Random Walks on Digraphs

J. J. P. Veerman*

October 23, 2017

1 Introduction

Let $V = \{1, \cdots n\}$ be a vertex set and S a non-negative row-stochastic matrix (i.e. rows sum to 1). V and S define a digraph $G = G(V, S)$ and a directed graph Laplacian $\mathcal{L}$ as follows. If $(S)_{ij} > 0$ (in what follows we will leave out the parentheses) there is a directed edge $j \to i$. Thus the ith row of S identifies the edges coming *into* vertex i and their weights. This set of vertices are collectively the neighbors of i, and is denoted by $\mathcal{N}_i$. The diagonal elements S_{ii} are chosen such that each row sum equals 1. In particular, if a vertex i has no incoming edges, we choose $S_{ii} = 1$. For the purposes of this work, we define the *Laplacian* by

$$\mathcal{L} \equiv I - S \ ,$$

where I is the identity matrix.

It turns out, perhaps somewhat confusingly, that depending on the application one is interested in, sometimes S and $\mathcal{L}$ are the natural objects of study, and at other times it may be better to look at their transpose S^T and $\mathcal{L}^T$. As an example of the first, consider a very simple *consensus* model, where the components $x \in \mathbb{R}^n$ are individual approvals (or lack thereof) of some item or idea. Suppose that one's opinion changes with a rate given by some linear function of the perceived differences of opinion with one's neighbors. This naturally leads to the model

$$\dot{x} = g\mathcal{L}x \ , \tag{1.1}$$

where g is some arbitrary real parameter. What is important here, is that every opinion depends on *incoming* opinions:

$$\dot{x}_k = g \sum_{j \in \mathcal{N}_k} w_{kj}(x_k - x_j) \ ,$$

where the w_{kj} are non-negative weights with row-sum one. One sees that in effect, the rate of change depends on the difference of one's opinion with a weighted average of *incoming* opinions. Thus $\mathcal{L}$ has row-sum zero and $S = I - \mathcal{L}$ is *row-stochastic*. The same things holds in the study of flocks and more details for that model can be found in [2].

*Fariborz Maseeh Dept. of Math. and Stat., Portland State Univ., Portland, OR, USA

On the other hand, consider next a random walk on the same graph with n vertices. Given the probabilities $p(i)$ that a walker is at vertex i, the probabilities $p'(j)$ that the walker is at vertex j in the next time step. Since the probabilities must always be non-negative and sum to 1, the random walker is a map from the $(n-1)$-dimensional simplex to itself. Let us suppose, as before, that map is linear and represented by a $n \times n$ matrix A. So suppose $p = (1, 0, \cdots 0)^T$, then p', the first column of A, is a column of non-negative probabilities summing to one. In essence, at very vertex, the *outgoing* probabilities must sum to one, and A is *column-stochastic*. Thus, in this case, it is natural to consider $A \equiv S^T$ as the linear map determining the random walk. It is crucial to realize that according to this definition the random walker moves *in the direction opposite to the edge arrows* defined above.

The aim of this note is to characterize the null spaces of both $\mathcal{L}$ and $\mathcal{L}^T$, and to describe their relation to the dynamical processes just outlined. We build on work done in [1] and [2] where the role of the null space in the study of flocking was described. In section 2 we summarize the results of these papers. The main results of this work concern the random walk and its relation to the null space of $\mathcal{L}^T$. These follow in Section 3. In Section 4 we illustrate our results with two simple examples.

2 Prior Results

In [1] matrices M of the form

$$M = D - DS$$

were considered, where D is a non-negative diagonal matrix and S is row stochastic. The special case $\mathcal{L}$ arises by choosing $D = I$. We will assume that from here on. The following definitions (with the exception of Definition 2.3) and results are taken from [1] and [2].

Definition 2.1 *Given any real $N \times N$ matrix $M = D - DS$, we denote by $G(V, S)$ the directed graph with vertices $1, \cdots n$ and an edge $j \to i$ whenever $S_{ij} \neq 0$. For each vertex i, set $\mathcal{N}_i := \{j | j \to i\}$. We write $j \rightsquigarrow i$ if there exists a directed path in G_S from vertex j to vertex i. Furthermore, for any vertex j, we define $R(j)$ to be the set containing j and all vertices i such that $j \rightsquigarrow i$. We refer to $R(j)$ as the reachable set of vertex j.*

Definition 2.2 *A set R of vertices in a graph will be called a reach if it is a maximal reachable set; in other words, R is a reach if $R = R(i)$ for some i and there is no j such that $R(i) \subset R(j)$ (properly). Since our graphs all have finite vertex sets, such maximal sets exist and are uniquely determined by the graph. For each reach R_i of a graph, we define the exclusive part of R_i to be the set $H_i = R_i \backslash \cup_{j \neq i} R_i$. Likewise, we define the common part of R_i to be the set $C_i = R_i \backslash H_i$.*

Note that, by definition, the pairwise empty intersection of two exclusive sets is empty: $H_i \cap H_j = \emptyset$ if $i \neq j$. The common sets can, however, intersect. Note further that each reach R contains at least one vertex r such that its reachable set $R(r)$ equals the entire reach. Such a vertex is called a root of R. By definition, any root must be contained in the exclusive part of its reach.

Definition 2.3 *Given a digraph G. Then each reach R_i contains a set of roots P_i. The set P_i is called the root set and is contained in H_i.*

Theorem 2.4 *Suppose $M = D - DS$, where D is a nonnegative $N \times N$ diagonal matrix and S is stochastic. Suppose G_S has k reaches, denoted R_1 through R_k, where we denote the exclusive and common parts of each R_i by H_i, C_i respectively. Then the null space of M has a basis γ_1, γ_2, ... γ_k in $\mathbb{R}^n$ whose elements satisfy:*
(i) $\gamma_i(v) = 1$ for $v \in H_i$;
(ii) $\gamma_i(v) \in (0, 1)$ for $v \in C_i$;
(iii) $\gamma_i(v) = 0$ for $v \notin R_i$;
(iv) $\sum_i \gamma_i = \mathbf{1}_n$.

Notice that Theorem 2.5 does not completely determine the basis of the null space. The values of the basis vectors on the common depend on the weights of the relevant edges. This will be illustrated with an example in Section 4.

Theorem 2.5 *The eigenvalue 0 of a Laplacian matrix of the form $D - DS$ with k reaches has algebraic and geometric multiplicity k.*

Theorem 2.6 *Any nonzero eigenvalue of a Laplacian matrix of the form $D - DS$, where D is nonnegative diagonal and S is stochastic, has (strictly) positive real part.*

The consequences of this for the simple consensus model mentioned earlier are easy to describe. Let Δ be the subspace of $\mathbb{R}^n$ spanned by all the (generalized) eigenspaces of $\mathcal{L}$ other than its null space and denote by $\{\gamma_i\}_{i=k+1}^n$ a basis of Δ. Then the vectors $\{\gamma_i\}_{i=1}^n$ form a basis for $\mathbb{R}^n$. The initial condition $x(0)$ can be written in terms of this basis as

$$x(0) = \sum_{i=1}^{n} \alpha_i \gamma_i \; .$$

Set $g < 0$. By Theorem 2.6, the non-zero eigenvalues of $g\mathcal{L}$ all have negative real part. Thus the solution of the differential equation (1.1) is:

$$x(t) = \sum_{i=1}^{k} \alpha_i \gamma_i + \mathcal{R}(t) \; ,$$

where $\mathcal{R}(t)$ is as function that decreases to 0 exponentially fast as time t tends to infinity (though it may have large transients). This means that, for large t, the opinion vector x is entirely determined by the null space of $\mathcal{L}$. The conclusion for the more complicated flocking models is very similar. We refer the interested reader to [2].

3 Random Walks on Directed Graphs

Denote the transpose of S by A: $A = S^T$. Everything else is as defined in Section 2. In particular, edges $i \to j$ and directed paths $i \rightsquigarrow j$ have directions determined by S.

Definition 3.1 *Let G be the digraph with (weighted) adjacency row stochastic matrix S. T_S is the random walk on the digraph G given by the transition probabilities: prob$(j \to i) = A_{ij} = S_{ji}$.*

Recall that the edges have the opposite direction, i.e. $i \to j$ is an edge if A_{ij} =prob$(j \to i) > 0$. We will abbreviate T_S with T since no confusion is likely.

Definition 3.2 *A probability vector or a (discrete)* measure *is a vector in* $\mathbb{R}^n$ *such that for all* i, $p(i) \geq 0$ *and* $\sum_i p(i) = 1$. *The* support, supp(p), *of the measure* p, *is the set of vertices on which* p *takes a positive value.*

To conform to the formal definition of probability measure, we note that $p(\emptyset) = 0$, and for any vertex set $W \subseteq V(G)$, $p(W) = \sum_{i \in W} p(i)$.

Definition 3.3 *Let* T *be a random walk. The* push forward T_*p *of the measure* p *is given by*

$$(T_*p)(i) = \sum_j prob(j \to i)p(j) = \sum_j A_{ij}p(j) \ ,$$

The pull back T^*p *of the measure* p *is given by*

$$(T^*p)(i) = \sum_j prob(i \to j)p(j) = \sum_i A_{ji}p(j) \ ,$$

Therefore if the probabilities at time step 0 are given by the vector p, then at time step 1 they are given by the vector $T_*p = Ap$. Note that this corresponds to *left* multiplication of S by p. It is easy to directly verify that Ap is a probability vector if p is a probability vector. First of all, if p is a probability vector, then

$$(Ap)(i) = \sum_j A_{ij}p(j) \geq 0 \ ,$$

and then, of course,

$$\sum_i (Ap)(i) = \sum_i \sum_j A_{ij}p(j) = \sum_j \left(\sum_i A_{ij} \right) p(j) = 1 \ .$$

So that Ap satisfies Definition 3.2.

Definition 3.4 T *has a forward invariant probability measure* p *if* $Ap = p$. $K \subseteq V(G)$ *is a forward invariant set under* T, *if* $p|_{K^c} = 0$ *implies* $Ap|_{K^c} = 0$, *where* K^c *is the complement of* K *in* $V(G)$.

We now use the notation of the previous section.

Lemma 3.5 *Given a random walk random walk* T, *every exclusive set* H_i *and its root set* P_i *are forward invariant sets under* T.

Proof: Suppose C_i is the common part of the reach corresponding to H_i. A walker landing in H_i can only leave the exclusive part if the graph G has an edge in the opposite direction (Definition 3.1). This contradicts Definitions 2.2 and 2.3.

Similarly, a walker can only leave P_i if the graph G has edge from $j \in V(G) \backslash P_i$ into P_i. But this would mean that j is a root of R_i, which is a contradiction. ∎

Theorem 3.6 *Let G be a graph with Laplacian $\mathcal{L} = I - S$ with k reaches as in Theorem 2.4. Given $j \in V(G)$ and $m \in \{1, \cdots k\}$. The probability that a random walker under T starting at $j \in V$ ends up in P_m equals $\gamma_m(j)$.*

Proof: Let $q(j)$ be the probability that a random walker starting at $j \in V$ reaches P_m for some fixed m. Then $q : V \to [0, 1]$ is well-defined and is constant in time. Since, by Lemma 3.5, P_m is forward invariant, $q(j)$ is also equal to the probability that the walker starting at j ends up and stays in P_m.

The probability $q(j)$ concerns the future (under T) of the walker on j. Therefore it is equal to the appropriately weighted average of $q(i)$ of j's successors under T. Thus

$$q(j) = \sum_i \text{prob}(j \to i) q(i) = \sum_j A_{ij} q(i) \ .$$

From this we conclude:

$$q = A^T q \quad \Longleftrightarrow \quad q = Sq \quad \Longleftrightarrow \quad \mathcal{L} q = (I - S)q = 0$$

This proves that q is in the null space of M:

$$q(j) = \sum_i \alpha_i \gamma_i(j)$$

However, again by Lemma 3.5, if j is a vertex in P_m, then $q(j) = 1$, and if j is in any P_j with $j \neq m$, then $q(j) = 0$. Thus $\alpha_m = 1$ and $\alpha_\ell = 0$ if $\ell \neq m$. ■

Lemma 3.7 *Let G be a digraph with reach R consisting of an exclusive part H, containing the root set P, and a common part C. Under the random walk T on G, there is a unique invariant measure p whose support equals P.*

Proof: Consider a reach R with its root set P and denote the vertex set $R \backslash P$ by Y and the vertex set $V \backslash R$ by Z. Since directed paths in G cannot leave the reach R, we have $A_{PZ} = A_{YZ} = 0$. Similarly, directed paths in G cannot go from Z to P, nor from Y to P (see Lemma 3.5). Therefore, upon permuting vertices, the matrix A equals

$$A = \begin{pmatrix} A_{PP} & A_{PY} & \mathbf{0} \\ \mathbf{0} & A_{YY} & \mathbf{0} \\ \mathbf{0} & A_{ZY} & A_{ZZ} \end{pmatrix} \ .$$

The matrix A_{PP} is the transpose of the matrix S_{PP}. Clearly P has at least one root v. Lemma 3.1 in [1] shows that the eigenvalue 1 in S_{PP} has algebraic and geometric multiplicity one. Its transpose A_{PP} has the same characteristic polynomial, its eigenvalue 1 also has algebraic multiplicity 1 (and therefore geometric multiplicity 1). Similarly, the proof of Theorem 2.7 in [1] establishes that the spectral radius of S_{YY} is strictly less than 1, and the same holds for A_{YY}.

We solve for p in $Ap = p$, where $p = (a_P, b_Y, c_Z)^T$. This gives

$$A_{PP} a_P + A_{PY} b_Y = a_P \ , \quad A_{YY} b_Y = b_Y \quad \text{and} \quad A_{YZ} b_Y + A_{ZZ} c_Z = c_Z \ .$$

The middle equation can only be satisfied if $b_Y = 0$. This shows that the support of p is in P.

Now we assume that there is a vertex $k \in P$ such that $p(k) = 0$. Then

$$0 = p(k) = \sum_j A_{kj} p(j) \ .$$

Since k is a root, it has outgoing edges. So there must be j such that $A_{kj} > 0$. Since p is a probability measure, its components are non-negative. Therefore $p(j)$ must be zero. By induction, one proves that for all vertices i such that $k \rightsquigarrow i$, $p(i) = 0$. Since k is a root, that implies that $p = 0$ on the entire reach R, which is absurd. Therefore $p(k) > 0$. ■

Theorem 3.8 *Let G be a graph with Laplacian $\mathcal{L} = I - S$ with k reaches as in Theorem 2.4. Then the eigenvalue 0 of $\mathcal{L}^T$ has algebraic and geometric multiplicity k. Furthermore, the null space of $\mathcal{L}^T$ has a basis $\bar{\gamma}_1$, $\bar{\gamma}_2$, ... $\bar{\gamma}_k$ in $\mathbb{R}^n$ whose elements satisfy:*
(i) For all $i \in \{1, \cdots k\}$ and all $v \in \{1, \cdots n\}$: $\bar{\gamma}_i(v) \geq 0$;
(ii) $\sum_{v \in P_i} \bar{\gamma}_i(v) = 1$;
(iii) $\bar{\gamma}_i(v) = 0$ for $v \notin P_i$.
Proof: Denote the union $\cup_{i=1}^k R_i \backslash P_i$ by Z. Permuting rows and columns, we can write the matrix A in block diagonal form:

$$A = \begin{pmatrix} A_{P_1P_1} & \mathbf{0} & \cdots & \mathbf{0} & A_{P_1Z} \\ \mathbf{0} & A_{P_2P_2} & \cdots & \mathbf{0} & A_{P_2Z} \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ \mathbf{0} & \cdots & \mathbf{0} & A_{P_kP_k} & A_{P_kZ} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & A_{ZZ} \end{pmatrix} .$$

Each of the diagonal blocks, except A_{ZZ}, is rooted and so has eigenvalue 1 with algebraic and geometric multiplicity 1. As before, A_{ZZ} has spectral radius less than 1. The characteristic polynomial of A is the product of the characteristic polynomials of the diagonal blocks, and the result follows. ■

Notice that, just like Theorem 2.5, Theorem 3.8 does not completely determine the null space. The examples in the next section will show that indeed random walks on a given digraph but with different weights on the edges can have non-trivially different invariant measures.

4 The Kernel of $\mathcal{L}$ Versus the Kernel of $\mathcal{L}^T$

We saw that the components $\gamma_i(j)$ of the vectors γ_i equal the probability that a random walker starting at j reaches H_i. In contrast, the null space $\bar{\mathcal{E}}_0$ of $\mathcal{L}^T = I - A$ is spanned by vectors $\{\bar{\gamma}_i\}_{i=1}^k$ which give the different invariant measures with support in R_i associated with the random walk T.

We will now show that the role of these invariant measures $\bar{\gamma}_i$ in the random walk is similar to that of the vectors γ_i in the consensus model discussed at the end of Section 2. Let $\bar{\Delta}$ be the subspace of $\mathbb{R}^n$ spanned by all the (generalized) eigenspaces of $\mathcal{L}^T$ other than its null space and denote by $\{\bar{\gamma}_i\}_{i=k+1}^n$ a basis of $\bar{\Delta}$. Then the vectors $\{\bar{\gamma}_i\}_{i=1}^n$ form a basis for $\mathbb{R}^n$. Clearly, any probability vector $p^{(0)}$ can be written uniquely as

$$p^{(0)} = \sum_{i=1}^n \alpha_i \bar{\gamma}_i \ .$$

Lemma 4.1 *Each of the basis vectors of* $\{\bar{\gamma}_i\}_{i=k+1}^n$ *of* $\bar{\Delta}$ *has the property that the sum of its components equals zero.*

Proof: First, assume that v is an eigenvector of $\mathcal{L}^T$ with eigenvalue $\lambda \neq 1$. We obtain

$$\sum_{i,j} A_{ij} v(j) = \lambda \sum_i v(i) \quad \Longrightarrow \quad (1-\lambda) \sum_i v(i) = 0 \ ,$$

because A is column stochastic. Thus $\sum_i v(i) = 0$. Next, suppose the lemma is false for some generalized eigenvector w of $\mathcal{L}^T$. Then, by Theorem 3.8, the associated eigenvalue must be different from 1. Thus we have a situation where there is a vector v with $\sum_i v(i) = 0$ and

$$(A - \lambda I) w = v \ .$$

Taking the sum of the components on both sides yields the result. ∎

Theorem 4.2 *Let* G *be a graph with Laplacian* $\mathcal{L} = I - S$ *with* k *reaches as in Theorem 2.4. Let* $p^{(0)} = \sum_{i=1}^n \alpha_i \bar{\gamma}_i$ *be the initial probability distribution. Assume* S *no eigenvalues with modulus 1, except* $\lambda = 1$*. Then the probability measure at time step* ℓ*,* $p^{(\ell)}$*, is given by:*

$$p^{(\ell)} = A^\ell p^{(0)} = \sum_{i=1}^k \alpha_i \bar{\gamma}_i + A^\ell \delta \ ,$$

for some $\delta \in \bar{\Delta}$*, and* $\left|A^\ell \delta\right|$ *tends to zero exponentially fast as* ℓ *tends to zero.*

Proof: Lemma 4.1 says that $\sum_j \bar{\gamma}_i(j) = 0$ for $i > k$. Therefore $\sum_{j=1}^n \sum_{i=1}^k \alpha_i \bar{\gamma}_i(j) = 1$. Furthermore, all components are non-negative. This shows that $\sum_{i=1}^k \alpha_i \bar{\gamma}_i$ is a probability measure.

By Gershgorin's theorem, eigenvalues of $A = S^T$ have modulus at most 1. The eigenvalue 1 has algebraic multiplicity k and the corresponding eigenspace is spanned by $\{\bar{\gamma}_i\}_{i=1}^k$. All other eigenvalues must have modulus strictly less than 1. Thus, setting $\delta = \sum_{i=k+1}^n \alpha_i \bar{\gamma}_i$, we have $|A^\ell \delta| < K \lambda^\ell$, for some $K > 0$ and some $\lambda \in (0,1)$. ∎

Remark: Even though the convergence is exponentially fast, very large transients may occur for intermediate values of ℓ, especially for graphs that have many vertices and are highly asymmetric.

We give two simple examples of the contrast between the null spaces of $\mathcal{L}$ and $\mathcal{L}^T$. As the first example suppose that $V = \{1, 2\}$ and

$$S = \begin{pmatrix} x & 1-x \\ 1 & 0 \end{pmatrix} \qquad \text{or} \qquad \mathcal{L} = \begin{pmatrix} 1-x & x-1 \\ -1 & 1 \end{pmatrix} \ ,$$

and A is again the transpose of S.

We take $x \in [0,1]$. The eigenvalues of S (and A) are 1 and $x - 1$. Thus when $x = 0$, there is an eigenvalue -1, violating the condition on the eigenvalues of Theorem4.2. The kernel of $\mathcal{L} = I - S$ is the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Indeed graph $G(V,S)$ has one reach and both vertices are (of course) in it.

There is an invariant measure and it is given by $(2-x)^{-1}\begin{pmatrix} 1 \\ 1-x \end{pmatrix}$ which spans the null space of $\mathcal{L}^T = I - S^T$. Note that this also works for $x = 0$ when the "random" walker hops back and forth from 1 to 2 to 1, etcetera. In that case the probability is simply 1/2 that the walker is at vertex 1 and 1/2 that he is at vertex 2.

The next example is modified from [2]. Its vertex set is $V = \{1, 2, 3, 4, 5\}$ and its (directed) Laplacian is given by

$$\mathcal{L} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ x-1 & 0 & 1 & -x & 0 \\ 0 & 0 & -1/2 & 1 & -1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $x \in [0, 1]$. The reaches are given by $R_1 = \{1, 2, 3, 4\}$ and $R_2 = \{3, 4, 5\}$ with exclusive parts $H_1 = \{1, 2\}$ and $H_2 = \{5\}$. Its root sets are $P_1 = \{1\}$ and $P_2 = \{5\}$. Since this graph has two reaches, the eigenspace corresponding to the zero eigenvalue is two-dimensional. It is easy to verify that it is spanned by

$$\gamma_1 = (2-x)^{-1}\begin{pmatrix} 2-x \\ 2-x \\ 2(1-x) \\ 1-x \\ 0 \end{pmatrix} \quad \text{and} \quad \gamma_2 = (2-x)^{-1}\begin{pmatrix} 0 \\ 0 \\ x \\ 1 \\ 2-x \end{pmatrix}$$

To find the invariant measures we look for the kernel of the transpose of the Laplacian. This time the kernel is spanned by the measures

$$\bar{\gamma}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \bar{\gamma}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Since each of the two reaches has a unique root, the invariant measures are uniquely determined. As an example, a walker starting at vertex 3 has probability $\frac{2(1-x)}{2-x}$ to end up at vertex 1 and probability $\frac{x}{2-x}$ to end up at vertex 3. The associated invariant measure is $\frac{2(1-x)}{2-x}\bar{\gamma}_1 + \frac{x}{2-x}\bar{\gamma}_2$.

References

[1] J. S. Caughman, J. J. P. Veerman, *Kernels of Directed Graph Laplacians*, Electronic Journal of Combinatorics 13, No 1, R39, 2006.

[2] J. J. P. Veerman, G. Lafferriere, John S. Caughman, A. Williams, *Flocks and Formations*, J. Stat.Phys. 121, Vol 5-6, 901-936, 2005.