

OPTIMIZATION OF MATRIX SEMIRINGS FOR CLASSIFICATION SYSTEMS

D. Y. GAO, A. V. KELAREV  and J. L. YEARWOOD

(Received 5 April 2011)

Abstract

The max-plus algebra is well known and has useful applications in the investigation of discrete event systems and affine equations. Structural matrix rings have been considered by many authors too. This article introduces more general structural matrix semirings, which include all matrix semirings over the max-plus algebra. We investigate properties of ideals in this construction motivated by applications to the design of centroid-based classification systems, or classifiers, as well as multiple classifiers combining several initial classifiers. The first main theorem of this paper shows that structural matrix semirings possess convenient visible generating sets for ideals. Our second main theorem uses two special sets to determine the weights of all ideals and describe all matrix ideals with the largest possible weight, which are optimal for the design of classification systems.

2010 Mathematics subject classification: primary 16Y60; secondary 62H30.

Keywords and phrases: max-plus algebras, semirings, matrix constructions, classification systems.

1. Introduction

The max-plus algebra is well known and has useful applications in the investigation of discrete event systems and affine equations (see [4, 13, 14]). On the other hand, structural matrix rings have also been considered by many authors (see, for example, [7, 15, 28] and [17, Section 3.14]). The present article introduces more general structural matrix semirings, incorporating all matrix semirings over the max-plus algebra.

We investigate properties of ideals in this construction motivated by applications to the design of centroid-based classification systems, or classifiers, as well as for multiple classifiers combining several initial classifiers (see the monograph [27], recent articles [24, 25] and Section 2 below). The concept of an ideal is very important and has been considered in various branches of mathematics (see, for example

The first author has been supported by grant FA9550-10-1-0487 from US Air Force Office of Scientific Research (AFOSR). The second author was supported by Discovery grant DP0449469 from the Australian Research Council. The third author was supported by a Queen Elizabeth II Fellowship, Discovery grant DP0211866, and Linkage grant LP0990908 from the Australian Research Council.

© 2011 Australian Mathematical Publishing Association Inc. 0004-9727/2011 \$16.00

[1–3, 8, 9, 15, 16, 20, 23, 25, 26, 31–33]). In particular, visible generating sets have been found for certain ideals of group rings in [8] and polynomial quotient rings in [25].

Our first main theorem shows that every structural matrix semiring possesses a convenient visible generating set \mathcal{V} for matrix ideals; see Theorem 3.2 in Section 3. Our second main theorem demonstrates that every structural matrix semiring has two special sets, \mathcal{G} and \mathcal{G}_{\max} , such that the weight of each nonzero matrix ideal I (respectively, each matrix ideal I of largest possible weight) is equal to the minimum of the weights of the elements in $I \cap \mathcal{G}$ (respectively, $I \cap \mathcal{G}_{\max}$); see Theorem 3.3 in Section 3. This means that the set \mathcal{G} can be used to determine the weights of all ideals in this construction, and the set \mathcal{G}_{\max} gives us a description of all matrix ideals with largest possible weight, which are optimal for the design of classification systems.

2. Motivation and preliminaries

Semirings and matrix constructions over them can be used to generate convenient sets of centroids for centroid-based classification systems, or classifiers, and to design combined multiple classifiers capable of correcting the errors of individual initial classifiers. We include concise preliminaries in this section and refer the reader to [13, 17–19, 21, 22, 27, 29] for more detailed discussions.

The classification process begins with a representation of data in a standard vector space F^n , where F can be regarded as a semifield. A *semifield* is a semiring, where the set of all nonzero elements forms a group with respect to multiplication. Recall that a *semiring* is a set F with two binary operations, addition $+$ and multiplication \cdot , such that the following conditions are satisfied:

- (S1) $(F, +)$ is a commutative semigroup with zero 0 ;
- (S2) (F, \cdot) is a semigroup;
- (S3) multiplication distributes over addition;
- (S4) zero 0 annihilates F , that is, $0 \cdot F = F \cdot 0 = 0$.

It is also often assumed that every semiring satisfies the following additional property:

- (S5) (F, \cdot) has an identity element 1 .

Here we look at more general semirings, which do not have to satisfy (S5), because such more general terminology adds the convenience of allowing us to consider more general subsets as subsemirings without assuming that all subsemirings contain the identity element. By analogy with a similar situation in ring theory, we call every semiring satisfying (S5) a *semiring with identity element*. Both terminologies are essentially equivalent, since it is always easy to adjoin an identity element to every semiring that does not have one. Semirings have been used in broad areas (see [4, 13, 14]). Originally, our investigation of semirings was motivated by the development of methods useful for duality theory (see [6, 10–12]).

Every centroid-based classification system, or classifier, selects special elements c_1, \dots, c_k in F^n , called *centroids* (see [5, 27]). For $i = 1, \dots, k$, each centroid c_i

defines its class $N(c_i)$ consisting of all vectors v such that c_i is the nearest centroid of v . Then every vector in F^n , representing the attributes of features of an instance, is assigned to the class of its nearest centroid.

On the other hand, multiple classifiers are often used in analysis of data to combine individual initial classifiers (see, for example, [30]). A well-known method for the design of multiple classifiers consists in designing several simpler initial or individual classifiers, and then combining them into one multiple classification scheme with several classes. This method is very effective, and it is often recommended for various applications (see [27, Section 7.5]).

Denote the number of initial classifiers being combined by n . If o_1, \dots, o_n are the outputs of the initial classifiers, then the sequence (o_1, \dots, o_n) is called a *vector of outputs* of the initial classifiers. In order to define the multiple classifier and enable correction of errors of the initial classifiers, a set of centroids c_1, \dots, c_k is again selected in F^n . For $i = 1, \dots, k$, the class $N(c_i)$ of the centroid c_i is again defined as the set of all observations with the vector outputs of the initial classifiers having c_i as its nearest centroid.

The design of multiple classifiers by combining individual classifiers is quite common in the literature. We refer to [24, 25, 27] for a list of properties required of the sets of centroids. In particular, it is essential to find sets of centroids with large weights and small numbers of generators. The *weight* $\text{wt}(v)$ of $v \in F^n$ is the number of nonzero components or coordinates in v . The *weight* of a set $C \subseteq F^n$ is the minimum weight of a nonzero element in C . For more information we refer the reader to [24, 25].

The *information rate* of a class set C in F^n can be defined as $\log_{|F|}(|C|)/n$. It reflects the proportion of output of the individual initial classifiers used to produce the outcomes of the multiple classification, as opposed to additional efforts spent on increasing reliability and correcting classification errors.

All sequences of the centroid set C can be written down in a matrix M to discuss their properties. If M has two identical columns, this means that two initial classifiers produce identical outputs. This duplication is very inefficient, even though it could help to correct classification errors. Therefore, in a situation like this, one of these classifiers can be removed and a better scheme can be devised. Likewise, it is undesirable to have strong correlation or functional dependencies between very small sets of columns in M or between the initial classifiers.

According to [27, Section 7.5], for a classifier with a class set C to be efficient, the class C must satisfy the following most essential basic properties.

- (1) The set C must have a large weight.
- (2) The information rate of C must be large.
- (3) A small set of generators for the set C is essential in order to simplify computer storage and manipulation of the set.
- (4) If all vectors of C are recorded in a matrix M , then there should not be strong correlation or functional dependencies between small sets of columns of M . In particular, the matrix M should not have duplicate columns.

Thus, in particular, it is essential to find sets of centroids with large weights and small numbers of generators. For additional references and discussion of experimental research related to these properties we refer the reader to [24, 25, 29, 30].

Let F be a semifield, and let $M_m(F)$ be the semiring of all $m \times m$ matrices over F . Let ϱ be a binary relation on the set $[1 : m] = \{1, \dots, m\}$. For $i, j \in [1 : m]$, denote by $e_{i,j}$ the standard elementary matrix in $M_m(F)$ with 1 in the intersection of i th row and j th column and zeros in all other cells or entries. It is well known and easy to verify that the set $M_\varrho(F) = \bigoplus_{(i,j) \in \varrho} F e_{i,j}$ is a subsemiring of $M_m(F)$ if and only if the relation ϱ is transitive. In this case $M_\varrho(F)$ is called a *structural matrix semiring*. Many interesting results on structural matrix rings have been obtained in the literature (see, for example, [7, 15, 28]). Known facts and references concerning structural matrix rings can also be found in [17]—see Section 3.14 and the bibliography.

If $|\varrho| = n$, then the additive semigroup of $M_\varrho(F)$ is isomorphic to F^n and we can introduce multiplication in F^n by identifying it with $M_\varrho(F)$. Further, we are going to look at the sets of centroids generated in $M_\varrho(F)$ as ideals.

Denote the set of all positive integers by \mathbb{N} and put $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Every set of elements $g_1, \dots, g_k \in M_\varrho(F)$ generates an *ideal* $\text{id}(g_1, \dots, g_k)$ in $M_\varrho(F)$, namely, the set of all sums of these elements and their multiples, that is, the set of all sums of the form

$$\sum_{j=1}^{m_1} \ell_{1,j} g_1 r_{1,j} + \dots + \sum_{j=1}^{m_k} \ell_{k,j} g_k r_{k,j}, \tag{2.1}$$

where $k \in \mathbb{N}$ and $\ell_{i,j}, r_{i,j} \in M_\varrho(F) \cup \{1\}$, for $i = 1, \dots, k$.

3. Main results

DEFINITION 3.1. A set U in a semiring Q is called a *visible generating set* or a *visible set of generators* if, for any subset V of U , the weight of the ideal $\text{id}(V)$ in Q is equal to the minimum of the weights of the generators $v \in V$.

This notion was introduced in [8] and was also studied in [25]. It is related to the concept of a visible basis, investigated by many authors (see [17, 25] for additional references). Let ϱ be a binary relation on the set $[1 : m]$. We define two sets

$$\begin{aligned} L &= \{i \mid \exists k : (k, i) \in \varrho\}, \\ R &= \{j \mid \exists k : (j, k) \in \varrho\} \end{aligned}$$

and the relations

$$\begin{aligned} \varrho_\ell &= \{(i, j) \in \varrho \mid i \in L\}, \\ \varrho_r &= \{(i, j) \in \varrho \mid j \in R\}. \end{aligned}$$

For any $i \in [1 : m]$, we introduce notation for the sets

$$\begin{aligned} \varrho(i) &= \{j \mid (i, j) \in \varrho\}, \\ \varrho^{-1}(i) &= \{j \mid (j, i) \in \varrho\}. \end{aligned}$$

For any $d \in \mathbb{N}$, let $\mathcal{P}_L(d)$ be the set of all pairs (S, i) such that $i \in [1 : m]$, $S \subseteq \varrho^{-1}(i) \setminus L$ and $|S| = d$. Denote by M_L the largest positive integer such that the set $\mathcal{P}_L(M_L)$ is not empty, or zero if such integers do not exist.

Similarly, for any $d \in \mathbb{N}$, let $\mathcal{P}_R(d)$ stand for the set of all pairs (i, S) such that $i \in [1 : m]$, $S \subseteq \varrho(i) \setminus R$ and $|S| = d$. Denote by M_R the largest positive integer such that the set $\mathcal{P}_R(M_R)$ is not empty, or zero if such integers do not exist.

Put $M_Z = |\varrho \setminus (\varrho_r \cup \varrho_\ell)|$, and let

$$M = \max\{1, M_Z, M_L, M_R\}.$$

Next, we put

$$g_Z = \sum_{(i,j) \in \varrho \setminus (\varrho_r \cup \varrho_\ell)} e_{i,j} = \sum_{(i,j) \in \varrho, i \notin L, j \notin R} e_{i,j},$$

and introduce the following subsets of $M_\varrho(F)$:

$$\begin{aligned} \mathcal{V}_H &= \begin{cases} \{e_{i,j} \mid (i, j) \in \varrho\} & \text{if } M = 1 \\ \emptyset & \text{otherwise;} \end{cases} \\ \mathcal{V}_Z &= \begin{cases} \{g_Z\} & \text{if } M = |\varrho \setminus (\varrho_r \cup \varrho_\ell)| \\ \emptyset & \text{otherwise;} \end{cases} \\ \mathcal{V}_L &= \left\{ \sum_{j \in S} e_{j,i} \mid (S, i) \in \mathcal{P}_L(M) \right\}; \\ \mathcal{V}_R &= \left\{ \sum_{j \in S} e_{i,j} \mid (i, S) \in \mathcal{P}_R(M) \right\}. \end{aligned}$$

Finally, put

$$\mathcal{V} = \mathcal{V}(F, \varrho) = \mathcal{V}_H \cup \mathcal{V}_Z \cup \mathcal{V}_L \cup \mathcal{V}_R.$$

THEOREM 3.2. *Let F be a semifield, and let $M_\varrho(F)$ be a structural matrix semiring over F , where ϱ is a transitive relation. Then the set $\mathcal{V}(F, \varrho)$ is a visible generating set for ideals in $M_\varrho(F)$.*

Denote by \mathcal{G}_Z the set of all nonzero elements

$$g = \sum_{(i,j) \in \varrho \setminus (\varrho_r \cup \varrho_\ell)} f_{i,j} e_{i,j} \in M_\varrho(F),$$

where $f_{i,j} \in F$. Let

$$\begin{aligned} \mathcal{G}_L &= \left\{ \sum_{j \in S} f_j e_{j,i} \in M_\varrho(F) \mid (S, i) \in \bigcup_{d=1}^\infty \mathcal{P}_L(d), 0 \neq f_j \in F \right\}, \\ \mathcal{G}_R &= \left\{ \sum_{j \in S} f_j e_{i,j} \in M_\varrho(F) \mid (i, S) \in \bigcup_{d=1}^\infty \mathcal{P}_R(d), 0 \neq f_j \in F \right\}. \end{aligned}$$

Put $\mathcal{G}_H = \{fe_{i,j} \mid (i, j) \in \varrho, 0 \neq f \in F\}$ and

$$\mathcal{G} = \mathcal{G}_H \cup \mathcal{G}_Z \cup \mathcal{G}_L \cup \mathcal{G}_R.$$

For any $d \in \mathbb{N}$, let $\mathcal{P}_Z(d)$ be the set of all subsets S of $\varrho \setminus (\varrho_r \cup \varrho_\ell)$ such that $|S| = d$. Denote by \mathcal{G}_{\max}^Z the set of all nonzero elements $g = \sum_{(i,j) \in S} f_{i,j}e_{i,j} \in M_\varrho(F)$, for all $S \in \mathcal{P}_Z(M)$, where $0 \neq f_{i,j} \in F$. Let \mathcal{G}_{\max}^L be the set of all nonzero elements $g = \sum_{j \in S} f_j e_{j,i} \in M_\varrho(F)$, for all pairs $(S, i) \in \mathcal{P}_L(M)$ and all $0 \neq f_j \in F$. Denote by \mathcal{G}_{\max}^R the set of all nonzero elements $g = \sum_{j \in S} f_j e_{i,j} \in M_\varrho(F)$, for all pairs $(i, S) \in \mathcal{P}_R(M)$ and all $0 \neq f_j \in F$. Set

$$\mathcal{G}_{\max}^H = \begin{cases} \mathcal{G}_H & \text{if } M = 1 \\ \emptyset & \text{otherwise.} \end{cases}$$

Finally, put

$$\mathcal{G}_{\max} = \mathcal{G}_{\max}^H \cup \mathcal{G}_{\max}^Z \cup \mathcal{G}_{\max}^L \cup \mathcal{G}_{\max}^R.$$

THEOREM 3.3. *Let F be a semifield, and let $M_\varrho(F)$ be a structural matrix semiring over F , where ϱ is a transitive relation. Then the following conditions hold:*

- (i) *for every nonzero ideal I of $M_\varrho(F)$, the weight of I is equal to the minimum of the weights of the generators in $I \cap \mathcal{G}$;*
- (ii) *for every nonzero ideal I of the largest possible weight in $M_\varrho(F)$, the weight of I is equal to the minimum of the weights of the generators in $I \cap \mathcal{G}_{\max}$.*

4. Proofs

For any $x = \sum_{(i,j) \in \varrho} f_{i,j}e_{i,j} \in M_\varrho(F)$, the *support* of x is the set

$$\text{supp}(x) = \{(i, j) \in \varrho \mid f_{i,j} \neq 0\}.$$

For any semiring Q , the *left annihilator* of Q is the set

$$\text{Ann}_\ell(Q) = \{x \in Q \mid xQ = 0\},$$

and the *right annihilator* of Q is the set

$$\text{Ann}_r(Q) = \{x \in Q \mid Qx = 0\}.$$

LEMMA 4.1. *For any semifield F and any structural matrix semiring $M_\varrho(F)$, where ϱ is a transitive relation, the following equalities hold:*

$$\text{Ann}_r(M_\varrho(F)) = M_{\varrho \setminus \varrho_\ell}(F), \tag{4.1}$$

$$\text{Ann}_\ell(M_\varrho(F)) = M_{\varrho \setminus \varrho_r}(F). \tag{4.2}$$

PROOF. Take any element r in $\text{Ann}_r(M_\varrho(F))$. It can be recorded as $r = \sum_{(i,j) \in \varrho} f_{i,j}e_{i,j}$, where $f_{i,j} \in F$. Pick any pair (i, j) in ϱ_ℓ . There exists $k \in [1 : m]$ such that $(k, i), (k, j) \in \varrho$. Hence $e_{k,i} \in M_\varrho(F)$ and $f_{i,j}e_{k,i}e_{i,j}$ is a summand of the product $e_{k,i}r$ in $M_\varrho(F)$.

Since $r \in \text{Ann}_r(M_\varrho(F))$, we get $f_{i,j} = 0$. It follows that r belongs to $M_{\varrho \setminus \varrho_\ell}(F)$, and so $\text{Ann}_r(M_\varrho(F)) \subseteq M_{\varrho \setminus \varrho_\ell}(F)$.

To prove the reverse inclusion, let us pick any element r in $M_{\varrho \setminus \varrho_\ell}(F)$. It can be written down as $r = \sum_{(i,j) \in \varrho \setminus \varrho_\ell} f_{i,j} e_{i,j}$, where $f_{i,j} \in F$. In order to verify that $M_\varrho(F)r = 0$, it suffices to show that $e_{a,b}r = 0$ for all $(a, b) \in \varrho$. Suppose to the contrary that $e_{a,b}r \neq 0$ for some $(a, b) \in \varrho$. Then it is clear that at least one of the summands $e_{a,b}f_{i,j}e_{i,j}$ is nonzero for some $(i, j) \in \varrho \setminus \varrho_\ell$. The definition of a structural matrix semiring implies that $f_{i,j} \neq 0$, $b = i$, $(a, i) = (a, b) \in \varrho$ and $(a, j) \in \varrho$. Hence $(i, j) \in \varrho_\ell$. This contradicts the choice of (i, j) in $\varrho \setminus \varrho_\ell$ and shows that $e_{a,b}r = 0$ for all $(a, b) \in \varrho$. Therefore $M_\varrho(F)r = 0$, which means that $r \in \text{Ann}_r(M_\varrho(F))$. Thus $\text{Ann}_r(M_\varrho(F)) \supseteq M_{\varrho \setminus \varrho_\ell}(F)$.

These two inclusions show that equality (4.1) always holds. The proof of equality (4.2) is dual and we omit it. □

LEMMA 4.2. *Let $M_\varrho(F)$ be a structural matrix semiring over a semifield F , where ϱ is a transitive relation, and let $g \in \mathcal{G}_Z$. Then $\text{id}(g) = \{cg \mid c \in \mathbb{N}_0\}$ and $\text{wt}(\text{id}(g)) = \text{wt}(g)$.*

PROOF. By the definition of \mathcal{G}_Z , we get

$$g = \sum_{(i,j) \in \varrho \setminus (\varrho_r \cup \varrho_\ell)} f_{i,j} e_{i,j} \in M_\varrho(F),$$

where $f_{i,j} \in F$. It follows from (2.1) that

$$\text{id}(g) = \left\{ \sum_{i=1}^k \ell_i g r_i \mid k \in \mathbb{N}; \ell_i, r_i \in M_\varrho(F) \cup \{1\} \right\}.$$

Since $\text{supp}(g) \cap \varrho_\ell = \emptyset$, equality (4.1) in Lemma 4.1 shows that $\ell_i g = 0$ for every $\ell_i \in M_\varrho(F)$. Equality (4.2) in Lemma 4.1 shows that $g r_i = 0$ for every $r_i \in M_\varrho(F)$, because $\text{supp}(g) \cap \varrho_r = \emptyset$. Hence we may assume that $\ell_i = r_i = 1$, for $i = 1, \dots, k$. Therefore $\ell_i g r_i = g$, for all $i = 1, \dots, k$, and we get

$$\text{id}(g) = \{cg \mid c \in \mathbb{N}_0\},$$

as required. It follows that $\text{wt}(\text{id}(g)) = \text{wt}(g)$, which completes the proof. □

LEMMA 4.3. *Let $M_\varrho(F)$ be a structural matrix semiring over a semifield F , where ϱ is a transitive relation, and let $g = \sum_{s \in S} f_s e_{s,i} \in \mathcal{G}_L$, where $i \in [1 : m]$, $(S, i) \in \mathcal{P}_L(d)$ for some $d \in \mathbb{N}$, and $0 \neq f_s \in F$. Then*

$$\text{id}(g) = \left\{ cg + \sum_{(i,b) \in \varrho} h_b g e_{i,b} \mid c \in \mathbb{N}_0, h_b \in F \right\} \tag{4.3}$$

and $\text{wt}(\text{id}(g)) = \text{wt}(g)$.

PROOF. It follows from (2.1) that $\text{id}(g)$ contains the set in the right-hand side of (4.3). To prove the reverse inclusion, choose any element x in $\text{id}(g)$. By (2.1),

$$x = \sum_{j=1}^k \ell_j g r_j, \tag{4.4}$$

for some $\ell_j, r_j \in M_\varrho(F) \cup \{1\}$. We may assume that only nonzero summands $\ell_j r_j$ are included in the sum (4.4).

Since $(S, i) \in \mathcal{P}_L(d)$, we get $(s, i) \in \varrho \setminus \varrho_\ell$, for all $s \in S$. The equality (4.1) of Lemma 4.1 tells us that $g \in \text{Ann}_r(M_\varrho(F))$; whence $\ell_j g = 0$ for every $\ell_j \in M_\varrho(F)$. Hence we may assume that all the ℓ_j in (4.4) are equal to 1. Therefore, we can rewrite (4.4) as

$$x = c g + \sum_{j=1}^k g r_j, \tag{4.5}$$

for some $c \in \mathbb{N}_0, r_j \in M_\varrho(F)$. Keeping in mind that $M_\varrho(F) = \bigoplus_{(a,b) \in \varrho} F e_{a,b}$, without loss of generality the distributive law allows us to assume that every element $r_j \in M_\varrho(F)$ in (4.5) is *homogeneous*, that is, $r_j = h' e_{a,b}$ for some $h' \in F, (a, b) \in \varrho$. Then it follows that $r_j = h_b e_{i,b}$, for some $h_b \in F$, because $g r_j \neq 0$. Therefore, we can rewrite (4.5) in the required form as in the right-hand side of (4.3). This shows that (4.3) always holds.

In order to prove that $\text{wt}(\text{id}(g)) = \text{wt}(g)$, let us take any nonzero element x in $\text{id}(g)$ and represent it as a sum in the right-hand side of (4.3).

Since F is a semifield, we get $\text{wt}(hg) = \text{wt}(g) = |S|$, if $h \neq 0$. The transitivity of ϱ implies that $\varrho(i) \subseteq \varrho(s)$, for all s in S . Likewise, since F is a semifield, it follows that

$$\text{wt}(h_b g e_{i,b}) = \text{wt}(g) = |S|,$$

whenever $h_b \neq 0$. Since $h_b g e_{i,b} = \sum_{s \in S} (h_b f_s) e_{s,b}$, it follows from (4.3) that $\text{wt}(x) \geq \text{wt}(g)$ for each nonzero element x in $\text{id}(g)$. Therefore $\text{wt}(x) \geq \text{wt}(g)$. The reverse inequality is obvious, and so we get $\text{wt}(\text{id}(g)) = \text{wt}(g) = |S|$. This completes the proof. □

LEMMA 4.4. *Let $M_\varrho(F)$ be a structural matrix semiring over a semifield F , where ϱ is a transitive relation, and let $g = \sum_{s \in S} f_s e_{i,s} \in \mathcal{G}_R$, where $i \in [1 : m], (i, S) \in \mathcal{P}_R(d)$ for some $d \in \mathbb{N}$, and $0 \neq f_s \in F$. Then*

$$\text{id}(g) = \left\{ c g + \sum_{(a,i) \in \varrho} h_a e_{a,i} g \mid c \in \mathbb{N}_0, h_a \in F \right\}$$

and $\text{wt}(\text{id}(g)) = \text{wt}(g)$.

PROOF. The proof is omitted, since it is dual to the proof of Lemma 4.3. □

PROOF OF THEOREM 3.2. If $M = 1$, then $\mathcal{V} = \mathcal{V}_H$ and it is clear that this set is a visible generating set. Further, we assume that $M > 1$. Then $\mathcal{V}_H = \emptyset$.

Take any subset V in $\mathcal{V} = \mathcal{V}(F, \varrho)$ and consider the ideal $I = \text{id}(V)$ in $M_\varrho(F)$. Let x be a nonzero element of minimal weight in I , so that $\text{wt}(x) = \text{wt}(I)$. The weight of each element in \mathcal{V} is equal to M . Consequently, in order to prove that \mathcal{V} is a visible generating set, it suffices to show that $\text{wt}(x) = M$. Since $V \subseteq \text{id}(V)$, the minimality of $\text{wt}(x)$ shows that $\text{wt}(x) \leq M$. Therefore it remains to verify that $\text{wt}(x) \geq M$.

By (2.1), we get

$$x = \sum_{i=1}^k x_i, \tag{4.6}$$

where $x_i \in \text{id}(g_i)$, for some $g_i \in V$. Without loss of generality we may assume that the positive integer k has a minimum value in (4.6). Lemmas 4.2, 4.3 and 4.4 imply that $\text{wt}(x_i) \geq \text{wt}(g_i)$, for all $i = 1, \dots, k$. Therefore, if $k = 1$, then the desired inequality $\text{wt}(x) \geq M$ follows.

In the rest of the proof, we assume that $k > 1$. We are going to show that this assumption always leads to a contradiction with the minimality of k . This will mean that $k = 1$, which will complete the proof.

If $\text{supp}(x_i) \cap \text{supp}(x_j) = \emptyset$, for all $1 \leq i, j \leq k, i \neq j$, then $\text{wt}(x) = \sum_{i=1}^k \text{wt}(x_i) \geq M$, and so we are done in this case. Further, we assume that the supports of some of the summands in the right-hand side of (4.6) intersect. To simplify notation we may assume that $\text{supp}(x_1) \cap \text{supp}(x_2) \neq \emptyset$. Then there exists $(a, b) \in \varrho$ such that $(a, b) \in \text{supp}(x_1) \cap \text{supp}(x_2)$. There are several cases to consider.

Case 1. $g_1, g_2 \in \mathcal{V}_Z$. Since $\mathcal{V}_Z = \{g_Z\}$, we see that $g_1 = g_2 = g_Z$, and so we can replace two summands x_1 and x_2 in (4.6) by just one summand $x_1 + x_2$, because $x_1 + x_2 \in \text{id}(g_Z)$. This contradicts the minimality of k , and shows that Case 1 is impossible.

Case 2. At least one of the elements g_1, g_2 belongs to \mathcal{V}_L . Without loss of generality we may assume that $g_1 \in \mathcal{V}_L$. Then it follows that $g_1 = \sum_{s \in S} e_{j,b}$, for some $(S, b) \in \mathcal{P}_L(M)$, where $a \in S$ and $|S| = M$. There are three subcases to consider.

Subcase 2.1. $g_2 \in \mathcal{V}_Z$. Since $(a, b) \in \text{supp}(x_2)$, it follows from Lemma 4.2 that $b \notin R$. Besides, it follows from $(S, b) \in \mathcal{P}_L(M)$ that $S \cap L = \emptyset$. Therefore $(s, b) \in \varrho \setminus (\varrho_\ell \cup \varrho_r)$, for all $s \in S$. However, the definition of M shows that $M = |S| = |\varrho \setminus (\varrho_\ell \cup \varrho_r)|$. Hence

$$\varrho \setminus (\varrho_\ell \cup \varrho_r) = \{(s, b) \mid s \in S\}.$$

Therefore $g_1 = g_Z$. This means that Subcase 2.1 coincides with Case 1, and so it is impossible.

Subcase 2.2. $g_2 \in \mathcal{V}_L$. Then $g_2 = \sum_{s \in S_2} e_{j,b}$, for some $(S_2, b) \in \mathcal{P}_L(M)$, where $a \in S$ and $|S| = M$. It follows from the maximality of M and Lemma 4.3 that

$$S = \varrho^{-1}(b) \setminus L = S_2.$$

Hence $g_1 = g_2$. Therefore we can replace two summands x_1 and x_2 in (4.6) by just one summand $x_1 + x_2$, because $x_1 + x_2 \in \text{id}(g_1)$. This contradicts the minimality of k , and shows that Subcase 2.2 is impossible.

Subcase 2.3. $g_2 \in \mathcal{V}_R$. Then it follows from $(a, b) \in \text{supp}(x_2)$ and Lemma 4.4 that $b \notin R$. As in Subcase 2.1, this yields $g_1 = g_Z$ again. Thus, Subcase 2.3 is impossible.

Case 3. At least one of the elements g_1, g_2 belongs to \mathcal{V}_R . This case is dual to Case 2 above, and so a dual proof shows that it is also impossible. This completes the proof. \square

PROOF OF THEOREM 3.3. The proof of condition (ii) is similar to that of (i), the only difference being that a few small clarifications have to be inserted in the proof of (i) to prove (ii). In order to avoid unnecessary duplication we will give a complete proof for condition (i), and in parentheses will indicate small additions required for the proof of (ii).

Let us start by choosing a nonzero ideal I of $M_\varrho(F)$. (For the proof of (ii) in addition here we assume that I has the largest possible weight in $M_\varrho(F)$.)

Clearly, $\text{wt}(I) = 1$ if and only if $I \cap \mathcal{G}_H \neq \emptyset$. Therefore in the case where $\text{wt}(I) = 1$ the theorem is straightforward, and so in the rest of the proof we assume that $\text{wt}(I) > 1$. Choose a nonzero element g of minimal weight in I so that $\text{wt}(g) = \text{wt}(I)$. There are several cases to consider.

Case 1. $g \in \text{Ann}_\ell(M_\varrho(F)) \cap \text{Ann}_r(M_\varrho(F))$. Lemma 4.1 shows that $r \in M_{\varrho \setminus \varrho_\ell}(F) \cap M_{\varrho \setminus \varrho_\ell}(F)$. Hence $g \in \mathcal{G}_Z$. This means that in this case $\text{wt}(I)$ is equal to the weight of the element g from $I \cap \mathcal{G}_Z$, and so condition (i) holds.

(For the proof of (ii) here we add that it follows from the maximality of $\text{wt}(I)$ and Lemma 4.2 that $\text{wt}(g) = M_Z$. Therefore $g \in \mathcal{G}_{\max}^Z$, as required for (ii).)

Case 2. $g \in \text{Ann}_r(M_\varrho(F)) \setminus \text{Ann}_\ell(M_\varrho(F))$. Then $ge_{i,t} \neq 0$ for some $(i, t) \in \varrho$. Obviously, $\text{wt}(ge_{i,t}) \leq \text{wt}(g)$. By the minimality of the weight $\text{wt}(g)$, we get $\text{wt}(ge_{i,t}) = \text{wt}(g)$, because $ge_{i,t} \in I$. Therefore there exists a subset $S \subseteq \varrho$ such that $g = \sum_{j \in S} f_j e_{j,i}$, where $0 \neq f_j \in F$. Clearly, $|S| = \text{wt}(g)$. It follows that $g \in \mathcal{G}_L$. Therefore in this case I has an element of weight $\text{wt}(I)$ belonging to \mathcal{G}_L , and so condition (i) holds.

(For the proof of (ii) here we add that the maximality of $\text{wt}(I)$ and Lemma 4.3 imply that $|S| = M_L = M$. Therefore $(S, i) \in \mathcal{P}_{|M|}$ and $\text{wt}(g) = M$, as required for (ii).)

Case 3. $g \in \text{Ann}_\ell(M_\varrho(F)) \setminus \text{Ann}_r(M_\varrho(F))$. The proof in this case is dual to that for Case 2, and so we omit it.

Case 4. $g \notin \text{Ann}_r(M_\varrho(F)) \cup \text{Ann}_\ell(M_\varrho(F))$. By Lemma 4.1, we get $g \notin M_{\varrho \setminus \varrho_\ell}(F) \cup M_{\varrho \setminus \varrho_\ell}(F)$. Hence there exist $(a, b), (c, d) \in \varrho$ such that $e_{a,b} g e_{c,d} \neq 0$. However, $e_{a,b} g e_{c,d} \in I$ and $\text{wt}(e_{a,b} g e_{c,d}) = 1$. This contradicts the assumption that $\text{wt}(I) > 1$ and completes the proof. \square

Acknowledgements

The authors are grateful to the referee for corrections and suggestion of a possible direction for future research.

References

- [1] T. A. Abrahamsen and O. Nygaard, 'On λ -strict ideals in Banach spaces', *Bull. Aust. Math. Soc.* **83**(2) (2011), 231–240.
- [2] D. D. Anderson and S. Chun, 'Some remarks on principal prime ideals', *Bull. Aust. Math. Soc.* **83**(1) (2011), 130–137.
- [3] R. R. Andruszkiewicz, 'On maximal essential extensions of rings', *Bull. Aust. Math. Soc.* **83**(2) (2011), 329–337.
- [4] F. Baccelli, G. Cohen, G. J. Olsder and J.-E. Quadrat, *Synchronization and Linearity: An Algebra for Discrete Event System* (Wiley Interscience, New York, 1992).
- [5] A. M. Bagirov and J. L. Yearwood, 'A new nonsmooth optimization algorithm for minimum sum-of-squares clustering problems', *European J. Oper. Res.* **170** (2006), 578–596.
- [6] D. M. Clark and B. A. Davey, *Natural Dualities for the Working Algebraist* (Cambridge University Press, Cambridge, 1998).
- [7] S. Dăscălescu and L. van Wyk, 'Do isomorphic structural matrix rings have isomorphic graphs?', *Proc. Amer. Math. Soc.* **124** (1996), 1385–1391.
- [8] V. Drensky and P. Lakatos, 'Monomial ideals, group algebras and error-correcting codes', in: *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, Lecture Notes in Computer Science, 357 (ed. T. Mora) (Springer, Berlin, 1989), pp. 181–188.
- [9] T. Dube, 'Real ideals in pointfree rings of continuous functions', *Bull. Aust. Math. Soc.* **83**(2) (2011), 338–352.
- [10] D. Easdown, J. East and D. G. FitzGerald, 'A presentation of the dual symmetric inverse monoid', *Internat. J. Algebra Comput.* **18** (2008), 357–374.
- [11] D. Y. Gao, *Duality Principles in Nonconvex Systems: Theory, Methods and Applications* (Kluwer Academic Publishers, Dordrecht, 1999).
- [12] D. Y. Gao, *Complementarity, Duality and Symmetry in Nonlinear Mechanics* (Kluwer Academic Publishers, Dordrecht, 2004).
- [13] J. S. Golan, *Semirings and Their Applications* (Kluwer Academic Publishers, Dordrecht, 1999).
- [14] J. S. Golan, *Semirings and Affine Equations over them: Theory and Applications* (Kluwer Academic Publishers, Dordrecht, 2003).
- [15] B. W. Green and L. van Wyk, 'On the small and essential ideals in certain classes of rings', *J. Aust. Math. Soc. Ser. A* **46** (1989), 262–271.
- [16] A. V. Kelarev, 'On classical Krull dimension of group-graded rings', *Bull. Aust. Math. Soc.* **55** (1997), 255–259.
- [17] A. V. Kelarev, *Ring Constructions and Applications* (World Scientific, River Edge, NJ, 2002).
- [18] A. V. Kelarev, *Graph Algebras and Automata* (Marcel Dekker, New York, 2003).
- [19] A. V. Kelarev, R. Göbel, K. M. Rangaswamy, P. Schultz and C. Vinsonhaler, *Abelian Groups, Rings and Modules*, Contemporary Mathematics, 273 (American Mathematical Society, New York, 2001).
- [20] A. V. Kelarev and D. S. Passman, 'A description of incidence rings of group automata', *Contemp. Math.* **456** (2008), 27–33.
- [21] A. V. Kelarev and C. E. Praeger, 'On transitive Cayley graphs of groups and semigroups', *European J. Combin.* **24**(1) (2003), 59–72.
- [22] A. Kelarev, J. Ryan and J. Yearwood, 'Cayley graphs as classifiers for data mining: the influence of asymmetries', *Discrete Math.* **309**(17) (2009), 5360–5369.
- [23] A. V. Kelarev and P. Solé, 'Error-correcting codes as ideals in group rings', in: *Abelian Groups, Rings and Modules*, Contemporary Mathematics, 273 (eds. A. V. Kelarev, R. Göbel,

- K. M. Rangaswamy, P. Schultz and C. Vinsonhaler) (American Mathematical Society, New York, 2001), pp. 11–18.
- [24] A. V. Kelarev, P. A. Watters and J. L. Yearwood, ‘Rees matrix constructions for clustering of data’, *J. Aust. Math. Soc. Ser. A* **87** (2009), 377–393.
- [25] A. V. Kelarev, J. L. Yearwood and P. W. Vamplew, ‘A polynomial ring construction for classification of data’, *Bull. Aust. Math. Soc.* **79** (2009), 213–225.
- [26] S. Mendes-Gonçalves and R. P. Sullivan, ‘The ideal structure of semigroups of transformations with restricted range’, *Bull. Aust. Math. Soc.* **83**(2) (2011), 289–300.
- [27] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* (Elsevier, Amsterdam, 2005).
- [28] L. van Wyk, ‘Matrix rings satisfying column sum conditions versus structural matrix rings’, *Linear Algebra Appl.* **249** (1996), 15–28.
- [29] J. L. Yearwood and M. Mammadov, *Classification Technologies: Optimization Approaches to Short Text Categorization* (Idea Group Inc., Hershey, Pennsylvania, 2010).
- [30] J. Yearwood, D. Webb, L. Ma, P. Vamplew, B. Ofoghi and A. Kelarev, ‘Applying clustering and ensemble clustering approaches to phishing profiling’, in: *Data Mining and Analytics 2009, Proc. 8th Australasian Data Mining Conference: AusDM 2009, (1–4 December 2009, Melbourne, Australia)*, CRPIT, 101; pp. 25–34.
- [31] D. T. Yost, ‘M-ideals, the strong 2-ball property and some renorming theorems’, *Proc. Amer. Math. Soc.* **81**(2) (1981), 299–303.
- [32] D. T. Yost, ‘Semi-M-ideals in complex Banach spaces’, *Rev. Roumaine Math. Pures Appl.* **29**(7) (1984), 619–623.
- [33] D. T. Yost, ‘Banach spaces isomorphic to proper M-ideals’, *Colloq. Math.* **56**(1) (1988), 99–106.

D. Y. GAO, School of Science, Information Technology and Engineering,
University of Ballarat, P.O. Box 663, Ballarat, Victoria 3353, Australia
e-mail: d.gao@ballarat.edu.au

A. V. KELAREV, School of Science, Information Technology and Engineering,
University of Ballarat, P.O. Box 663, Ballarat, Victoria 3353, Australia
e-mail: a.kelarev@ballarat.edu.au

J. L. YEARWOOD, School of Science, Information Technology and Engineering,
University of Ballarat, P.O. Box 663, Ballarat, Victoria 3353, Australia
e-mail: j.yearwood@ballarat.edu.au