- All matrices with nonpositive off-diagonal entries whose principal minors are nonnegative are M-matrices. All matrices with nonpositive off-diagonal entries whose principal minors are positive are nonsingular M-matrices.
- If $\mathbf{A} = \mathbf{M} - \mathbf{N}$ is a splitting of a nonsingular M-matrix for which $\mathbf{M}^{-1} \geq \mathbf{0}$, then the linear stationary iteration (15.1.21) converges for all initial vectors $\mathbf{x}(0)$ and for all right-hand sides $\mathbf{b}$. In particular, Jacobi's method converges.

15.2 PERRON–FROBENIUS THEORY

At a mathematics conference held a few years ago our friend Hans Schneider gave a memorable presentation titled "Why I Love Perron–Frobenius" in which he made the case that the Perron–Frobenius theory of nonnegative matrices is not only among the most elegant theories in mathematics, but it is also among the most useful. One might sum up Hans's point by saying that Perron–Frobenius is a testament to the fact that beautiful mathematics eventually tends to be useful, and useful mathematics eventually tends to be beautiful. The applications involving PageRank, HITS, and other ranking schemes [103] help to underscore this principle.

A matrix $\mathbf{A}$ is said to be *nonnegative* when each entry is a nonnegative number (denote this by writing $\mathbf{A} \geq \mathbf{0}$). Similarly, $\mathbf{A}$ is a *positive matrix* when each $a_{ij} > 0$ (write $\mathbf{A} > \mathbf{0}$). For example, the hyperlink matrix $\mathbf{H}$ and the stochastic matrix $\mathbf{S}$ (from Chapter 4) that are at the foundation of PageRank are nonnegative matrices, and the Google matrix $\mathbf{G}$ is a positive matrix. Consequently, properties of positive and nonnegative matrices govern the behavior of PageRank, and the Perron–Frobenius theory reveals these properties by describing the nature of the dominant eigenvalues and eigenvectors of positive and nonnegative matrices.

Perron

So much of the mathematics of PageRank, HITS, and associated ideas involves nonnegative matrices and graphs. This section provides you with the needed ammunition to handle these concepts. Perron's 1907 theorem provides the insight for understanding the eigenstructure of positive matrices. Perron's theorem for positive matrices is stated below, and the proof is in [127].

Perron's Theorem for Positive Matrices

If $\mathbf{A}_{n\times n} > \mathbf{0}$ with $r = \rho(\mathbf{A})$, then the following statements are true.

1. $r > 0$.
2. $r \in \sigma(\mathbf{A})$ (r is called the *Perron root*).
3. $alg\ mult_{\mathbf{A}}(r) = 1$ (the Perron root is simple).
4. There exists an eigenvector $\mathbf{x} > \mathbf{0}$ such that $\mathbf{Ax} = r\mathbf{x}$.
5. The *Perron vector* is the unique vector defined by

$$\mathbf{Ap} = r\mathbf{p}, \qquad \mathbf{p} > \mathbf{0}, \|\mathbf{p}\|_1 = 1,$$

   and, except for positive multiples of $\mathbf{p}$, there are no other nonnegative eigenvectors for $\mathbf{A}$, regardless of the eigenvalue.

6. r is the only eigenvalue on the spectral circle of $\mathbf{A}$.
7. $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$, (the Collatz–Wielandt formula),

$$\text{where } f(\mathbf{x}) = \min_{\substack{1\le i\le n \\ x_i\neq 0}} \frac{[\mathbf{Ax}]_i}{x_i} \text{ and } \mathcal{N} = \{\mathbf{x} \mid \mathbf{x} \geq \mathbf{0} \text{ with } \mathbf{x} \neq \mathbf{0}\}.$$

Extensions to Nonnegative Matrices

Perron's theorem for positive matrices is a powerful result, so it's only natural to ask what happens when zero entries creep into the picture. Not all is lost if we are willing to be flexible. The next theorem (the proof of which is in [127]) says that a portion of Perron's theorem for positive matrices can be extended to nonnegative matrices by sacrificing the existence of a positive eigenvector for a nonnegative one.

Perron's Theorem for Nonnegative Matrices

For $\mathbf{A}_{n\times n} \geq \mathbf{0}$ with $r = \rho(\mathbf{A})$, the following statements are true.

- $r \in \sigma(\mathbf{A})$, (but $r = 0$ is possible).
- There exists an eigenvector $\mathbf{x} \geq \mathbf{0}$ such that $\mathbf{Ax} = r\mathbf{x}$.
- The Collatz–Wielandt formula remains valid.

Frobenius

This is as far as Perron's theorem can be generalized to nonnegative matrices without additional hypothesis. For example, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ shows that properties 1, 3, and 4 in Perron's theorem for positive matrices do not hold for general nonnegative matrices, and

$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ shows that property 6 is also lost. Rather than accepting that the major issues concerning spectral properties of nonnegative matrices had been settled, F. G. Frobenius had the insight in 1912 to look below the surface and see that the problem doesn't stem just from the existence of zero entries, but rather from the positions of the zero entries. For example, properties 3 and 4 in Perron's theorem do not hold for

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \text{ but they are valid for } \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Frobenius's genius was to see that the difference between $\mathbf{A}$ and $\mathbf{B}$ is in terms of matrix reducibility (or irreducibility) and to relate these ideas to spectral properties of nonnegative matrices. The next section introduces these ideas.

Graph and Irreducible Matrices

A *graph* is a set of nodes $\{N_1, N_2, \ldots, N_n\}$ and a set of edges $\{E_1, E_2, \ldots, E_k\}$ between the nodes. A *connected graph* is one in which there is a sequence of edges linking any pair of nodes. For example, the graph shown on the right-hand side of Figure 15.1 is undirected and connected.

A *directed graph* is a graph containing directed edges. A directed graph is said to be *strongly connected* if for each pair of nodes (N_i, N_k) there is a sequence of directed edges leading from N_i to N_k. The graph on the left-hand side of Figure 15.1 is directed but *not* strongly connected (e.g., you can't get from N_3 to N_1).

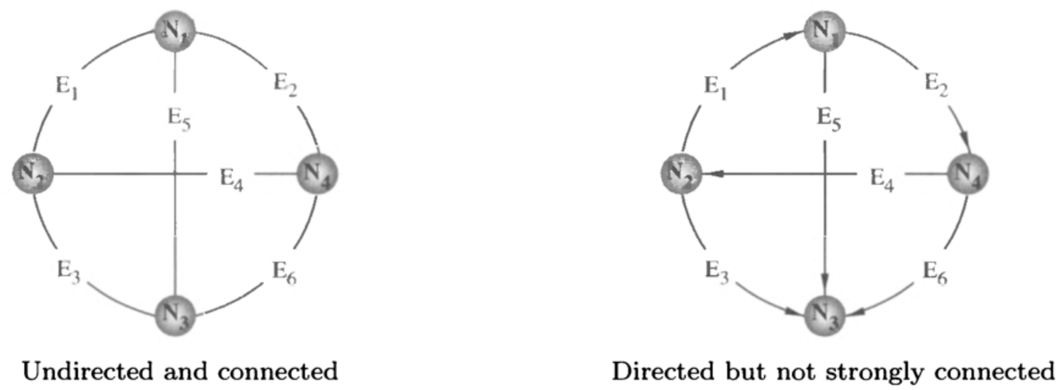


Figure 15.1

Each graph defines two useful matrices—an *adjacency matrix* and an *incidence matrix*. For a graph $\mathcal{G}$ containing nodes $\{N_1, N_2, \ldots, N_n\}$, the *adjacency matrix* $\mathbf{L}_{n\times n}$ is the $(0,1)$-matrix having

$$l_{ij} = \begin{cases} 1 & \text{if there is an edge from } N_i \text{ to } N_j, \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathcal{G}$ is undirected, then its adjacency matrix $\mathbf{L}$ is symmetric (i.e., $\mathbf{L} = \mathbf{L}^T$). For example, the adjacency matrices for the two graphs shown in Figure 15.1 are

$$\mathbf{L}_1 = \begin{array}{c} \\ N_1 \\ N_2 \\ N_3 \\ N_4 \end{array} \begin{array}{c} \begin{array}{cccc} N_1 & N_2 & N_3 & N_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \quad \mathbf{L}_2 = \begin{array}{c} \\ N_1 \\ N_2 \\ N_3 \\ N_4 \end{array} \begin{array}{c} \begin{array}{cccc} N_1 & N_2 & N_3 & N_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

For an undirected graph $\mathcal{G}$ with nodes $\{N_1, N_2, \ldots, N_n\}$ and edges $\{E_1, E_2, \ldots, E_k\}$, the *incidence matrix* $\mathbf{C}_{n\times k}$ is the $(0,1)$-matrix having

$$c_{ij} = \begin{cases} 1 & \text{if node } N_i \text{ touches edge } E_j, \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathcal{G}$ is a directed graph, then its incidence matrix is the $(0,-1,1)$-matrix having

$$c_{ij} = \begin{cases} 1 & \text{if edge } E_j \text{ is directed } \textit{toward} \text{ node } N_i, \\ -1 & \text{if edge } E_j \text{ is directed } \textit{away from} \text{ node } N_i, \\ 0 & \text{if edge } E_j \text{ neither begins nor ends at node } N_i. \end{cases}$$

For example, the incidence matrices for the two graphs shown in Figure 15.1 are

$$\mathbf{C}_1 = \begin{array}{c} \\ N_1 \\ N_2 \\ N_3 \\ N_4 \end{array} \begin{array}{c} \begin{array}{cccccc} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \text{ and } \mathbf{C}_2 = \begin{array}{c} \\ N_1 \\ N_2 \\ N_3 \\ N_4 \end{array} \begin{array}{c} \begin{array}{cccccc} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -1 \end{pmatrix} \end{array}.$$

There is a direct connection between the connectivity of a directed graph and the rank of its incidence matrix.

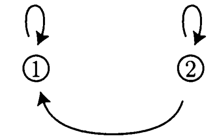
Connectivity and Rank

A directed graph with n nodes and incidence matrix $\mathbf{C}$ is connected if and only if

$$rank\,(\mathbf{C}) = n - 1. \tag{15.2.1}$$

For undirected graphs, arbitrarily assign directions to the edges to make the graph directed and apply (15.2.1) [127, p. 203].

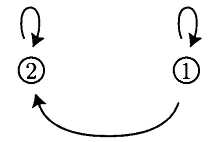
Instead of starting with a graph to build a matrix, we can also do it in reverse—i.e., start with a matrix and build a graph. Given a matrix $\mathbf{A}_{n\times n}$, the graph of $\mathbf{A}$ is defined to be the *directed* graph $\mathcal{G}(\mathbf{A})$ on a set of nodes $\{N_1, N_2, \ldots, N_n\}$ in which there is a directed edge leading *from* N_i *to* N_j if and only if $a_{ij} \neq 0$. For example, if $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$, then the graph $\mathcal{G}(\mathbf{A})$ looks like this:



Any product of the form $\mathbf{P}^T\mathbf{A}\mathbf{P}$ in which $\mathbf{P}$ is a permutation matrix (a matrix obtained from the identity matrix $\mathbf{I}$ by permuting its rows or columns) is called a *symmetric permutation* of $\mathbf{A}$. The effect of a symmetric permutation to a matrix is to interchange rows in the same way as columns are interchanged. The effect of a symmetric permutation on the graph of a matrix is to relabel the nodes. Consequently, the directed graph of a matrix in invariant under a symmetric permutation. In other words, $\mathcal{G}(\mathbf{P}^T\mathbf{AP}) = \mathcal{G}(\mathbf{A})$ whenever $\mathbf{P}$ is a permutation matrix. For example, if $\mathbf{P}$ is the permutation matrix $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and if we again use $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$, then

$$\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}, \tag{15.2.2}$$

and the graph $\mathcal{G}(\mathbf{P}^T\mathbf{AP})$ looks like this:



Matrix $\mathbf{A}_{n\times n}$ is said to be a *reducible matrix* when there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix}, \quad \text{where } \mathbf{X} \text{ and } \mathbf{Z} \text{ are both square.} \tag{15.2.3}$$

For example, the matrix $\mathbf{A}$ in (15.2.2) is clearly reducible. Naturally, an *irreducible matrix* is a matrix that is not reducible.

As the following theorem shows, the concepts of matrix irreducibility (or reducibility) and strong connectivity (or lack thereof) are intimately related.

Irreducibility and Connectivity

A square matrix $\mathbf{A}$ is irreducible if and only if its directed graph is strongly connected. In other words, $\mathbf{A}$ is irreducible if and only if for each pair of indices (i, j) there is a sequence of entries in $\mathbf{A}$ such that $a_{ik_1}a_{k_1k_2}\cdots a_{k_tj} \neq 0$. Equivalently, $\mathbf{A}$ is irreducible if for all permutation matrices $\mathbf{P}$,

$$\mathbf{P}^T\mathbf{AP} \neq \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix}, \quad \text{where } \mathbf{X} \text{ and } \mathbf{Z} \text{ are square.}$$

For example, can you determine if

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 9 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

is reducible or irreducible? It would be a mistake to try to use the definition because deciding on whether or not there exists a permutation matrix $\mathbf{P}$ such that (15.2.3) holds by sorting through all 6×6 permutation matrices is pretty hard. However, the above theorem makes the question easy. Examining $\mathcal{G}(\mathbf{A})$ reveals that it is strongly connected (every node is accessible by some sequence of paths from every other node), so $\mathbf{A}$ must be irreducible.

The Perron–Frobenius Theorem

Frobenius's contribution was to realize that while properties 1, 3, 4, and 6 in Perron's theorem for positive matrices can be lost when zeros creep into the picture (i.e., for nonnegative matrices), the trouble is not simply the existence of zero entries, but rather the problem is the location of the zero entries. In other words, Frobenius realized that the lost properties 1, 3, and 4 are in fact *not lost* when the zeros are in just the right locations—namely the locations that ensure that the matrix is irreducible. Unfortunately irreducibility alone still does not save property 6—it remains lost (more about this issue later).

Below is the formal statement of the Perron–Frobenius theorem—the details concerning the proof can be found in [127].

Perron–Frobenius Theorem

If $\mathbf{A}_{n\times n} \geq \mathbf{0}$ is irreducible, then each of the following is true.

1. $r = \rho(\mathbf{A}) > 0$.
2. $r \in \sigma(\mathbf{A})$ (r is the *Perron root*).
3. $alg\ mult_{\mathbf{A}}(r) = 1$. (the Perron root is simple).
4. There exists an eigenvector $\mathbf{x} > \mathbf{0}$ such that $\mathbf{Ax} = r\mathbf{x}$.
5. The *Perron vector* is the unique vector defined by

$$\mathbf{Ap} = r\mathbf{p}, \qquad \mathbf{p} > \mathbf{0}, \|\mathbf{p}\|_1 = 1,$$

   and, except for positive multiples of $\mathbf{p}$, there are no other nonnegative eigenvectors for $\mathbf{A}$, regardless of the eigenvalue.
6. r need not be the only eigenvalue on the spectral circle of $\mathbf{A}$.
7. $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$, (the Collatz–Wielandt formula),

$$\text{where } f(\mathbf{x}) = \min_{\substack{1\leq i\leq n\\ x_i\neq 0}} \frac{[\mathbf{Ax}]_i}{x_i} \text{ and } \mathcal{N} = \{\mathbf{x}\,|\,\mathbf{x}\geq\mathbf{0} \text{ with } \mathbf{x}\neq\mathbf{0}\}.$$

Primitive Matrices

The only property in Perron's theorem for positive matrices on page 168 that irreducibility is not able to salvage is the sixth property, which states that there is only one eigenvalue on the spectral circle. Indeed, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is nonnegative and irreducible, but the eigenvalues ± 1 are both on the unit circle. The property of having (or not having) only one eigenvalue on the spectral circle divides the set of nonnegative irreducible matrices into two important classes.

Primitive Matrices

- A matrix $\mathbf{A}$ is defined to be a *primitive matrix* when $\mathbf{A}$ is a nonnegative irreducible matrix that has only one eigenvalue, $r = \rho(\mathbf{A})$, on its spectral circle.
- A nonnegative irreducible matrix having $h > 1$ eigenvalues on its spectral circle is said to be *imprimitive,* and h is called the *index of imprimitivity.*
- If $\mathbf{A}$ is imprimitive, then the h eigenvalues on the spectral circle are

  $$\{r,\, r\omega,\, r\omega^2,\, \ldots,\, r\omega^{h-1}\}, \quad \text{where} \quad \omega = \mathrm{e}^{2\pi \mathrm{i}/h}.$$

  In other words, they are the h^{th} roots of $r = \rho(\mathbf{A})$, and they are uniformly spaced around the circle. Furthermore each eigenvalue $r\omega^k$ on the spectral circle is simple.

So what's the big deal about having only one eigenvalue on the spectral circle? Well, primitivity is important because it's precisely what determines whether or not the powers of a normalized nonnegative irreducible matrix will have a limiting value, and this is the fundamental issue concerning the existence of the PageRank vector. The precise wording of the theorem is as follows.

Limits and Primitivity

A nonnegative irreducible matrix $\mathbf{A}$ with $r = \rho(\mathbf{A})$ is primitive if and only if $\lim_{k\to\infty}(\mathbf{A}/r)^k$ exists, in which case

$$\lim_{k\to\infty}\left(\frac{\mathbf{A}}{r}\right)^k = \frac{\mathbf{p}\mathbf{q}^T}{\mathbf{q}^T\mathbf{p}} > \mathbf{0}, \tag{15.2.4}$$

where $\mathbf{p}$ and $\mathbf{q}^T$ are the respective right-hand and left-hand Perron vectors for $\mathbf{A}$.

If $\mathbf{A}_{n\times n} \geq \mathbf{0}$ is irreducible but imprimitive so that there are $h > 1$ eigenvalues on the spectral circle, then it can be demonstrated [127] that each of these eigenvalues is simple and that they are distributed uniformly on the spectral circle in the sense that they are the h^{th} roots of $r = \rho(\mathbf{A})$—i.e., the eigenvalues on the spectral circle are given by

$$\{r,\, r\omega,\, r\omega^2,\, \ldots,\, r\omega^{h-1}\}, \quad \text{where} \quad \omega = \mathrm{e}^{2\pi \mathrm{i}/h}.$$

Given a nonnegative matrix, do we really have to compute the eigenvalues and count how many fall on the spectral circle to check for primitivity? No! There are simpler tests.

Tests for Primitivity

For a square nonnegative matrix $\mathbf{A}$, each of the following is true.

- $\mathbf{A}$ is primitive if $\mathbf{A}$ is irreducible and has at least one positive diagonal element.
- $\mathbf{A}$ is primitive if and only if $\mathbf{A}^m > \mathbf{0}$ for some $m > 0$.

The first test above only provides a sufficient condition for primitivity, while the second condition is both necessary and sufficient—the first test is cheaper but not conclusive, while the second is more expensive, but absolutely conclusive. For example, to determine whether or not the irreducible matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 3 & 4 & 0 \end{pmatrix}$ is primitive, the first test doesn't apply because the diagonal of $\mathbf{A}$ is entirely zeros, so we are forced to apply the second test by computing powers of $\mathbf{A}$. But the job is simplified by noticing that if $\mathbf{B}$ is the Boolean matrix defined by

$$b_{ij} = \begin{cases} 1 & \text{if } a_{ij} > 0, \\ 0 & \text{if } a_{ij} = 0, \end{cases}$$

then $[\mathbf{B}^k]_{ij} > 0$ if and only if $[\mathbf{A}^k]_{ij} > 0$ for every $k > 0$. Therefore, we only need to compute powers of $\mathbf{B}$ (it can be shown that no more than $n^2 - 2n + 2$ powers are required), and these powers require only Boolean operations *AND* and *OR*. The matrix $\mathbf{A}$ in this example is primitive because the powers of $\mathbf{B}$ are

$$\mathbf{B} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \ \mathbf{B}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \ \mathbf{B}^3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \ \mathbf{B}^4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \ \mathbf{B}^5 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

While we might prefer our matrices to be primitive, Mother Nature doesn't always cooperate. Mathematical models of physical phenomena that involve oscillations generally produce imprimitive matrices, where the number of eigenvalues on the spectral circle (the index of imprimitivity) corresponds to the period of oscillation. Consequently, it's worthwhile to have a grasp on the index of imprimitivity. While the powers of an irreducible matrix $\mathbf{A} \geq \mathbf{0}$ can tell us if $\mathbf{A}$ has more than one eigenvalue on its spectral circle, the powers of $\mathbf{A}$ provide no clue to the number of such eigenvalues. The issue is more complicated—the following theorem is the primary theoretical aid in determining the index of imprimitivity short of actually computing all eigenvalues.

Index of Imprimitivity

If $c(x) = x^n + c_{k_1}x^{n-k_1} + c_{k_2}x^{n-k_2} + \cdots + c_{k_s}x^{n-k_s} = 0$ is the characteristic equation of an imprimitive matrix $\mathbf{A}_{n \times n}$ in which only the terms with nonzero coefficients are listed (i.e., each $c_{k_j} \neq 0$, and $n > (n-k_1) > \cdots > (n-k_s)$), then the index of imprimitivity h is the greatest common divisor of $\{k_1, k_2, \ldots, k_s\}$.

Finally, it is often useful to decompose an imprimitive matrix, and the *Frobenius form* is the standard way of doing so.

Frobenius Form

For each imprimitive matrix $\mathbf{A}$ with index of imprimitivity $h > 1$, there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{23} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_{h-1,h} \\ \mathbf{A}_{h1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}, \qquad (15.2.5)$$

where the zero blocks on the main diagonal are square.

15.3 MARKOV CHAINS

The mathematical component of Google's PageRank vector is the stationary distribution of a discrete-time, finite-state Markov chain. So, to understand and analyze the mathematics of PageRank, it's necessary to have an appreciation of Markov chain concepts, and that's the purpose of this section. Let's begin with some definitions.

- A *stochastic matrix* is a nonnegative matrix $\mathbf{P}_{n\times n}$ in which each row sum is equal to 1. Some authors say "row-stochastic" to distinguish this from the case when each column sum is 1.
- A *stochastic process* is a set of random variables $\{X_t\}_{t=0}^{\infty}$ having a common range $\{S_1, S_2, \ldots, S_n\}$, which is called the *state space* for the process. Parameter t is generally thought of as time, and X_t represents the state of the process at time t. For example, consider the process of surfing the Web by successively clicking on links to move from one Web page to another. The state space is the set of all Web pages, and the random variable X_t is the Web page being viewed at time t.
  - To emphasize that time is considered discretely rather than continuously the phrase "*discrete-time* process" is often used, and the phrase "*finite-state* process" can be used to emphasize that the state space is finite rather than infinite. Our discussion is limited to discrete-time finite-state processes.
- A *Markov chain* is a stochastic process that satisfies the *Markov property*

  $$P(X_{t+1} = S_j \mid X_t{=}S_{i_t}, X_{t-1}{=}S_{i_{t-1}}, \ldots, X_0{=}S_{i_0}) = P(X_{t+1} = S_j \mid X_t = S_{i_t})$$

  for each $t = 0, 1, 2, \ldots$. The notation $P(E \mid F)$ denotes the conditional probability that event E occurs given event F occurs—a review some elementary probability is in order if this is not already a familiar concept.
  - The Markov property asserts that the process is memoryless in the sense that the state of the chain at the next time period depends only on the current state and not on the past history of the chain. For example, the process of surfing the Web is a Markov chain provided that the next page that the Web surfer visits doesn't depend on the pages that were visited in the past—the choice depends