0. Preliminaries. This small set of notes is intended as a complement to Atkinson’s Ch. 7 and part of Ch. 8, so the treatment will be somewhat sketchy, referring to the textbook quite a bit. Much of his §§7.1–7.3 is discussed at considerable length in the separate notes on linear algebra, the first link on the list on the home web page of this course.

1. A Few More Things about Vector Norms. There are a few numerical-analysis convergence theorems in which it becomes important to know the “Equivalence of Norms” theorem, Atkinson’s Theorem 7.7, p. 483. The theorem is basically a consequence of the fact that continuous real-valued functions defined on closed-and-bounded subsets of \( \mathbb{R}^n \) or \( \mathbb{C}^n \) (which looks, for metric purposes, like \( \mathbb{R}^{2n} \)) attain maximum and minimum values. Atkinson gives the standard proof. Note that it also applies to norms on the spaces of \( m \times n \) matrices, since these can be identified with \( \mathbb{R}^{mn} \) or \( \mathbb{C}^{mn} \) respectively: no matter how the \( m \)- and \( n \)-dimensional vector spaces have been normed, convergence of \( m \times n \) matrices is the same in all associated operator norms.

The natural, induced or operator norms on the spaces of \( n \times n \) matrices (also called matrix norms by some authors, e.g., Kincaid & Cheney) are discussed adequately by Atkinson on pp. 484–486, though he does not mention two important and frequently used facts about operator norms. The first is that the operator norm of an identity matrix is 1, with respect to any vector norm on \( \mathbb{R}^n \) or \( \mathbb{C}^n \), because

\[
\| I \| = \max \{ \| I x \| : \| x \| \leq 1 \} = 1
\]

is trivially true.\(^{(1)}\) The second is that the natural norm is the “amplification factor” for \( A \), in that \( \| A \| \) is the smallest number \( M \geq 0 \) for which the inequality \( \| A x \| \leq M \cdot \| x \| \) will hold for every \( x \) in \( \mathbb{R}^n \) or \( \mathbb{C}^n \).

Reason: on one hand, any such number \( M \) is larger than \( \| A \| \), because

\[
\| A x \| \leq M \cdot \| x \| \quad \Rightarrow \quad \max \{ \| A x \| : \| x \| \leq 1 \} \leq M \cdot 1 = M ;
\]

on the other hand, \( \| A \| \) is such a number, because for any \( 0 \neq x \) in \( \mathbb{R}^n \) or \( \mathbb{C}^n \) we may set \( z = (1/\| x \|) x \); then \( \| z \| = 1 \) and

\[
\| (1/\| x \|) A x \| = \| A z \| \leq \| A \| \quad \text{by definition}
\]

\[
\| A x \| \leq \| A \| \cdot \| x \| \quad \text{upon clearing fractions.}
\]

Atkinson’s unworked example (7.3.17) on p. 487, giving a concrete computable form for \( \| A \|_\infty \), is one of the happiest facts in numerical analysis. The concrete expression (7.3.19) for the natural matrix norm associated with the \( \ell^2 \) norm of vectors is also useful for thinking, but its numerical value for a particular matrix is much harder to compute. (For further information on these topics see also Kincaid & Cheney p. 201 ff., and their §5.4 or a discussion of singular value decompositions: Atkinson gives one as Theorem 7.5, pp. 478–479, and there is also a discussion in the longer linear-algebra notes.)

2. Schur’s Theorem, or the “Schur normal form” of matrices. I have some quarrel with the use of “normal” or “canonical” in this context, because the upper-triangular matrix unitarily (or sometimes orthogonally) equivalent to the given matrix is very far from uniquely determined. But Schur’s theorem is an important lemma in matrix numerical analysis: it can be used to replace the Jordan canonical form in many arguments. Its principal flaw for concrete numerical work, like that of the Jordan canonical form, is that the eigenvalues it requires cannot be determined in a numerically robust way.

Atkinson’s proof of this theorem is the standard one, an “induction by descent.” It is possible to phrase the descent as a constructive induction, up to determining the eigenvalues of the given matrix. For example (using Atkinson’s notation on pp. 474–476), after finding the eigenvalue \( \lambda_1 \), the orthonormal basis \( \{ u^{(1)}, \ldots, u^n \} \) whose first element is an eigenvector belonging to \( \lambda_1 \), and the unitary matrix \( U_1 \) of

\(^{(1)}\) Note that the “Frobenius norm” (which many people call the “Hilbert-Schmidt norm”; its infinite-dimensional analogue is important in the theory of integral equations) does not have this property and thus is not an operator norm.
which the $u^{(j)}$’s are the columns (this is Atkinson’s $P_1$), one forms $B_1 = U_1^* A U_1$ which has the form

$$B_1 = \begin{bmatrix}
\lambda_1 & * & \cdots & *

0 & & &

\vdots & & \ddots & \ddots

0 & \cdots & A_2 & \ddots

\end{bmatrix}.$$  

Expanding $\det(\lambda I_n - B_1)$ by its first column, one sees that the characteristic polynomial of $B_1$—call it $\chi_1(\lambda)$—is just that of $U_1^* A U_1$ and therefore the same as that of $A$, divided by $(\lambda - \lambda_1)$: in symbols, $\chi_1(\lambda) = \frac{\chi_A(\lambda)}{\lambda - \lambda_1}$. Thus the eigenvalues of $B_1$ are the same as those of $A$, according to multiplicity, except that the multiplicity of $\lambda_1$ has been decreased by one (and thus $\lambda_1$ has been deleted, if it was a simple root of $\chi_A(\lambda)$). We now play almost the same game with the $(n-1) \times (n-1)$ matrix $A_2$, finding an $(n-1) \times (n-1)$ unitary matrix $P_2$ for which the equation in $(n-1) \times (n-1)$ matrices:

$$P_2^* A_2 P_2 = B_2 = \begin{bmatrix}
\lambda_2 & * & \cdots & *

0 & & &

\vdots & & \ddots & \ddots

0 & \cdots & A_3 & \ddots

\end{bmatrix},$$

holds. Let $U_2$ be the partitioned $n \times n$ matrix $U_2 = \begin{bmatrix} I_1 & 0 & \cdots & 0 \\
0 & & &

\vdots & & & \\
0 & & & \\
\end{bmatrix}$ where $I_1$ is the $1 \times 1$ identity matrix. Then it is easy to see that

$$(U_1 U_2)^* A (U_1 U_2) = \begin{bmatrix}
\lambda_1 & * & \cdots & *

0 & \lambda_2 & * & \cdots & *

\vdots & \vdots & \ddots & \ddots & \ddots

0 & 0 & \cdots & A_3

\end{bmatrix}.$$  

Continuing in this fashion, we build up a product of $n \times n$ unitary matrices, the last having the form

$$\begin{bmatrix} I_{n-2} & 0 \\
0 & P_{n-1} \end{bmatrix}.$$  

The product $(U_1 U_2 \cdots U_{n-1})^* A (U_1 \cdots U_{n-1})$ will then have the upper-triangular form required by the theorem. Note that if all the eigenvalues of $A$ are real, then the unitary matrices $U_j$ can be taken to be (real and) orthogonal rather than (complex) unitary. As a very simple example, the matrix

$$\begin{bmatrix} 2 & 1 \\
-1 & 0 \end{bmatrix}$$

has characteristic polynomial $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$ but only the single eigenvector $[1, -1]^T$, which we can normalize to $[\sqrt{2}/2, -\sqrt{2}/2]^T$. We pick this as the first column of $U_1$, and the only possible real choice of the second column (up to sign, which we can choose to give positive orientation if we wish) is $[\sqrt{2}/2, \sqrt{2}/2]^T$. It is routine to verify that

$$\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\
\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\
-1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\
-\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\
0 & 1 \end{bmatrix}.$$  

As a corollary of Schur’s theorem, we see that a conjugate-symmetric matrix $A$ can be diagonalized by a unitary equivalence, replacing $A$ by $U^* A U$: $U$ can be found for which $U^* A U$ is upper-triangular, but because $U^* A U$ is still conjugate-symmetric, it must be diagonal. In particular, if $A$ is a real symmetric matrix, then its eigenvalues are real—but then the unitary matrix rendering $A$ upper-triangular can be taken to be a (real) orthogonal matrix $Q$, and we see that $Q^T A Q$ is (real and) diagonal. Since the columns of $Q$ are eigenvectors of $A$, this analysis also gives us an orthonormal basis of $\mathbb{R}^n$ whose elements are eigenvectors of $A$. This fact is sometimes called the **Principal Axes Theorem**.

Schur’s Theorem also furnishes us with a way of seeing that the behavior of $\|A^k\|$ for large $k$ is governed by the **spectral radius** $r_s(A) = \max \{ |\lambda| : \lambda \in \sigma(A) \}$ of $A$, where as usual $\sigma(A)$ is the **spectrum**= set of
The eigenvalues of $A$. To see this, let $0 < \epsilon < 1$ be given and let $E$ be the $n \times n$ diagonal matrix whose diagonal elements are $[\epsilon^1, \epsilon^2, \ldots, \epsilon^n]$ (and whose inverse is therefore the $n \times n$ diagonal matrix whose diagonal elements are $[\epsilon^{-1}, \epsilon^{-2}, \ldots, \epsilon^{-n}]$). If $C$ is an $n \times n$ matrix, then left-multiplying it by a diagonal matrix multiplies the $i$-th row of $C$ by the $i$-th diagonal entry of the multiplier, while right-multiplying it by a diagonal matrix multiplies the $j$-th column of $C$ by the $j$-th diagonal entry of the multiplier. It follows that if $C = (c_{ij})$, then the $ij$-th entry in $E^{-1}CE$ is $\epsilon^{-1}c_{ij}\epsilon^j = c_{ij}\epsilon^{i-j}$. If now $C$ is an upper-triangular matrix, then its diagonal elements will be unchanged, its below-diagonal elements (those for which $i > j$) will remain zero, but its above-diagonal elements (for which $j > i$) will be multiplied by a positive integer power of $\epsilon$. By taking $\epsilon > 0$ sufficiently small, therefore, we can make $E^{-1}CE$ as close as we wish in matrix norm—say the $\ell^2$ matrix norm for definiteness—to the diagonal matrix whose diagonal elements are $[c_{11}, \ldots, c_{nn}]$. Let a matrix $A$ be given. Using Schur’s theorem, we can find a unitary matrix $U$ for which $U^*AU$ is upper-triangular, and then by choosing a suitable small $\epsilon > 0$ and forming $E^{-1}U^*AU = (UE)^{-1}A(UE)$, we can produce a matrix which is as close as we please (in the $\ell^2$ matrix norm) to the diagonal matrix whose diagonal elements are the eigenvalues $[\lambda_1, \ldots, \lambda_n]$ of $A$, enumerated according to multiplicity. We can now put a new norm on $\mathbb{C}^n$ (and thus also on $\mathbb{R}^n$) as follows: set $\|x\|_0 = \|(UE)^{-1}x\|_2$. In words, measure $x$ by the $\ell^2$-norm length of its image under $(UE)^{-1}$. It is obvious that this is a norm, and if $y$ is chosen with $\|y\|_2 = 1$ such that $\|(UE)^{-1}A(UE)y\|_2 = \|(UE)^{-1}A(UE)\|_2$ (that is, $y$ is a unit vector whose image under $(UE)^{-1}A(UE)$ is of maximal $\ell^2$-length), then setting $x = UEy$ we have $\|x\|_0 = 1$ and $\|Ax\|_0 = \|A(UE)y\|_2 = \|(UE)^{-1}A(UE)\|_2$. Thus $\|A\|_0 = \|(UE)^{-1}(AUE)\|_2$, and since $(UE)^{-1}A(UE)$ can be as close as we please to the diagonal matrix with diagonal elements the eigenvalues $[\lambda_1, \ldots, \lambda_n]$ of $A$, its $\ell^2$ matrix norm can be as close to $r_\sigma(A)$ as we please, so an appropriate choice of $E$ will make the matrix norm $\|A\|_0$ as close to $r_\sigma(A)$ as we please (though in general equality cannot be attained). This gives us the following

**Theorem:** Let $A$ be an $n \times n$ matrix. In order that $\|A^k\| \to 0$ in (every) matrix norm as $k \to \infty$, it is necessary and sufficient that $r_\sigma(A) < 1$.

**Proof.** If $r_\sigma(A) < 1$, then as we have just seen there is a norm on $\mathbb{C}^n$ for which $\|A\|_0$ is as close to $r_\sigma(A)$ as we wish, and therefore we may choose a norm for which $\|A\|_0 < 1$. Since $\|A^k\|_0 \leq \|A\|_0^k$ and the r. h. s. tends to zero (geometrically fast), $A^k \to 0$. On the other hand, if $r_\sigma(A) \geq 1$ then there is a (nonzero) eigenvector $x \in \mathbb{C}^n$ of $A$ belonging to an eigenvalue $\lambda \in \sigma(A)$ with $|\lambda| \geq 1$. We then have $\|A^kx\| = |\lambda|^k\|x\|$ in any norm you please on $\mathbb{C}^n$, and obviously the r. h. s. does not tend to zero with increasing $k \in \mathbb{N}$.

Another consequence of this “renorming theorem” is the following one: while this result is rarely of use in computing spectral radii of particular concrete matrices, but it is amusing and has theoretical implications.

**Theorem:** Let $\|\cdot\|$ be any associated matrix norm. Then for every $n \times n$ matrix $A$,

$$r_\sigma(A) = \lim_{m \to \infty} \|A^m\|^{1/m}.$$

**Proof.** As we observed in the previous theorem, if $\lambda$ is an eigenvalue of $A$ whose absolute value is maximal, so $|\lambda| = r_\sigma(A)$, then

$$r_\sigma(A)^m = |\lambda|^m \leq \|A^m\|$$

and taking $m$-th roots gives

$$r_\sigma(A) \leq \|A^m\|^{1/m} \text{ for all natural numbers } m.$$

On the other hand, for any given $\epsilon > 0$ we can find a new norm $\|\cdot\|_{\epsilon}$ on vectors which will have the property that

$$\|A\|_{\epsilon} \leq r_\sigma(A) + \frac{\epsilon}{2},$$

and thus

$$\|A^m\|_{\epsilon} \leq \|A\|_{\epsilon}^m \leq \left(r_\sigma(A) + \frac{\epsilon}{2}\right)^m.$$

(2) Recall that for any $1 \leq p \leq \infty$, the $\ell^p$ matrix norm of a diagonal matrix is the largest absolute value of an element on the diagonal.
Atkinson’s Theorem 7.7, p. 483—applied to the associated matrix norms of $\| \cdot \|$ and $\| \cdot \|_c$—tells us that there exists a constant $c > 0$ for which

$$c \cdot \| B \| \leq \| B \|_c$$

for every $n \times n$ matrix $B$. With $B = A^m$ this gives

$$c \| A^m \| \leq \| A^m \|_c \leq \left[ r_\sigma(A) + \frac{\epsilon}{2} \right]^m$$

from which follows

$$\| A^m \|^{1/m} \leq \frac{1}{\sqrt{c}} \left[ r_\sigma(A) + \frac{\epsilon}{2} \right]$$

with the constant $c$ independent of $m$. The limit of the r. h. s. as $m \to \infty$ is obviously $r_\sigma(A) + \epsilon/2$, so for sufficiently large $m$ we must have

$$r_\sigma(A) \leq \| A^m \|^{1/m} \leq r_\sigma(A) + \epsilon$$

and that proves the limit assertion of the theorem.

3. Matrices with Nonnegative Entries. This is material which should be more widely known to applied mathematicians, since many linear-algebraic problems that arise in applications come equipped with natural constraints on the signs of their matrices. The word “positive” and symbol “$\geq 0$” are frequently used to denote $(n \times n)$ symmetric matrices $A$ for which $A x \cdot x \geq 0$ holds for all $x \in \mathbb{R}^n$ (or $x \in \mathbb{C}^n$, if $A$ is conjugate-symmetric), although this condition only forces positivity on the diagonal elements of $A$ (for example, $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is positive in this sense). To try to avoid confusion, therefore, we shall write $A \succeq 0$ for matrices all of whose entries are nonnegative, and call them matrices with nonnegative entries. Similarly, we shall write $x \succeq 0$ and say that $x \in \mathbb{R}^n$ is a nonnegative vector if all its components are nonnegative.

The most important fact about these matrices is a theorem variously ascribed to Frobenius, Perron, and others:

**Theorem:** Let $A \succeq 0$ be an $n \times n$ matrix with nonnegative entries. Then

(a) The spectral radius $r_\sigma(A)$ belongs to the spectrum of $A$ (i.e., it is an eigenvalue), and it is the eigenvalue of $A$ of largest absolute value;

(b) There is\(^{(3)}\) a nonnegative eigenvector $0 \preceq x \in \mathbb{R}^n$ belonging to the eigenvalue $r_\sigma(A)$;

(c) Each eigenvalue of $A$ is a pole of the resolvent. The order of the pole at $r_\sigma(A)$ is largest among the orders of the poles of all eigenvalues $\lambda \in \sigma(A)$ for which $|\lambda| = r_\sigma(A)$;

(d) The resolvent of $A$, which is the matrix-valued function $\lambda \mapsto (\lambda I - A)^{-1}$ defined on $\mathbb{C} \setminus \sigma(A)$, takes values that are matrices with nonnegative entries for $\lambda \in (r_\sigma(A), \infty) \subseteq \mathbb{R}$, and

(e) it does not take such values for $\lambda \in (0, r_\sigma(A)) \subseteq \mathbb{R}$.

The proof of this theorem requires some preparation. The principal lemma is a theorem of A. Pringsheim on analytic functions whose power-series coefficients are nonnegative:

\(^{(3)}\) Under additional assumptions about $A$ one can show that this eigenvector is uniquely determined (up to positive scalar multiples, of course), that each of its entries is strictly positive, and that the order of the pole of the resolvent at $r_\sigma(A)$ is 1. To prove these things would take us too far afield from numerical analysis, although all of them are useful in various parts of applied mathematics. For example, the Perron-Frobenius theorem in the form seen here implies the existence of stationary probability distributions for (finite) Markov chains, and the more refined version implies their uniqueness.
[Pringsheim’s] Theorem: Let \( \sum_{n=0}^{\infty} a_n z^n \) be a power series with radius of convergence \( \rho > 0 \), all of whose coefficients \( a_n \geq 0 \). Then the function \( f(z) = \sum_{n=0}^{\infty} a_n z^n \) defined by the power series has a singularity at the point \( \rho \in \mathbb{R}^+ \subseteq \mathbb{C} \); that is, there is no analytic function defined in any set that is the union of two open discs \( \{ z : |z| < r \} \cup \{ z : |z - \rho| < \delta \} \) centered at 0 and at \( r \) respectively, which extends \( f(z) \) to that union.

Proof of Pringsheim’s Theorem. Suppose there were such an extension of \( f(z) \). Then it is geometrically clear(4) that there exist a real number \( z_0 \) with \( 0 < z_0 < \rho \) and a radius \( \epsilon > 0 \) such that the closed disc \( \{ z : |z - z_0| \leq \epsilon \} \) contains \( \rho \) and is contained in the union \( \{ z : |z| < r \} \cup \{ z : |z - \rho| < \delta \} \). Now because \( |z_0| < \rho \) we can find the power-series expansion of \( f(z) \) with center \( z_0 \) by formally rewriting the series \( \sum_{n=0}^{\infty} a_n z^n \) by substituting \( z = [(z - z_0) + z_0] \)—at least for real \( z_0 < z < \rho \):

\[
\sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} a_n [(z - z_0) + z_0]^n = \sum_{n=0}^{\infty} a_n \left[ \sum_{k=0}^{n} \binom{n}{k} z_0^{n-k} (z - z_0)^k \right]
\]

The interchange of the order of summation is justified by the positivity of the quantities involved. But now, since the power series \( \sum_{k=0}^{\infty} \left[ \sum_{n=k}^{\infty} \binom{n}{k} a_n z_0^{n-k} \right] (z - z_0)^k \) converges to \( f(z) \) at least on the line segment \( z_0 < z < \rho \), it converges—and converges to \( f(z) \)—at least in the disc \( \{ z : |z - z_0| < \rho - z_0 \} \). But since the disc \( \{ z : |z - z_0| \leq \epsilon \} \) is contained in the domain of analyticity of the extension of \( f(z) \), this series must converge (to the extended function) at each point of that disc, and in particular it must converge for \( z = z_0 + \epsilon \). But now if we formally retrace the steps in the long set-off equation above, we get

\[
\sum_{k=0}^{\infty} \left[ \sum_{n=k}^{\infty} \binom{n}{k} a_n z_0^{n-k} \right] (z - z_0)^k \bigg|_{z=z_0+\epsilon} = \sum_{n=0}^{\infty} a_n \left[ \sum_{k=0}^{n} \binom{n}{k} z_0^{n-k} \right] \epsilon^k
\]

\[
= \sum_{n=0}^{\infty} a_n \left[ \sum_{k=0}^{n} \binom{n}{k} z_0^{n-k} \right] = \sum_{n=0}^{\infty} a_n (z_0 + \epsilon)^n
\]

in which, again, the interchange of the order of summation is justified by the positivity of everything in sight. However, we now see that \( \sum_{n=0}^{\infty} a_n (z_0 + \epsilon)^n \) converges, and therefore the original series \( \sum_{n=0}^{\infty} a_n z^n \) converges (absolutely) for all \( z \in \mathbb{C} \) with \( |z| \leq z_0 + \epsilon \), and since \( z_0 + \epsilon > \rho \) the original series converges for \( |z| \) larger than its radius of convergence, a contradiction that proves Pringsheim’s theorem.

On our way to proving the Perron-Frobenius theorem, we now need to look at the resolvent series (or Neumann series) for \((\lambda I - A)^{-1}\), where \( A \) is an arbitrary matrix (not necessarily having nonnegative entries): this is the formal matrix series obtained by the formal geometric-series calculation

\[
(\lambda I - A)^{-1} = \frac{1}{\lambda - A} = \frac{1}{\lambda} \frac{1}{1 - \frac{A}{\lambda}} = \sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} A^k .
\]

This series clearly converges for \( |\lambda| > \|A\| \), and multiplying its sum by \((\lambda I - A)\) shows that in fact it is the inverse of that matrix. Since this is a convergent power series representing an analytic function (look (4) Draw a picture!}
at the entries of \((\lambda I - A)^{-1}\) as given by determinant formulas: each entry is a rational function of \(\lambda\) and thus analytic off the zeros of \(\det(\lambda I - A)\), i.e., analytic in \(\mathbb{C} \setminus \sigma(A)\), it converges and represents \((\lambda I - A)^{-1}\) in the “disc centered at infinity” \(\{\lambda : |\lambda| > r_\sigma(A)\}\). Now in the case in which \(A \geq 0\), it is clear that the coefficients of this series (or of each entry in the matrix sum of the series) are nonnegative numbers. By Pringsheim’s theorem (with \(\frac{1}{\lambda}\) taking the place of \(z\)), we see that at least one of the entries in the sum of the series must have a singularity at the point where the circle \(|\lambda| = r_\sigma(A)\) meets the positive real line, i.e., exactly at \(\lambda = r_\sigma(A)\); and since those entries are rational functions, that singularity must be a zero of the function \(\det(\lambda I - A) = \chi_A(\lambda)\)—in other words, it must be an eigenvalue of \(A\). (Obviously no eigenvalue of \(A\) has \(|\lambda| > r_\sigma(A)\).)

So we have now established assertion (a) of the Perron-Frobenius theorem. Indeed, we have also established assertion (d), since the geometric series formula for \((\lambda I - A)^{-1}\), valid for \(r_\sigma(A) < \lambda \in \mathbb{R}\), shows that \((\lambda I - A)^{-1} \geq 0\) for \(\lambda\) in that range. That each entry of \((\lambda I - A)^{-1}\) has at worst a pole (half of assertion (c)) at each \(\lambda \in \sigma(A)\) is obvious from the determinant formula for the inverse, and indeed it is clear that the order of the pole can be at most the order of the zero of \(\chi_A(\lambda)\) at that point. So if \(m\) is the largest order of the pole of \((\lambda I - A)^{-1}\) at \(r_\sigma(A)\) among all the entries of that matrix, then it is easy to see, entry-by-entry, that the limit \(\lim_{\lambda \to r_\sigma(A)} (\lambda - r_\sigma(A))^m (\lambda I - A)^{-1}\) exists and is not the zero matrix. If we take that limit as \(\lambda \downarrow r_\sigma(A)\) through (positive) real values in the form

\[
L = \lim_{\lambda \downarrow r_\sigma(A)} (\lambda - r_\sigma(A))^m \left[ \sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} A^k \right]
\]

then it is clear that the limit matrix \(0 \preceq L \neq 0\). Because

\[
(\lambda I - A) L = \lim_{\lambda \downarrow r_\sigma(A)} (\lambda - r_\sigma(A))^m (\lambda I - A) \left[ \sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} A^k \right] = \lim_{\lambda \downarrow r_\sigma(A)} (\lambda - r_\sigma(A))^m I = 0 ,
\]

it is clear that any nonzero column of \(L\), or alternatively the (necessarily nonzero) vector \(L \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\), will be a nonnegative eigenvector of \(A\) belonging to the eigenvalue \(r_\sigma(A)\), and that gives us assertion (b).

We still have to do a small amount of cleaning up. If \(\lambda_0\) is any eigenvalue of \(A\) with \(|\lambda_0| = r_\sigma(A)\), then we may write \(\lambda_0 = r_\sigma(A) \cdot e^{i\theta}\) and make the estimate for \(\rho > r_\sigma(A)\)

\[
\left\| (\rho e^{i\theta} - \lambda_0)^m ((\rho e^{i\theta}) I - A)^{-1} \right\|_\infty = (\rho - r_\sigma(A))^m \left\| \sum_{k=0}^{\infty} \frac{1}{\rho^{k+1} e^{i(k+1)\theta}} A^k \right\|_\infty \\
\leq (\rho - r_\sigma(A))^m \left\| \sum_{k=0}^{\infty} \frac{1}{\rho^{k+1}} A^k \right\|_\infty \\
= \left\| (\rho - r_\sigma(A))^m \sum_{k=0}^{\infty} \frac{1}{\rho^{k+1}} A^k \right\|_\infty \to \|L\|_\infty \quad \text{as} \quad \rho \downarrow |\lambda_0| .
\]

Here the inequality on the second line follows easily from the fact that the \(\| \cdot \|_\infty\)-norm of a matrix is the sum of the absolute values of the entries in a row of the matrix. It follows from this estimate that no entry of \((\rho e^{i\theta} - \lambda_0)^m ((\rho e^{i\theta}) I - A)^{-1}\) can become unbounded as \(\rho \downarrow |\lambda_0| = r_\sigma(A)\), and that implies that the poles of the entries of \((\lambda I - A)^{-1}\) cannot have order \(\geq m\), establishing the remaining part of assertion (d). Finally, suppose that \(\lambda_0 \in \mathbb{R}\) has the property that \((\lambda_0 I - A)^{-1} \geq 0\). It is well known and easy to prove (the spectral mapping theorem) that \(\sigma((\lambda_0 I - A)^{-1}) = \left\{ \frac{1}{\lambda_0 - \lambda} : \lambda \in \sigma(A) \right\}\). Under the mapping
\[ \zeta \rightarrow \frac{1}{\lambda_0 - \zeta}, \text{ the circle } \{\zeta : |\zeta| = r_\sigma(A)\} \text{ goes over to a circle of which the real axis is a diameter that cuts the circle at } \frac{1}{\lambda_0 - r_\sigma(A)} \text{ and } \frac{1}{\lambda_0 + r_\sigma(A)}; \text{ then } \sigma((\lambda_0 I - A)^{-1}) \text{ is contained in the union of that circle and its interior. If } 0 < \lambda_0 < r_\sigma(A), \text{ then the two points at which the real axis cuts the circle would be } \frac{1}{\lambda_0 - r_\sigma(A)} < 0 \text{ and } \frac{1}{\lambda_0 + r_\sigma(A)} > 0, \text{ with absolute values } \frac{1}{r_\sigma(A) - \lambda_0} \text{ and } \frac{1}{\lambda_0 + r_\sigma(A)} \text{ respectively. The former of these (the negative one) has the larger absolute value and belongs to } \sigma((\lambda_0 I - A)^{-1}) \text{ because } r_\sigma(A) \in \sigma(A); \text{ therefore } -r_\sigma((\lambda_0 I - A)^{-1}) \in \sigma((\lambda_0 I - A)^{-1}) \text{ but } r_\sigma((\lambda_0 I - A)^{-1}) \notin \sigma((\lambda_0 I - A)^{-1}), \text{ which is a situation that assertion (a) tells us is impossible if } (\lambda_0 I - A)^{-1} \geq 0. \text{ We must therefore have } \lambda_0 > r_\sigma(A), \text{ and thus we have proved assertion (e).}

A contrapositive version of assertion (d) is occasionally useful: if } A \geq 0 \text{ and the vector } x_0 \text{ and positive real } \lambda_0 \text{ are such that } x_0 \not\preceq 0 \text{ but } (\lambda_0 I - A)x_0 \succeq 0, \text{ then one of } \lambda_0 \in \sigma(A) \text{ or } \lambda_0 < r_\sigma(A) \text{ must hold. Otherwise we would have } (\lambda_0 I - A)^{-1} \geq 0 \text{ and therefore } x_0 = (\lambda_0 I - A)^{-1}(\lambda_0 I - A)x_0 \succeq 0.

Just as an example: consider the Markov transition matrix } A = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{bmatrix}. \text{ Clearly } A \succeq 0 \text{ and } \|A\|_\infty = 1, \text{ and } [1,1,1]^T \text{ is a nonnegative column eigenvector of } A \text{ belonging to the eigenvalue } 1; \text{ but we need a row (left) eigenvector of } A \text{ belonging to } 1 \text{ to give } A\text{'s stationary probability distribution. It is routine to solve for it and get } [2/13, 6/13, 5/13] \text{ as a row vector fixed under right-multiplication by } A; \text{ the remaining eigenvalues are } -\frac{1}{2} \pm \frac{\sqrt{3}}{6}, \text{ both of which are indeed of absolute value } < 1. \text{ Since } \lambda = 1 \text{ is a simple root of the characteristic polynomial } \chi_A(\lambda) = \lambda^3 - \frac{5}{6} \lambda - \frac{1}{6}, \text{ the poles in the entries of } (\lambda I - A)^{-1} \text{ are of order at most } 1, \text{ and some pole has order exactly } 1. \text{ It is a routine computation}^{(5)} \text{ to show that}

\[
\lim_{\lambda \to 1}(\lambda - 1)(\lambda I - A)^{-1} = \begin{bmatrix} 2/13 & 6/13 & 5/13 \\ 2/13 & 6/13 & 5/13 \\ 2/13 & 6/13 & 5/13 \end{bmatrix}
\]

which (since its column vectors are multiples of } [1,1,1]^T \text{ as well) illustrates that the proof of the Perron-Frobenius result actually can generate numbers (at least in small dimensions).

4. Iterative Solution of Linear Systems. Much of this material can be found in §11 of the separate notes on linear algebra, but in somewhat disguised form: it can be viewed as a form of iterative correction.

Suppose that the coefficient matrix } A \text{ of a system of linear equations } Ax = b \text{ is fairly sparse: that is, most of its entries are zeros, and suppose also that its diagonal entries are quite large in comparison to its other entries. Or—to be more general—suppose it somehow may make sense to write } A \text{ in some form } A = B - C \text{ in which } B \text{ is an invertible matrix for which } B^{-1}z \text{ is somehow easy or cheap to compute, while } C \text{ is “whatever is left of } A.” \text{ One can then treat } B \text{ as an “approximate inverse of } A, \text{ view the equation } Ax = b \text{ as being of the form } Bx - Cx = b \text{ and therefore of the form } Bx = Cx + b \text{ and therefore of the form } x = B^{-1}Cx + B^{-1}b, \text{ so that the solution } x \text{ appears as a fixed-point of the mapping } x \mapsto F(x) \equiv B^{-1}Cx + B^{-1}b.

From the results of the preceding §8, we now know that if } r_\sigma(B^{-1}C) \text{ can be shown to be } < 1, \text{ we shall have } \|B^{-1}C\| < 1 \text{ in some norm on } \mathbb{C}^n. \text{ In this norm the mapping } F : \mathbb{C}^n \to \mathbb{C}^n \text{ is a contraction, since } \|F(x) - F(z)\| = \|B^{-1}Cx - B^{-1}Cz\| \leq \|B^{-1}C\| \|x - z\|, \text{ and the iteration will then converge to a fixed point by general contraction-mapping fixed-point theory. The fixed point will then be a solution of the original equation.}

\(^{(5)}\) \text{ for MATLAB or Maple . . .}
The first and least complicated scheme of this type is Gauss-Jacobi iteration: suppose the matrix of coefficients $A$ is written as $A = D - L - U$, where $D = (d_{ii})$ is the diagonal of $A$, $L = (l_{ij}, i > j)$ the lower-triangular part—whose entries equal those of $A$ strictly below the diagonal, but are zeros on and above the diagonal—and $U = (u_{ij}, i > j)$ the upper-triangular part, defined similarly. The iteration then takes the form

$$F(x) = D^{-1}[L + U]x + D^{-1}b,$$

and the condition required to make $F(\cdot)$ a contraction mapping is that $r(\cdot D^{-1}[L + U]) < 1$. The easiest sufficient condition for this is that $A$ be strictly diagonally dominant, since then $\|D^{-1}[L + U]\|_\infty < 1$: the condition for strict diagonal dominance, namely that $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ hold for all rows $i = 1, \ldots, n$, is clearly logically equivalent to the condition

$$\max_i \left\{ \frac{\sum_{j \neq i} |l_{ij}| + |u_{ij}|}{|d_{ii}|} \right\} < 1$$

for $\|D^{-1}[L + U]\|_\infty < 1$. In certain important special cases the requirement of strict diagonal dominance can be relaxed slightly. For example, suppose that $A$ is tridiagonal, and that $|a_{ii}| \geq |a_{i,i-1}| + |a_{i,i+1}|$ holds for $1 < i < n$, with $|a_{11}| \geq |a_{12}|$ and $|a_{nn}| \geq |a_{n,n-1}|$; suppose that all the entries in the immediately sub- and super-diagonals are nonzero (so there are no zeros on the diagonal); and suppose strict inequality holds in at least one of these inequalities. Then no eigenvalue of $D^{-1}[L + U]$ can be $\geq 1$ in absolute value. Indeed, if $\lambda$ were such an eigenvalue and $x = [x_1, \ldots, x_n]^T$ an eigenvector that belonged to it, we would have

$$\lambda x_i = -\frac{a_{i,i-1}}{a_{ii}} x_{i-1} - \frac{a_{i,i+1}}{a_{ii}} x_{i+1} \quad \text{for } 1 < i < n,$$

with the obvious modifications at $i = 1$ and $i = n$. This would imply

$$|a_{ii}| \cdot |\lambda||x_i| \leq |a_{i,i-1}| |x_{i-1}| + |a_{i,i+1}| |x_{i+1}| \quad \text{for } 1 < i < n,$$

again with the obvious modifications at the endpoints. If all the $|x_i|$ are equal—equal to 1, with no loss of generality—this gives

$$|a_{ii}| \cdot |\lambda| \leq |a_{i,i-1}| + |a_{i,i+1}| \leq |a_{ii}| \quad \text{for } 1 < i < n,$$

and similarly at $i = 1$ and $i = n$; but since the second inequality is strict for at least one value of $i$, we can conclude that $|\lambda| < 1$. In any other case, there are indices for which $|x_i| = \|x\|_\infty$, which norm we may assume $= 1$ without loss of generality; choose such an $i$ such that one of the adjacent indices $i - 1$ or $i + 1$ has $|x_{i-1}| < 1$ or $|x_{i+1}| < 1$. Then we have

$$|a_{ii}| \cdot |\lambda| = |a_{ii}| \cdot |\lambda||x_i| \leq |a_{i,i-1}| |x_{i-1}| + |a_{i,i+1}| |x_{i+1}| < |a_{i,i-1}| + |a_{i,i+1}| \leq |a_{ii}|$$

(with the obvious modifications if $i = 1$ or $i = n$) because at least one of $|x_{i-1}| < 1$ or $|x_{i+1}| < 1$ holds, and so again $|\lambda| < 1.$(7)

From the “equation-by-equation” standpoint, Gauss-Jacobi iteration looks like this: if $x^{(k)}$ is the $k$-th iterate of $F(\cdot)$, then the $(k+1)$-st iterate is obtained by solving the equations (with the obvious modifications for $i = 1$ and $i = n$)

$$a_{ii}x_i^{(k+1)} = -a_{i1}x_1^{(k)} - \cdots - a_{i,i-1}x_{i-1}^{(k)} - a_{i,i+1}x_{i+1}^{(k)} - \cdots - a_{in}x_n^{(k)} + b_i$$

for $x_i^{(k+1)}$, where $i$ runs through $1, 2, \ldots, n$ in increasing order. This scheme uses “old” values of $x_j^{(k)}$ for the entire “pass” through the equations, even though the “new” values $x_j^{(k+1)}$ are available for all components $x_j$

(6) See the notes on spline interpolation for one class of examples; the matrices that occur in discretizing Sturm-Liouville operators form another class of examples.

(7) This is only the simplest of a large class of results of this kind that can be derived for “banded matrices.”
with \( j < i \). It seems perfectly reasonable that one should use the “new” values everywhere they are available, and **Gauss-Seidel iteration** does exactly that. The equations set off above are therefore replaced by the equations (with the obvious modifications for \( i = 1 \) and \( i = n \))

\[
a_{ii}x_i^{(k+1)} = -a_{i1}x_1^{(k+1)} - \cdots - a_{i,i-1}x_{i-1}^{(k+1)} - a_{i,i+1}x_{i+1}^{(k)} - \cdots - a_{in}x_n^{(k)}b_i
\]

for \( x_i^{(k+1)} \), where \( i \) runs through \( 1, 2, \ldots, n \) in increasing order. In terms of the matrices \( D, L \) and \( U \), these equations can be rewritten in the vector-matrix form

\[
DX^{(k+1)} = Lx^{(k+1)} + Ux^{(k)} + b
\]

or, upon solving this relation for \( x^{(k+1)} \), the form

\[
x^{(k+1)} = [D - L]^{-1}Ux^{(k)} + [D - L]^{-1}b,
\]

giving the iteration the form

\[
x \mapsto [D - L]^{-1}Ux + [D - L]^{-1}b.
\]

So now one has to investigate the spectral radius of the matrix \((D - L)^{-1}U\). One should note that in all cases in this setting the matrix \((D - L)\) is a lower-triangular matrix with positive elements on the diagonal, so \((D - L)^{-1}\) exists simply because it is possible to solve the equation \((D - L)x = y\) for any given \( y \in \mathbb{C}^n \) by simple forward-solution of the equations whose coefficient matrix is the lower-triangular matrix \(D - L\); the fact that \( U \) must have \( 0 \) in its spectrum (its first column consists only of zeros, so \( Ue_1 = 0 \) where \( e_1 \) is the first standard basis vector) gives one some hope that its spectral radius might be smaller, which would imply that these iterates might converge more rapidly than those of the Gauss-Jacobi iteration. Unfortunately, that is not always what happens; however, there are some important special cases in which it does. While the most obvious and important applications of the following result involve the matrices that have to be inverted to solve discretized (or finite-element) boundary value problems, the result can be stated in a form that applies to more general situations than those matrices represent.

**Lemma:** Suppose \( D, L \) and \( U \) are \( n \times n \) matrices with nonnegative entries, such that \((D - L)^{-1}\) and \((D - L)^{-1}\) exist and have nonnegative entries. Then

(a) If \( r_\sigma(D^{-1}(L + U)) \leq 1 \), then \( r_\sigma((D - L)^{-1}U) \leq r_\sigma(D^{-1}(L + U)) \).

(b) If \( r_\sigma(D^{-1}(L + U)) \geq 1 \), then \( r_\sigma((D - L)^{-1}U) \leq r_\sigma(D^{-1}(L + U)) \).

**Proof.** Of (a): First we consider the case of (a) in which \( \lambda_0 = r_\sigma((D - L)^{-1}U) \leq 1 \). If \( \lambda_0 \in \sigma(D^{-1}(L + U)) \) or if \( \lambda_0 = 0 \) there is nothing to prove, since in either of those cases automatically \( r_\sigma((D - L)^{-1}U) = \lambda_0 \leq r_\sigma(D^{-1}(L + U)) \). Otherwise, let \( x_0 \geq 0 \) be a nonnegative eigenvector of \((D - L)^{-1}U\) belonging to \( \lambda_0 \). Then we have

\[
(D - L)^{-1}Ux_0 = \lambda_0 x_0
\]

\[
Ux_0 = \lambda_0 (D - L)x_0
\]

\[
\lambda_0 Lx_0 + Ux_0 = \lambda_0 Dx_0
\]

\[
Lx_0 + Ux_0 = \lambda_0 Dx_0 + (1 - \lambda_0)Lx_0
\]

\[
(D^{-1}(L + U))x_0 = \lambda_0 x_0 + (1 - \lambda_0)D^{-1}Lx_0
\]

\[
-(1 - \lambda_0)D^{-1}Lx_0 = [\lambda_0 I - D^{-1}(L + U)]x_0.
\]

If \( \lambda_0 = 1 \) then the last set-off line shows that \( \lambda_0 \in \sigma(\lambda_0 I - D^{-1}(L + U)) \), a possibility we have already considered and excluded. So we can now write

\[
[\lambda_0 I - D^{-1}(L + U)]^{-1}((1 - \lambda_0)D^{-1}Lx_0) = -x_0 \leq 0.
\]

Because \((D - L)x_0 \geq 0 \) (this vector cannot be zero, since the r. h. s. of the last set-off line is nonzero), we see that the resolvent \([\lambda_0 I - D^{-1}(L + U)]^{-1}\) is not a matrix with nonnegative entries. But then \( \lambda_0 = r_\sigma((D - L)^{-1}U) < r_\sigma(D^{-1}(L + U)) \), as desired.
The proof of (b) is similar: let $1 \leq \lambda_1 = r_{\sigma}(D^{-1}(L + U))$. If $\lambda_1 \in \sigma((D - L)^{-1}U)$ there is nothing to prove. Otherwise, let $x_1 \succeq 0$ be a nonzero nonnegative eigenvector of $D^{-1}(L + U)$ belonging to $\lambda_1$; then
\[
D^{-1}(L + U)x_1 = \lambda_1 x_1 \\
Lx_1 + Ux_1 = \lambda_1 Dx_1 \\
(1 - \lambda_1)Jx_1 + Ux_1 = \lambda_1 (D - L)x_1 \\
(1 - \lambda_1)(D - L)^{-1}Lx_1 + (D - L)^{-1}Ux_1 = \lambda_1 x_1 \\
-(\lambda_1 - 1)(D - L)^{-1}Lx_1 = \lambda_1 x_1 - (D - L)^{-1}Ux_1 = (\lambda_1 I - (D - L)^{-1}U)x_1.
\]

We have already ruled out the possibility that $\lambda_1 \in \sigma((D - L)^{-1}U)$, so we must have $\lambda_1 > 1$ and $Lx_1 \neq 0$ and $\succeq 0$. Again the last set-of-line shows that $(\lambda_1 I - (D - L)^{-1}U)^{-1} \geq 0$ is impossible, so $r_{\sigma}(D^{-1}(L + U)) = \lambda_1 \leq r_{\sigma}((D - L)^{-1}U)$.

Finally, to finish proving (a) we must exclude the possibility that $\lambda_0 = r_{\sigma}((D - L)^{-1}U) > 1$ but somehow $r_{\sigma}(D^{-1}(L + U)) < 1$. Assuming $r_{\sigma}(D^{-1}(L + U)) < 1$, and with $\lambda_0$ and $x_0$ as in the proof of (a), if we had $\lambda_0 > 1$ we would have
\[
(D - L)^{-1}Ux_0 = \lambda_0 x_0 \\
Ux_0 = \lambda_0 (D - L)x_0 \\
\lambda_0 Lx_0 + Ux_0 = \lambda_0 Dx_0 + (\lambda_0 - 1)Ux_0 \\
\lambda_0 D^{-1}(L + U)x_0 = \lambda_0 x_0 + (\lambda_0 - 1)D^{-1}Ux_0 \\
-(\lambda_0 - 1)D^{-1}Ux_0 = \lambda_0 [I - D^{-1}(L + U)]x_0.
\]

Because we have assumed $r_{\sigma}(D^{-1}(L + U)) < 1$, the resolvent $[I - D^{-1}(L + U)]^{-1}$ has nonnegative entries, so applying it to both sides of the last set-of-line gives the expected contradiction
\[
-(\lambda_0 - 1) [I - D^{-1}(L + U)]^{-1} D^{-1}Ux_0 = \lambda_0 x_0
\]
where the l. h. s. is clearly a vector with no positive components and the r. h. s. is just as clearly a nonzero vector with some positive components.

**Theorem:** Suppose $D$, $L$ and $U$ are $n \times n$ matrices with nonnegative entries, such that $D^{-1}$ and $(D - L)^{-1}$ exist and have nonnegative entries. Then exactly one of the following alternatives can occur:

1. $r_{\sigma}((D - L)^{-1}U) \leq r_{\sigma}(D^{-1}(L + U)) < 1$;
2. $1 < r_{\sigma}((D^{-1}(L + U)) \leq r_{\sigma}((D - L)^{-1}U)$;
3. $r_{\sigma}((D - L)^{-1}U) = r_{\sigma}(D^{-1}(L + U)) = 1$.

**Proof.** The proof follows directly by checking cases in the immediately-preceding lemma.

The most important special case of what this theorem says, in words and a sort of picture, is that if the distribution of signs in the matrix $A$ follows the pattern
\[
\begin{bmatrix}
+ & - & - & \cdots & - \\
- & + & - & \cdots & - \\
- & - & + & \cdots & - \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
- & - & - & \cdots & +
\end{bmatrix}
\]
(where the diagonal entries must be strictly positive, but the off-diagonal entries may be negative or zeros) then the Gauss-Seidel iteration will converge at least as well as the Gauss-Jacobi or blow up at least as
badly as the Gauss-Jacobi, with an indeterminate sort of behavior if both iterations are given by matrices having spectral radius 1. However, one should note that the theorem is somewhat more general than this: for example, if $D$ has the form of a permutation matrix times a positive diagonal matrix, then it satisfies the requirement that both $D$ and $D^{-1}$ have nonnegative entries. The theorem can also be adapted to say something about the iterative solution of integral equations whose kernels satisfy certain positivity assumptions.