration involves only positive weights! As a remedy, we pick a cyclic path of positive weights, $12 \to 13 \to 23 \to 21 \to 31 \to 32\ (\to 12)$. The approximate equality associated with this cycle is

$$\frac{x_{12}}{v_{12}}\frac{x_{23}}{v_{23}}\frac{x_{31}}{v_{31}} \approx \frac{x_{13}}{v_{13}}\frac{x_{21}}{v_{21}}\frac{x_{32}}{v_{32}}.$$

The cycle is determined by the row and column indices on the left, $i_{(3)} = (1, 2, 3)$ and $j_{(3)} = (2, 3, 1)$. The column indices on the right are a permutation $j^*_{(3)} = (3, 1, 2)$ of $j_{(3)}$.

Generally, a *cycle* is a succession of matrix entries where a move within rows alternates with a move within columns, of the form

$$i_1 j_1 \to i_1 j_2 \to i_2 j_2 \to i_2 j_3 \to \cdots \to i_{q-1} j_{q-1} \to i_{q-1} j_q \to i_q j_q \to i_q j_1\ (\to i_1 j_1),$$

visiting any row or column at most once. Thus a cycle is determined by a vector $i_{(q)}$ of $q \geq 2$ distinct row indices, and a vector $j_{(q)}$ of q distinct column indices:

$$i_{(q)} = \left(i_1, i_2, \ldots, i_{q-1}, i_q\right), \qquad j_{(q)} = \left(j_1, j_2, \ldots, j_{q-1}, j_q\right).$$

We use an asterisk to designate cyclic permutation of column indices,

$$j^*_{(q)} := \left(j_2, j_3, \ldots, j_q, j_1\right).$$

In component-wise notation, this means $j^*_p = j_{p+1}$ for $p < q$, and $j^*_q = j_1$.

We contend that such cycles permit an appropriate standard of comparison, namely

$$\prod_{p \leq q} \frac{x_{i_p j_p}}{v_{i_p j_p}} \approx \prod_{p \leq q} \frac{x_{i_p j^*_p}}{v_{i_p j^*_p}}.$$

We speak of a "cycle on $\mathrm{supp}(v)$" when the corresponding entries in the weight matrix v are all positive, $(i_p, j_p), (i_p, j^*_p) \in \mathrm{supp}(v)$ for all $p \leq q$. Again delimiting the range of variation of the approximate equalities by means of the basic relations $s(x_{ij}) \leq x_{ij} \leq s(x_{ij} + 1)$, we define the set of *critical inequalities* for the matrix problem to be

$$\prod_{p \leq q} \frac{s(x_{i_p j_p})}{v_{i_p j_p}} \leq \prod_{p \leq q} \frac{s(x_{i_p j^*_p} + 1)}{v_{i_p j^*_p}} \quad \text{for all cycles on } \mathrm{supp}(v).$$

While for vector problems the system of critical inequalities is condensed into the single max–min inequality, no such condensation is available for matrix problems. Yet multipliers that fit in-between the critical inequalities play an essential role. Their existence is derived from a theorem of the alternative.

Theorem 7.1 (*Multiplier Existence*). *Let the apportionment matrix $x \in \mathbb{N}_v^{k \times \ell}$ be feasible. Then x obeys the critical inequalities if and only if there exist "row multipliers" $\rho_1, \ldots, \rho_k > 0$ and "column multipliers" $\gamma_1, \ldots, \gamma_\ell > 0$ satisfying*

$$\frac{s\left(x_{ij}\right)}{v_{ij}} \leq \rho_i \gamma_j \leq \frac{s\left(x_{ij} + 1\right)}{v_{ij}} \quad \textit{for all } (i, j) \in \mathrm{supp}(v).$$

Proof. The proof of the converse part is a one-liner. With all multipliers given, we get, for any cycle on $\mathrm{supp}(v)$ with index vectors $i_{(q)}$ and $j_{(q)}$,

$$\prod_{p\le q}\frac{s\left(x_{i_pj_p}\right)}{v_{i_pj_p}}\le\left(\prod_{p\le q}\rho_{i_p}\right)\left(\prod_{p\le q}\gamma_{j_p}\right)=\left(\prod_{p\le q}\rho_{i_p}\right)\left(\prod_{p\le q}\gamma_{j_p^*}\right)\le\prod_{p\le q}\frac{s\left(x_{i_pj_p^*}+1\right)}{v_{i_pj_p^*}}.$$

The proof of the direct part is more demanding, as the existence of the multipliers needs to be established. To this end we linearize the assertion by taking logarithms:

$$\log\frac{s\left(x_{ij}\right)}{v_{ij}}\le\alpha_i+\beta_j\le\log\frac{s\left(x_{ij}+1\right)}{v_{ij}}\qquad\text{for all }(i,j)\in\mathrm{supp}(v),$$

setting $\log(0/v_{ij})=-\infty$. The existence of $\alpha_i\equiv\log\rho_i\in\mathbb{R}$ and $\beta_j\equiv\log\gamma_j\in\mathbb{R}$ is at issue.

For $i\le k$ and $j\le\ell$ we define the non-empty intervals I_{ij} through

$$I_{ij}:=\begin{cases}\left[\log\dfrac{s\left(x_{ij}\right)}{v_{ij}},\log\dfrac{s\left(x_{ij}+1\right)}{v_{ij}}\right] & \text{in case } v_{ij}>0 \text{ and } s(x_{ij})>0, \quad (1)\\ \left(-\infty,\log\dfrac{s\left(x_{ij}+1\right)}{v_{ij}}\right] & \text{in case } v_{ij}>0 \text{ and } s(x_{ij})=0, \quad (2)\\ (-\infty,\infty) & \text{in case } v_{ij}=0. \quad (3)\end{cases}\tag{I}$$

In the linear space $\mathbb{R}^{k\times\ell}$ of rectangular real matrices, let the subspace L consist of the $k\times\ell$ matrices z with entries $z_{ij}=\alpha_i+\beta_j$, where $\alpha_i,\beta_j\in\mathbb{R}$ are arbitrary. The existence of a matrix $z\in L$ with components $z_{ij}\in I_{ij}$ follows from alternative (a) of Theorem 22.6 in Rockafellar (1970), as soon as we exclude alternative (b).

Relative to the matrix inner product $\langle u,y\rangle:=\sum_{i\le k}\sum_{j\le\ell}u_{ij}y_{ij}$, the orthogonally complementary subspace $L^\perp$ of L consists of the $k\times\ell$ matrices y with all row sums and all column sums equal to zero. For any cycle with index vectors $i_{(q)}$ and $j_{(q)}$, we define the $k\times\ell$ *elementary cycle matrix* $e[i_{(q)};j_{(q)}]$ to have entries (i_p,j_p) equal to 1 and entries $e_{i_pj_p^*}$ equal to -1, for all $p\le q$, and zeros elsewhere. It is not hard to verify that the nonzero scalar multiples of the elementary cycle matrices furnish the elementary vectors of the subspace $L^\perp$, in the terminology of Rockafellar (1970, p. 203).

Alternative (b) of Rockafellar's Theorem 22.6 states that there exists some elementary cycle matrix $e[i_{(q)};j_{(q)}]$ such that every matrix u, with entries $u_{ij}\in I_{ij}$ for all $i\le k$ and $j\le\ell$, fulfills $\langle u,e[i_{(q)};j_{(q)}]\rangle>0$, that is, $\sum_{p\le q}u_{i_pj_p}>\sum_{p\le q}u_{i_pj_p^*}$. Thus the intervals $I_{i_pj_p}$ are seen to be bounded from below, and must originate from an interval in line (I1). The intervals I_{i_p,j_p^*} are bounded from above, and may stem from (I1) or (I2). In particular, the cycle involved has positive weights, and the values $u_{i_pj_p}:=\log s\left(x_{ij}\right)/v_{ij}$ and $u_{i_pj_p^*}:=\log s\left(x_{ij}+1\right)/v_{ij}$ are finite. This yields

$$\prod_{p\le q}\frac{s(x_{i_pj_p})}{v_{i_pj_p}}>\prod_{p\le q}\frac{s(x_{i_pj_p^*}+1)}{v_{i_pj_p^*}}.$$

Now the critical inequalities invalidate alternative (b), whence follows alternative (a). □

There is an alternate proof. Define A,B,C to be the $k\times\ell$ matrices with respective entries $a_{ij}=v_{ij}$, $b_{ij}=s(x_{ij})$, $c_{ij}=s(x_{ij}+1)$, and let $\le$ denote the component-wise matrix ordering. The inequalities in our Theorem 7.1 then take the form $B\le XAY\le C$, where X and Y

Table 1
Biproportional divisor method with standard rounding

	125	SP *47*	SVP *25*	FDP *22*	Greens *12*	CVP *9*	SenL *5*	AL *5*	*District divisor*
"1"	*12*	3516–4	1709–2	2413–3	1080–2	639–1	247–0	184–0	*800*
"2"	*16*	4264–6	1806–3	1062–2	860–2	539–1	339–1	503–1	*650*
"3"	*13*	3103–6	758–1	566–1	867–2	467–1	137–0	940–2	*530*
"4"	*10*	3626–4	1349–2	1487–2	956–1	471–1	359–0	280–0	*800*
"5"	*17*	4968–5	2423–2	4354–5	1939–2	968–1	485–1	411–1	*960*
"6"	*16*	3632–6	2724–5	1266–2	730–1	946–1	482–1	230–0	*590*
"7"	*12*	4103–5	2135–2	2066–3	885–1	647–1	446–0	363–0	*893*
"8"	*19*	4105–7	3333–5	1607–3	771–1	949–1	636–1	247–1	*603*
"9"	*10*	1970–4	1516–3	486–1	211–0	446–1	344–1	65–0	*500*
Party divisor		*1.02*	*1.01*	*0.9*	*0.87*	*1.07*	*1*	*0.815*	

The entries v–x show the vote counts v and seat numbers x. To obtain x, the votes v are divided by the associated district and party divisors, and then rounded. In district "1", party SP wins $v = 3516$ votes and gets $x = 4$ seats, since $v/(800 \times 1.02) = 4.3 \searrow 4$. The divisors (right and bottom, in italics) are such that the prespecified district seats and the overall party seats (left and top, in italics) are met exactly.

are diagonal matrices with the row and column multipliers on the diagonal. Thus Theorem 4.1 of Golitschek et al. (1983) yields our Theorem 7.1.

The inequalities $s(x_{ij}) \le \rho_i v_{ij} \gamma_j \le s(x_{ij}+1)$ show that the seat numbers x_{ij} result from a rounding operation, $x_{ij} \in [\![\rho_i v_{ij} \gamma_j]\!]$. We define the *apportionment set* through

$$A_v(r,c) := \{x \in \mathbb{N}^{k\times\ell} \mid x1_\ell = r \text{ and } x'1_k = c \text{ and } \exists \rho_1,\dots,\rho_k,\gamma_1,\dots,\gamma_\ell > 0\ \forall i \le k, j \le \ell :\ x_{ij} \in [\![\rho_i v_{ij} \gamma_j]\!]\}.$$

The set remains the same whether we evaluate the original weights v_{ij}, or the row normalized weights v_{ij}/v_{i+}, or the column normalized weights v_{ij}/v_{+j}, or the globally normalized weights v_{ij}/v_{++}, as is easily seen by adjusting row or column multipliers appropriately.

Example. The data in Table 1 are from the 2002 Zurich City Parliament election. As with all election data, large vote counts need to be scaled down to small seat numbers. It is then more convenient to use *divisors*, $1/\rho_i$ and $1/\gamma_j$, rather than multipliers. The house size is 125. The district magnitudes were specified prior to the election, proportionally to population counts. The overall party seats are obtained from evaluating the cumulative vote counts. Standard rounding is used (Webster/Sainte-Laguë).

8. The matrix optimization problem

The optimization formulation from Section 5 smoothly carries over from vector problems to matrix problems. The goal function $F_v : \mathbb{N}_v^{k\times\ell} \to (0,\infty)$ takes the form

$$F_v(x) := \prod_{(i,j)\in \mathrm{supp}(v)} \prod_{n \le x_{ij}, s(n)>0} \frac{s(n)}{v_{ij}},$$

and the matrix optimization problem (M) reads:

$$\text{Minimize} \quad F_v(x), \tag{M1}$$

$$\text{subject to} \quad x1_\ell = r \quad \text{and} \quad x'1_k = c, \tag{M2}$$

$$\text{over the set } x \in \mathbb{N}_v^{k\times\ell}. \tag{M3}$$

The *feasibility set* $R_v(r, c)$ is defined to comprise all matrices x satisfying (M3) and (M2), as in Balinski and Demange (1989a,b). We assume the feasibility set to be non-empty, $R_v(r, c) \neq \emptyset$. Hence the optimal value is positive, $\min_{x\in R_v(r,c)} F_v(x) > 0$.

For feasible matrices $x \in R_v(r, c)$ all marginals are fixed, whence a unit transfer triggers a chain reaction giving rise to cycles. We recall, from the proof of Theorem 7.1, that a cycle with index vectors $i_{(q)} = (i_1, \ldots, i_q)$ and $j_{(q)} = (j_1, \ldots, j_q)$ induces an elementary cycle matrix $e[i_{(q)}; j_{(q)}]$. Hence we define the *elementary transfer along the cycle with index vectors* $i_{(q)}$ *and* $j_{(q)}$ as a mapping $E[i_{(q)}; j_{(q)}] : \mathbb{N}_v^{k\times\ell} \to \mathbb{N}_v^{k\times\ell}$, through

$$E[i_{(q)}; j_{(q)}](x) \coloneqq x - e[i_{(q)}; j_{(q)}]$$

provided $v_{i_p j_p} > 0$ and $v_{i_p j_p^*} > 0$, and either $x_{i_p j_p} \geq 1$ and $s(1) > 0$, or else $x_{i_p j_p} \geq 2$ and $s(1) = 0$, for all $p \leq q$. That is, the entries (i_p, j_p) of the matrix x are decreased by one unit, and the entries (i_p, j_p^*) are increased. That the entries are eligible for this action is the proviso in the definition. In all other cases we set $E[i_{(q)}; j_{(q)}](x) \coloneqq x$. If x is feasible, so is its transform $E[i_{(q)}; j_{(q)}](x)$. An analysis of elementary transfers yields the following.

Theorem 8.1 (*Optimality*). *The following three statements are equivalent, for every feasible apportionment matrix* $x \in R_v(r, c)$*:*

(1) *x is a member of the apportionment set $A_v(r, c)$.*
(2) *x satisfies the critical inequalities, for all cycles on* $\operatorname{supp}(v)$.
(3) *x is an optimal solution of problem (M).*

Proof. The equivalence of (1) and (2) is Theorem 7.1. The equivalence of (2) and (3) follows as in Theorem 5.1, by replacing x_j by x_{ij}, and y_j by y_{ij}, and μ by $\rho_i\gamma_j$. □

The next corollary on multiple solutions is due to Balinski and Demange (1989a, Theorem 5, Lemma 2, 5). We present an alternate derivation paralleling our Corollary 5.2.

Corollary 8.2 (*Multiple Solutions*). *The following three statements are equivalent, for every optimal apportionment matrix* $x \in A_v(r, c)$*:*

(1) *There exists a matrix $y \in A_v(r, c)$ different from x.*
(2) *There exists a cycle on* $\operatorname{supp}(v)$ *for which the critical inequality holds with equality.*

Proof. Assume (1). As in the proof of Theorem 5.1, $F_v(x) = F_v(y)$ forces equality in $\prod_{n\leq x_{ij}, s(n)>0} \left(s(n)/v_{ij}\right) \leq (\rho_i\gamma_j)^{x_{ij}-y_{ij}} \prod_{n\leq y_{ij}, s(n)>0} \left(s(n)/v_{ij}\right)$, for all $(i, j) \in \operatorname{supp}(v)$. It follows that

$$x_{ij} > y_{ij} \Longrightarrow x_{ij} = y_{ij} + 1 \quad \text{and} \quad \rho_i\gamma_j = \frac{s(x_{ij})}{v_{ij}},$$

$$x_{ij} < y_{ij} \Longrightarrow x_{ij} + 1 = y_{ij} \quad \text{and} \quad \rho_i\gamma_j = \frac{s(x_{ij}+1)}{v_{ij}}.$$

The matrix $x - y$ is non-null, with row and column sums zero, and $v_{ij} = 0$ entailing $x_{ij} - y_{ij} = 0 - 0 = 0$. Hence we can construct a cycle $i_{(q)} = (i_1, \ldots, i_q)$ and $j_{(q)} = (j_1, \ldots, j_q)$ such that

$x_{i_p j_p} > y_{i_p j_p}$ and $x_{i_p j_p^*} < y_{i_p j_p^*}$, for all $p \le q$. The cycle lives on $\mathrm{supp}(v)$, and satisfies

$$\frac{s(x_{i_p j_p})}{v_{i_p j_p}} = \rho_{i_p}\gamma_{j_p} \quad \text{and} \quad \rho_{i_p}\gamma_{j_p^*} = \frac{s(x_{i_p j_p^*}+1)}{v_{i_p j_p^*}},$$

for all $p \le q$. Equality obtains, in the associated critical inequality:

$$\prod_{p\le q}\frac{s\left(x_{i_p j_p}\right)}{v_{i_p j_p}} = \left(\prod_{p\le q}\rho_{i_p}\right)\left(\prod_{p\le q}\gamma_{j_p}\right) = \left(\prod_{p\le q}\rho_{i_p}\right)\left(\prod_{p\le q}\gamma_{j_p^*}\right) = \prod_{p\le q}\frac{s\left(x_{i_p j_p^*}+1\right)}{v_{i_p j_p^*}}.$$

Conversely, assume (2), with a cycle with row indices $i_{(q)}$ and column indices $j_{(q)}$. Then $y := E[i_{(q)}; j_{(q)}](x) \neq x$ lies in $R_v(r,c)$. Extracting factors, as in the proof of Theorem 5.1, we find $F_v(y) = F_v(x)$. Thus y is optimal as well, and (1) is established. □

Corollary 8.3 (*Uniqueness*). *The following three statements are equivalent, for every optimal apportionment matrix $x \in A_v(r,c)$:*

(1) *The set $A_v(r,c)$ is a singleton, $A_v(r,c) = \{x\}$.*
(2) *For every cycle on* $\mathrm{supp}(v)$*, the critical inequality is strict.*

Proof. (1) and (2) are the negations of the assertions in Corollary 8.2. □

Just as the optimization formulations of the primal vector and matrix problems run parallel, so does the pairing of primal and dual optimization problems. For every feasible matrix $x \in R_v(r,c)$, the *row multipliers* $\alpha = (\alpha_1, \ldots, \alpha_k) \in (0,\infty)^k$ and *column multipliers* $\beta = (\beta_1, \ldots, \beta_\ell) \in (0,\infty)^\ell$ constitute the dual variables and induce lower bounds,

$$\begin{aligned} F_v(x) &= \left(\prod_{i\le k}\alpha_i^{r_i - x_{i+}}\right)\left(\prod_{j\le \ell}\beta_j^{c_j - x_{+j}}\right) F_v(x) \\ &\ge \inf_{y\in\mathbb{N}_v^{k\times\ell}}\left(\prod_{i\le k}\alpha_i^{r_i - y_{i+}}\right)\left(\prod_{j\le \ell}\beta_j^{c_j - y_{+j}}\right) F_v(y) =: G_v(\alpha,\beta). \end{aligned}$$

Lemma 8.4 (*Dual Goal Function*). *For all $\alpha \in (0,\infty)^k$ and $\beta \in (0,\infty)^\ell$ we have*

$$G_v(\alpha,\beta) = \left(\prod_{i\le k}\alpha_i^{r_i - z_{i+}}\right)\left(\prod_{j\le \ell}\beta_j^{c_j - z_{+j}}\right) F_v(z),$$

whenever the entries of the matrix $z \in \mathbb{N}^{k\times\ell}$ fulfill $z_{ij} \in [\![\alpha_i v_{ij} \beta_j]\!]$ for all $i \le k$ and $j \le \ell$.

Proof. The proof is a direct adaptation of the proof of Lemma 5.4. □

The dual problem (N) reads:

$$\text{Maximize } G_v(\alpha,\beta), \tag{N1}$$
$$\text{subject to } \alpha \in (0,\infty)^k \text{ and } \beta \in (0,\infty)^\ell. \tag{N2}$$

Primal and dual problems enjoy relations paralleling those of Theorem 5.5.

Theorem 8.5 (*Duality*). *Problems (M) and (N) both have optimal solutions, and share the same optimal value. Moreover, for all feasible apportionment matrices* $x \in R_v(r, c)$ *and for all row multipliers* $\rho \in (0, \infty)^k$ *and column multipliers* $\gamma \in (0, \infty)^\ell$*, the following three statements are equivalent:*

(1) x *is an optimal solution of* (M) *and* (ρ, γ) *is an optimal solution of* (N).
(2) x *satisfies the critical inequalities for all cycles on* $\mathrm{supp}(v)$*, and the multipliers satisfy*

$$\frac{s\left(x_{ij}\right)}{v_{ij}} \le \rho_i \gamma_j \le \frac{s\left(x_{ij}+1\right)}{v_{ij}} \quad \text{for all } (i, j) \in \mathrm{supp}(v).$$

(3) *For all* $i \le k$ *and* $j \le \ell$ *we have* $x_{ij} \in [\![\rho_i v_{ij} \gamma_j]\!]$.

Proof. The assertions rearrange the results of Theorem 8.1 and Lemma 8.4. □

9. Concluding remarks

In this paper we have exhibited a novel optimality approach to vector and matrix apportionment problems. The vector problem (V) and the matrix problem (M) differ in their side conditions, but use the same goal function, a cumulative product of signposts of the underlying rounding rule. In contrast, traditional criteria are peculiar for just a single method, and do not extend from vector to matrix problems.

Our primal–dual pairing complies with the well-known Fenchel duality theory. Though the primal programs are set up as integer problems, taking the logarithm of the goal function and interpolating linearly makes the goal function convex, see Gaffke and Pukelsheim (2008). Owing to the total unimodularity of the (linear) restrictions (V2) and (M2), the relaxed (continuous) problem has an optimal *integer* solution. Standard arguments establish duality. Since our apportionment problems are such that duality is easily obtained directly, we chose not to refer to the general theory.

A referee kindly alerted us of the available literature on diagonal equivalence of nonnegative matrices. Indeed, the sandwich Theorem 4.1 of Golitschek et al. (1983) supersedes our Theorem 7.1. Conversely, our optimality approach may be useful to investigate diagonal equivalence. For a recent paper on the subject, with more current references, see Hershkowitz and Schneider (2003).

The optimization approach entails practical merits in suggesting various algorithms, and in providing a frame for comparing them. The AS and TT algorithms that are mentioned in Section 1 solve the dual of the matrix apportionment problem, by varying the multipliers until the side conditions are met. Other algorithms operate on the primal problem, adjusting the seat allocations while obeying the side conditions until the goal function is minimum. A detailed discussion of primal and dual algorithms is presented in the companion paper Gaffke and Pukelsheim (2008).

Acknowledgements

We are very grateful to Michel Balinski for challenging remarks and valuable suggestions, and extend our thanks to Günter Rote and the three referees for pertinent comments that improved the paper significantly. In particular, the references Golitschek et al. (1983) and Hershkowitz and Schneider (2003) are owed to a referee. F.P. acknowledges the hospitality of the Laboratoire d'Économétrie, École Polytechnique, Paris, during a sabbatical visit 2005–2006.

References

Bacharach, M., 1965. Estimating nonnegative matrices from marginal data. International Economic Review (Osaka) 6, 294–310.

Bacharach, M., 1970. Biproportional Matrices & Input–Output Change. Cambridge, UK.

Balinski, M., 2004. Le suffrage universel inachevé, Paris.

Balinski, M., 2006. Apportionment: Uni-and bidimensional, In Simeone and Pukelsheim (2006), 43–53.

Balinski, M., Demange, G., 1989a. An axiomatic approach to proportionality between matrices. Mathematics of Operations Research 14, 700–719.

Balinski, M., Demange, G., 1989b. Algorithms for proportional matrices in reals and integers. Mathematical Programming 45, 193–210.

Balinski, M., Pukelsheim, F., 2006. Matrices and politics. In: Liski, E.P., Isotalo, J., Niemelä, J., Puntanen, S., Styan, G.P.H. (Eds.), Festschrift for Tarmo Pukkila on his 60th Birthday, Tampere, 233–242.

Balinski, M., Pukelsheim, F., 2007. Das Neue Zürcher Zuteilungsverfahren: Proportionalität und Kohärenz im Doppelpack. Spektrum der Wissenschaft 4, 76–80.

Balinski, M., Rachev, S.T., 1997. Rounding proportions: Methods of rounding. Mathematical Scientist 22, 1–26.

Balinski, M., Ramírez Gonzales, V., 1999. Parametric methods of apportionment, rounding and production. Mathematical Social Sciences 37, 107–122.

Balinski, M., Young, H.P., 2001. Fair Representation—Meeting the Ideal of One Man, One Vote. Second ed. Washington, DC.

Bishop, Y.M., Fienberg, S., Holland, P., 1975. Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA..

Bochsler, D., 2005. Biproportionale Wahlverfahren für den Schweizer Nationalrat. www.opus-bayern.de/uni-augsburg/volltexte/2005/160.

Carnal, H., 1993. Mathématiques et politique. Elemente der Mathematik 48, 27–32.

Carnal, H., Riedwyl, H., 1982. Wahlkreisverbandsarithmetik, Technischer Bericht No. 8, Institut für Mathematische Statistik und Versicherungslehre, Universität Bern.

Carnal, H., Riedwyl, H., 1984. Der Kanton Bern sucht ein neues, alle zufriedenstellendes Wahlgesetz, Der Bund, 135. Jahrgang, Nr. 256 (Mittwoch, 31. Oktober 1984), 2.

Deming, W.E., Stephan, F.F., 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. Annals of Mathematical Statistics 11, 427–440.

Dorfleitner, G., Klein, T., 1999. Rounding with multiplier methods: An efficient algorithm and applications in statistics. Statistical Papers 20, 143–157.

Gaffke, N., Pukelsheim, F., 2008. Vector and matrix apportionment problems and separable convex integer optimization. Mathematical Methods of Operations Research 67, 133–159.

Golitschek, M.v., Rothblum, U.G., Schneider, H., 1983. A conforming decomposition theorem, a piecewise linear theorem of the alternative, and scalings of matrices satisfying lower and upper bounds. Mathematical Programming 27, 291–306.

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.

Hershkowitz, D., Schneider, H., 2003. One-sided simultaneous inequalities and sandwich theorems for diagonal similarity and diagonal equivalence of nonnegative matrices. Electronic Journal of Linear Algebra 10, 81–101.

Joas, B., 2005. A graph theoretic solvability check for biproportional multiplier methods. Diploma-Thesis, Institut für Mathematik, Universität Augsburg.

Kalantari, B., Lari, I., Ricca, F., Simeone, B., 2008. On the complexity of general matrix scaling and entropy minimization via the RAS algorithm. Mathematical Programming, Series A 112, 371–401. doi:10.1007/s10107-006-0021-4.

Maier, S., 2006. Algorithms for biproportional apportionment methods. In Simeone and Pukelsheim (2006), 105–116.

Marshall, A.W., Olkin, I., 1968. Scaling of matrices to achieve specified row and column sums. Numerische Mathematik 12, 83–90.

Pukelsheim, F., 2000. Mandatszuteilungen bei Verhältniswahlen: Erfolgswertgleichheit der Wählerstimmen. Allgemeines Statistisches Archiv—Journal of the German Statistical Society 84, 447–459.

Pukelsheim, F., 2004. Bazi—A Java program for proportional representation, Oberwolfach Reports 1, 735–737.

Pukelsheim, F., Schuhmacher, C., 2004. Das neue Zürcher Zuteilungsverfahren für Parlamentswahlen. Aktuelle Juristische Praxis—Pratique Juridique Actuelle 5, 505–522.

Rockafellar, R.T., 1970. Convex Analysis. Princeton, NJ.

Rote, G., Zachariassen, M., 2007. Matrix scaling by network flow. In: Proceedings of the Eighteenth Annual ACM–SIAM Symposium on Discrete Algorithms, SIAM Proceedings in Applied Mathematics 125, pp. 848–854.

Schindler, P., 1999. Datenhandbuch zur Geschichte des Deutschen Bundestages 1949 bis 1999, Gesamtausgabe in drei Bänden und auf CD. Eine Veröffentlichung der Wissenschaftlichen Dienste des Deutschen Bundestages, Baden-Baden.

Simeone, B., Pukelsheim, F. (Eds.), 2006. Mathematics and Democracy—Recent Advances in Voting Systems and Collective Choice. Berlin.

Stephan, F.F., 1942. An iterative method of adjusting sample frequency tables when expected marginal totals are known. Annals of Mathematical Statistics 13, 166–178.

Zachariassen, P., Zachariassen, M., 2006. A comparison of electoral formulae for the Faroese Parliament. In Simeone and Pukelsheim (2006), pp. 235–251.