A VECTOR SPACE APPROACH TO THE ROAD COLORING PROBLEM

GREGORY BUDZBAN AND PHILIP FEINSILVER

Abstract. Let G be a strongly connected, aperiodic, two-out digraph with adjacency matrix A. Suppose $A = R + B$ are coloring matrices: that is, matrices that represent the functions induced by an edge-coloring of G. We introduce a matrix $\Delta = \frac{1}{2}(R - B)$ and investigate its properties. A number of useful conditions involving Δ which either are equivalent to or imply a solution to the road coloring problem are derived.

1. Introduction

Adler, Goodwyn, and Weiss [AGW] asked the following question: Suppose you are given a strongly connected directed graph. Is there a way to label the edges so that a synchronizing instruction exists? A precise definition of synchronizing instruction will be provided below. In their paper, Adler and his colleagues showed that aperiodicity of the graph was necessary for such an instruction to exist. The road coloring problem conjectures that aperiodicity is a sufficient condition as well.

The relationship between edge-colored directed graphs and finite automata is well known. (For the details, see [CKK] or [BF]). For this paper, all graphs are assumed to be two-out. Thus, an edge-coloring will induce two transformations on the set of vertices, which we will denote R and B. The semigroup $S = \langle R, B\rangle$, generated by R and B under composition is referred to as a coloring semigroup.

We will work almost exclusively in this paper with the standard matrix representation for finite transformations. That is, the matrix representing R will have a 1 in the i,jth component if R maps i to j. Since no confusion will result, we will denote these matrices as R and B as well. Similarly, the matrix semigroup generated by R and B will also be S. A knowledge of elementary semigroup theory is assumed of the reader, all of which can be found in [CP].

Definition 1.1. Let $S = \langle R, B\rangle$ be a coloring semigroup of a digraph $G = (V, E)$. Any finite transformation semigroup S has a minimal ideal or kernel K, which consists of the elements of minimal rank (see [H]). This common minimal rank is called the rank of the kernel. A synchronizing instruction for an edge-labeled digraph is any transformation of rank one.

Now unsolved for more than thirty years, there have been a number of recent papers on various aspects of the problem. Both Carbone [C] and Jonoska and Suen [JS] generalize a classic result of O'Brien [O]. In [BM] many earlier results, in particular some interesting results of Friedman [F], were translated into the language of semigroup theory and the problem was generalized to periodic graphs. Additional work on the generalized problem can be found in [B] and [HJ]. An interesting subcase when the adjacency matrix has both constant row and column sums was solved by Kari in [K]. In [CKK], Kari and his colleagues extended these techniques, and showed the importance of finding a labeling whose coloring semigroup has a structure known as a right group as its kernel, though they did not use this semigroup terminology in the paper. Continuing the analysis of the role that right groups play in the problem, [BF] showed that the existence of a right group kernel implied a graph was synchronizable and found a sufficient condition for such a structure to exist based on the Lie algebra generated by the matrices induced by the edge labeling.

The structure of this paper will be as follows. After this introduction, we will define a difference matrix in section 2, and show how recolorings of the edges have a simple characterization in terms of this matrix. In section 3, we will show how this difference matrix can be used to characterize the existence of synchronizing instructions. Because of the asymptotic properties of the (normalized) adjacency matrix, the vector space spanned by certain rank one matrices play an important role in the theory. This role is explored in section 4. In section 5, a condition based on a vector space generated in part by powers of Δ is shown to be sufficient to imply a solution to the problem. Finally, in section 6, we will introduce an algorithm for locating a synchronizing coloring, if one exists according to the condition introduced in section 5.

2. The matrix Δ

We will use the following notation. All matrices are $n \times n$, where n is the number of vertices in the graph. The **normalized adjacency matrix** $T = \frac{1}{2}(R + B)$, while the **coloring difference matrix** $\Delta = \frac{1}{2}(R - B)$. Since the graph is strongly connected and aperiodic, $\lim_{n\to\infty} T^n = \Omega$ exists, where Ω is a positive stochastic matrix with identical rows. Each row of Ω is a probability vector which we will refer to as π. In what follows, important structures will be the real vector space spanned by S, denoted $\text{vec}(S)$, and the span of the kernel K, denoted $\text{vec}(K)$.

Any eventual solution to the conjecture must find a systematic method for analyzing the possible recolorings of the graph. To this end, let ϕ_i be the diagonal matrix with -1 in the ith diagonal position, and 1s along the rest of the diagonal. We refer to ϕ_i as the ith elementary recoloring matrix, where an elementary recoloring of vertex i is a recoloring that switches the red and blue labels for the edges leading from i, and leaves all other labels in the graph unchanged. Notice that given the coloring difference matrix $\Delta = \frac{1}{2}(R-B)$, then $\phi_i\Delta$ is the coloring difference matrix induced by an elementary recoloring of i. It is clear any possible difference matrix Δ' is of the form $\Delta' = \phi_{i_1}\phi_{i_2}\ldots\phi_{i_j}\Delta$ where $\phi_{i_1}\phi_{i_2}\ldots\phi_{i_j}$ are elementary recoloring matrices. Because all possible coloring differences matrices are of this form, the following proposition is immediate.

Proposition 2.1. *All coloring difference matrices have the same rank.*

3. The matrix Δ annihilates only rank one kernels

Denote the all 1's (column) vector by $\mathbf{u}$. We say that a word in S has "length n" if it can be written as a product of n R's and B's.

Lemma 3.1. *Let the vector* $\mathbf{y}$ *satisfy* $\Delta w\mathbf{y} = 0$ *for all words* $w \in S$. *Then* $\mathbf{y}$ *is a constant vector, i.e., a multiple of* $\mathbf{u}$.

Proof. Notice that $\Delta \mathbf{y} = 0$, implies that $R\mathbf{y} = B\mathbf{y}$. Thus, $T\mathbf{y} = \frac{1}{2}(R+B)\mathbf{y} = R\mathbf{y} = B\mathbf{y}$. Inductively, assume that any word w_n of length n satisfies $w_n\mathbf{y} = T^n\mathbf{y}$. Now $Rw_n\mathbf{y} = Bw_n\mathbf{y} = Tw_n\mathbf{y} = T^{n+1}\mathbf{y}$ implies that any word w_{n+1} of length $n+1$ satisfies $w_{n+1}\mathbf{y} = T^{n+1}\mathbf{y}$.

Now let e be an idempotent in the kernel, of length l. Then

$$e\mathbf{y} = T^l\mathbf{y} = e^2\mathbf{y} = T^{2l}\mathbf{y} = \cdots = T^{nl}\mathbf{y}$$

for all positive integers n. Letting $n \to \infty$ we get $\Omega\mathbf{y} = c\mathbf{u}$, where the constant $c = \pi \cdot \mathbf{y}$. In other words, for every i in the range of e, $y_i = c$. It follows from Theorem 4 in [BM] that every vertex i is in the range of some idempotent e in the kernel, thus $\mathbf{y} = c\mathbf{u}$. □

Corollary 3.2. *Let the vector* $\mathbf{y}$ *satisfy* $\Delta T^n\mathbf{y} = 0$ *for all integer* $n \geq 0$. *Then* $\mathbf{y}$ *is a constant vector, i.e., a multiple of* $\mathbf{u}$.

Proof. As above, from $\Delta\mathbf{y} = 0$, we get $R\mathbf{y} = B\mathbf{y} = T\mathbf{y}$. Proceeding inductively as in the above proof $w_n\mathbf{y} = T^n\mathbf{y}$ for all words of length n implies that $\Delta w_n\mathbf{y} = 0$ or

$$Rw_n\mathbf{y} = Bw_n\mathbf{y} = Tw_n\mathbf{y} = T^{n+1}\mathbf{y}.$$

Hence, $\Delta w\mathbf{y} = 0$ for all words w and by the preceding Lemma, $\mathbf{y}$ is constant. □

Corollary 3.3. $\Delta k = 0$ *for all* $k \in K$ *if and only if* K *is rank one.*

Proof. Suppose $\Delta k = 0$ for all $k \in K$. Let e be an idempotent in K. Then for all words w, $we \in K$, so that $\Delta we = 0$, for all w. By the Lemma, the columns of e are constant vectors. Since any row has exactly one nonzero entry, one column is all ones and the rest are all zeros. The converse is immediate. □

Corollary 3.4. $\Delta e = 0$ *for all idempotents in* K *if and only if* K *is rank one.*

Proof. For any $k \in K$, since K is completely simple and, therefore, a union of isomorphic groups, there is an idempotent such that $k = ek$. Hence $\Delta k = \Delta ek = 0$ and by the previous Corollary, K is rank one. □

4. Special rank one matrices in Vec(S)

In this section we will focus on the vector spaces generated by either the matrix semigroup or its kernel. Any rank-one matrix with identical rows in vec(S) will be referred to as rank-one special, or ROSP. A ROSP in vec(K) will be a KROSP. Note, that the ROSP's form a vector space of dimension at most n. In general, we will identify a ROSP with its first row. Observe that if M is any ROSP, then MN is a ROSP for any matrix N in vec(S). In addition, if w is any word in S, $wM = M$. That is, M is a right zero for elements in S.

A *range vector* is a vector consisting of 0's and 1's with 1's corresponding to the range of an idempotent e in K, that is, a vector whose components indicate the non-zero columns of e. A range matrix is a KROSP whose rows equal some range vector. If there is no chance of confusion, for simplicity, we identify the range vector with its corresponding KROSP.

Proposition 4.1. *The vector space spanned by the KROSPs equals:*

1. *Span of* $\{\Omega w : w \in S\}$.
2. *Span of* $\{Mk : k \in K\}$, *where* M *is any ROSP in vec(S) satisfying* $M\mathbf{u} \neq 0$.
3. *Span of the range matrices.*

Proof. For #1, first note that Ω is an element of vec(K), and therefore a KROSP (see the proof of Theorem [7] in [BM]). Since K is an ideal, $kw \in K$, for any $w \in S$. So $\Omega w \in$ vec(K) and hence a KROSP. Equality will follow from #2.

For #2, for $k \in K$, $Mk \in$ vec(K) and so is a KROSP. If the set $\{Mk : k \in K\}$ does not span the KROSPs, then there is a KROSP, X, with first row $\mathbf{x}$ such that

$$Mk\mathbf{x}^t = 0, \quad \text{for all} \quad k \in K.$$

Let $M\mathbf{u} = c\mathbf{u}$, with $c \neq 0$. Since X itself is a KROSP, $cXX^t = MXX^t = 0$. So $X = 0$ and the result follows.

For #3, by the proof of Theorem 7 in [BM], πk is a multiple of a range vector for every $k \in K$. So the result follows from #2, taking $M = \Omega$. □

The constant functions play a special role in any investigation of the road coloring problem. Let C_i denote the matrix of the constant function i; that is for C_i, the ith column is all ones. If S contains the constants C_i, they must be in K due to minimality of the rank of the kernel. Clearly, any element in the linear span of the constants C_i, $i = i, \ldots, n$ is a KROSP. So the KROSPs have dimension n if and only if vec(K) contains the constants C_i. In fact, we can say more.

Proposition 4.2. *The KROSPs have dimension n if and only if the rank of the kernel is one.*

Proof. One direction is immediate. For the other, assume that the KROSPs have dimension n. By the above, vec(K) contains the constants C_i.

Now $C_i = \sum_{k \in K} \alpha_k k$. Choose an idempotent $e \in K$. Then $eC_ie = \sum_{k \in K} \alpha_k eke$. Since K is a completely simple semigroup (see [CP]), then $eC_ie = C_j$ is in vec(G) where

$G = eKe$ is the subgroup for which e is the local identity element (again [CP]), and $ie = j$.

Suppose the rank of K is r. Then G is a group of block permutation matrices of rank r. In particular, the matrices of G have r nonzero columns. Notice $\Omega C_j = C_j$. On the other hand, $\Omega C_j = \sum_{k \in K} \alpha_k \Omega eke = C_j$. But, as stated above, the proof of Theorem 7 in [BM] shows that, for each $k \in K$, Ωeke is a fixed ROSP whose rows are multiples of the range vector for the group G. Thus $C_j = \Omega eke \sum_{k \in K} \alpha_k$ where Ωeke is a ROSP with r non-zero columns. But this is only possible if $r = 1$. □

5. The Rank Condition

Recall that π denotes the invariant distribution, the row vector satisfying $\pi T = \pi$, with $\pi \cdot \mathbf{u} = 1$.

Definition 5.1. Let d denote the rank of Δ. The *rank condition* holds for Δ if the span of $\{\pi\Delta, \pi\Delta^2, \ldots, \pi\Delta^d\}$ is a d-dimensional space, in other words, these vectors span the range (row space) of Δ.

As a consequence of the rank condition, any vector $\mathbf{y}$ satisfying $\Omega\Delta^i\mathbf{y} = 0$, $1 \le i \le d$, satisfies $\Delta\mathbf{y} = 0$ as well.

Once again, the proof of Theorem 7 in [BM] shows that πk is a multiple of a range vector for every $k \in K$. Specifically, if k is in K, then πk is a multiple of the range vector corresponding to the nonzero columns of k, and that multiple depends only on the collection of non-zero columns, that is, the range of k. Observe that for any $k \in K$, Rk and Bk are elements of K with the same range. Hence, $\Omega Rk = \Omega Bk$, or $\Omega\Delta k = 0$ for all $k \in K$.

Assume the rank condition. Let e be any idempotent in K. Let w be any word in S. Then Rwe and Bwe are elements of K with the same range. Hence, $\Omega Rwe = \Omega Bwe$ is a multiple of the range matrix corresponding to e. In other words, $\Omega\Delta we = 0$ for any w in S. But then $\Omega\Delta^i e = 0$, for all $i \ge 0$, since powers of Δ are in $\mathrm{vec}(S)$. From the rank condition, we have $\Delta e = 0$, for all idempotents $e \in K$. By Corollary 3.4, K is rank one.

We thus have the following

Theorem 5.2. *The rank condition implies that the kernel is rank one.*

The following Corollary is immediate, and is stated here only for completeness.

Corollary 5.3. *Let G be a strongly connected, aperiodic graph. Suppose G admits a coloring semigroup $S = \langle R, B \rangle$ such that the rank condition holds for the difference matrix $\Delta = \frac{1}{2}(R - B)$ associated with the coloring. Then S contains a synchronizing instruction.*

6. Coloring and recoloring

The rank condition **not** holding means that for the given coloring ϕ, there is a polynomial p_ϕ, of degree at most d, with $p_\phi(0) = 0$, such that $\Omega p_\phi(\phi\Delta) = 0$.

Introduce variables ε_i, $\varepsilon_i^2 = 1$. Then generically

$$\phi = \operatorname{diag}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = \begin{pmatrix} \varepsilon_1 & 0 & \cdots & 0 \\ 0 & \varepsilon_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \varepsilon_n \end{pmatrix}$$

We want the matrices $\{\Omega(\phi\Delta), \Omega(\phi\Delta)^2, \dots, \Omega(\phi\Delta)^d\}$ to be linearly independent. Since these are all KROSPS, we form a matrix H whose i^{th} row is $\pi(\phi\Delta)^i$. Then the rank of H is d if and only if $\det HH^* \neq 0$.

Remark. It is readily seen that Ω is always linearly independent of $\{\Omega\Delta^i\}_{i\geq 1}$, just apply any linear combination to the all 1's vector. So if $d = n-1$, you can include π as the first row of H and get an $n \times n$ matrix. Then you only need to check $\det H$.

Note that as a function of the ε-variables, $D = D(\phi) = \det HH^*$ is multilinear, since all squares resolve to 1. For each ε_i, we thus have the partial derivative of D with respect to ε_i by taking all terms containing that variable and factoring it out. Combining these partials into a single vector for $1 \leq i \leq n$, gives $\nabla_\varepsilon D(\phi)$. It follows that unless $D(\phi)$ is identically zero, both $D(\phi)$ and $\nabla_\varepsilon D(\phi)$ cannot vanish for all values ± 1 of the variables $\{\varepsilon_1, \dots \varepsilon_n\}$.

The following procedure will find the coloring that has a synchronizing instruction under the assumption that the rank condition holds for at least one coloring.

Procedure. Initialize $\phi_0 = \operatorname{diag}(1, 1, \dots, 1)$.

1. Form HH^*.
2. Calculate $D(\phi) = \det HH^*$.
3. If $D(\phi_0) \neq 0$, **DONE** else continue.
4. Compute $\nabla_\varepsilon D(\phi_0)$.
5. If any component of $\nabla_\varepsilon D(\phi_0)$ is nonzero, recolor the corresponding vertex, **DONE** else continue.
6. If all components of the gradient are zero, randomly pick a new ϕ_0. Repeat from Step 3.

References

[AGW] R. L. Adler, L. W. Goodwyn and B. Weiss, *Equivalence of topological Markov shifts*, Israel J. Math. **27** (1977), 49–63.

[B] G. Budzban, *Semigroups and the generalized road coloring problem*, Semigroup Forum **69** (2004), 201–208.

[BF] G. Budzban and P. Feinsilver, *Completely Simple Semigroups, Lie Algebras, and the Road Coloring Problem*, accepted for publication, Semigroup Forum.

[BM] G. Budzban and A. Mukherjea, *A semigroup approach to the road coloring problem*, Contemp. Math., vol. 261, Amer. Math. Soc., Providence, RI, 2000, pp. 195–207.

[C] A. Carbone, *Cycles of relatively prime length and the road coloring problem*, Israel J. Math. **123** (2001), 303–316.

[CKK] K. Culic, J. Kari, and J. Karhumaki, *A note on synchronized automata and road coloring problem*, Internat. J. of Found. Comp. Sci. **3** (2002), 459–471.

[CP] A.H. Clifford and G.B. Preston, "The Algebraic Theory of Semigroups", vol. 1, Amer. Math. Soc., Providence, RI, 1961.

[F] J. Friedman, *On the Road Coloring Problem*, Proc. Amer. Math. Soc. **110** (1990), 1133–1135.

[HJ] R. Hegde and K. Jain, *A min-max theorem about the road coloring conjecture*, submitted to Discrete Math.

[H] G. Högnäs, *Random Semigoup Acts on a Finite Set*, J. Australian Math. Soc. **23** (1977), 481–498.

[JS] N. Jonoska and S. Suen, *Monocyclic Decomposition of Graphs and the Road Coloring Problem*, Congr. Numer. **110** (1995), 201–209.

[K] J. Kari, *Synchronizing finite automata on Eulerian digraphs*, Theoretical Computer Science **295** (2003) 223–232.

[O] G. L. O'Brien, *The Road Coloring Problem*, Israel Journal of Mathematics **39** (1981), 145–154.

Department of Mathematics, Mailcode 4408, 1245 Lincoln Drive, Southern Illinois University at Carbondale, Carbondale, IL 62901

Department of Mathematics, Mailcode 4408, 1245 Lincoln Drive, Southern Illinois University at Carbondale, Carbondale, IL 62901