Appendix: Matrix Estimates and the Perron-Frobenius Theorem.

This Appendix will first present some well known estimates. For any $m \times n$ matrix $A = [a_{ij}]$ over the real or complex numbers, it will be convenient to use the notation

$$\|A\| \;=\; \sum_{i,j} |a_{ij}| \;.$$

For any product matrix AB, note that

$$\|AB\| \;=\; \sum_{i,k} \Big| \sum_j a_{ij} b_{jk} \Big| \;\le\; \sum_{i,j,j',k} |a_{ij} b_{j'k}| \;=\; \|A\|\,\|B\| \;. \qquad (A:1)$$

In particular, if A is a square matrix and λ is one of its eigenvalues, then we can find a non-zero column vector X with

$$\lambda X \;=\; AX \;, \qquad \text{hence} \qquad \lambda^k X \;=\; A^k X \qquad \text{and} \qquad \|\lambda^k X\| \;\le\; \|A^k\|\,\|X\| \;.$$

Dividing by $\|X\|$, it follows that

$$|\lambda|^k \;\le\; \|A^k\| \;.$$

(Here λ and X may be complex, even if A is real.)

Define the *spectral radius* $|\lambda|_{\max}$ of an $n \times n$ matrix A to be the largest of the absolute values $|\lambda_1|, \ldots |\lambda_n|$ of its n eigenvalues. Then it follows that

$$|\lambda|_{\max}^k \;\le\; \|A^k\| \;. \qquad (A:2)$$

Lemma A.1. *The sequence of powers A^k converges uniformly to zero if and only if the spectral radius satisfies $|\lambda|_{\max} < 1$.*

(However, examples such as

$$A \;=\; \begin{bmatrix} 0.9 & 1000 \\ 0 & 0.9 \end{bmatrix}$$

show that the convergence may be very slow.)

Proof of A.1. If $|\lambda|_{\max} \ge 1$, then $\|A^k\| \ge 1$ for all k by (A : 2). The proof for the case $|\lambda|_{\max} < 1$ will be by induction on n. Evidently the statement is clear when $n = 1$. For $n > 1$, we can always find a non-singular complex matrix S so that SAS^{-1} has zeros below the diagonal. We can then write

$$SAS^{-1} \;=\; \begin{bmatrix} P & Q \\ 0 & R \end{bmatrix} \;,$$

where P and R are square matrices. Since each eigenvalue of P or R is also an eigenvalue of A, we may assume inductively that $\|P^k\| \to 0$ and $\|R^k\| \to 0$ as $k \to \infty$. Now set

$$SA^kS^{-1} \;=\; \begin{bmatrix} P & Q \\ 0 & R \end{bmatrix}^k \;=\; \begin{bmatrix} P^k & Q(k) \\ 0 & R^k \end{bmatrix} \;,$$

where

$$Q(k+\ell) \;=\; P^k Q(\ell) + Q(k) R^\ell \;.$$

Let q_m be the maximum of $\|Q(k)\|$ over the range $2^m \le k < 2^{m+1}$. Choosing m large enough so that $\|P^k\| < 1/4$ and $\|R^k\| < 1/4$ for $k \ge 2^m$, it follows easily that

$q_{m+1} \le q_m/2$, and more generally, $q_{m+r} \le q_m/2^r$, tending to zero. It follows that $\|Q(k)\|$ tends to zero as $k \to \infty$, and hence that $\|A^k\|$ does also. □

Corollary A.2 *Let* $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ *be a complex power series with radius of convergence* $0 < r \le \infty$ *. If* A *is a real or complex* $n \times n$ *matrix with spectral radius satisfying* $|\lambda|_{\max} < r$ *, then the corresponding series*

$$f(A) \;=\; a_0 I + a_1 A + a_2 A^2 + \cdots \ ,$$

converges absolutely.

Proof. Choose η with $|\lambda|_{\max} < \eta < r$. Then the powers of A/η converge to zero by A.1, and hence are bounded,

$$\|A^k/\eta^k\| \;\le\; c$$

for some uniform constant c . On the other hand, the series $\sum |a_k \eta^k|$ converges, and

$$\|a_k A^k\| \;\le\; c\,|a_k \eta^k| \ ,$$

hence the series $\sum \|a_k A^k\|$ also converges. □

Example. If the spectral radius of A satisfies $|\lambda|_{\max} < 1$, then it follows that the series $\sum A^k$ converges absolutely, and the identity

$$(I - A)^{-1} \;=\; I + A + A^2 + A^3 + \cdots \qquad (A:3)$$

can be verified by multiplying both sides of this equation by $I - A$.

We next prove two formulas for the spectral radius of a matrix $A = [a_{ij}]$ in terms of the powers A^k . We will also make use of the trace

$$\operatorname{trace} A \;=\; a_{11} + a_{22} + \cdots + a_{nn}$$

of the matrix A , as well as the traces of its powers A^k .

Theorem A.3. *The spectral radius of any real or complex* $n \times n$ *matrix* A *is given by*

$$|\lambda|_{\max} \;=\; \lim_{k\to\infty} \|A^k\|^{1/k} \;=\; \limsup_{k\to\infty} \left|\operatorname{trace}(A^k)\right|^{1/k} .$$

Here it is essential to work with the *lim sup*, since the limit may not exist. For example, if

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad \text{then} \qquad \operatorname{trace}(A^k) \;=\; \begin{cases} 3 & \text{if} \quad k \equiv 0 \pmod 3 \\ 0 & \text{otherwise} \ , \end{cases}$$

so that $\limsup |\operatorname{trace} A^k|^{1/k} = +1$ but $\liminf |\operatorname{trace} A^k|^{1/k} = 0$. (The eigenvalues of this matrix are the three cube roots of unity.)

Proof of A.3. Since $\|A^k\| \ge |\lambda|^k_{\max}$ by (A : 2), it certainly follows that

$$\|A^k\|^{1/k} \;\ge\; |\lambda|_{\max} \ .$$

On the other hand, if $c = |\lambda|_{\max} + \epsilon$, then it follows from A.1 that the powers of A/c converge to zero. Hence we can choose some constant c' so that $\|A^k\| \le c^k c'$ for all k . It follows that

$$\|A^k\|^{1/k} \;\le\; c\,\sqrt[k]{c'} \qquad \text{which converges to} \qquad c \;=\; |\lambda|_{\max} + \epsilon \qquad \text{as} \qquad k \to \infty \ .$$

Since ϵ can be arbitrarily small, this proves that $\|A^k\|^{1/k} \to |\lambda|_{\max}$ as $k \to \infty$.

The inequality

$$\limsup_{k\to\infty} |\mathrm{trace}(A^k)|^{1/k} \;\leq\; |\lambda|_{\max}$$

follows immediately, since $|\mathrm{trace}(A^k)| \leq \|A^k\|$. To obtain a lower bound for this lim sup, let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A, and note that

$$\mathrm{trace}(A^k) \;=\; \lambda_1^k + \cdots + \lambda_n^k .$$

For each non-zero λ_j, let $\mu_j = \lambda_j/|\lambda_j|$. We will need the following.

Lemma A.4. *Let $\mu_1, \ldots, \mu_q$ be complex numbers with $|\mu_j| = 1$. For any $\epsilon > 0$ there exist infinitely many values of k such that $|\mu_j^k - 1| < \epsilon$ and hence $\mathrm{Re}\,(\mu_j^k) > 1 - \epsilon$ for every μ_j.*

Proof. If $p > 2\pi/\epsilon$, then we can cover the unit circle by p intervals $J_1, \ldots, J_p$ of length less than ϵ. Hence we can cover the torus $T^q = S^1 \times \cdots \times S^1 \subset {}^q$ by p^q boxes $J_{i(1)} \times \cdots \times J_{i(q)}$. Consider the sequence of points $\boldsymbol{\mu}_h = (\mu_1^h, \ldots, \mu_q^h)$ on the torus T^q. Given any $p^q + 1$ of these points, there must be at least two, say $\boldsymbol{\mu}_h$ and $\boldsymbol{\mu}_\ell$, belonging to the same box. Setting $k = |h - \ell|$, it follows that $|\mu_j^k - 1| = |\mu_j^h - \mu_j^\ell| < \epsilon$ for all j, as required. Since ϵ can be arbitrarily small, we can construct infinitely many such integers k by this procedure. □

The proof of A.3 continues as follows. Taking $\epsilon = 1/2$, we can choose k as above, and conclude that the real part satisfies $\mathrm{Re}\,((\lambda_j/|\lambda_j|)^k) > 1/2$ whenever $\lambda_j \neq 0$. Multiplying by $|\lambda_j|^k$, it follows that

$$\mathrm{Re}\,(\lambda_j^k) \;\geq\; \tfrac{1}{2}\,|\lambda_j|^k$$

for *every* eigenvalue λ_j. Thus

$$|\mathrm{trace}(A^k)| \;\geq\; \mathrm{Re}\,(\mathrm{trace}(A^k)) \;=\; \sum_j \mathrm{Re}\,(\lambda_j^k) \;\geq\; \tfrac{1}{2}\sum_j |\lambda_j|^k \;\geq\; \tfrac{1}{2}\,|\lambda|_{\max}^k$$

for infinitely many values of k. The required inequality

$$\limsup_{k\to\infty} |\mathrm{trace}(A^k)|^{1/k} \;\geq\; |\lambda|_{\max}$$

follows immediately. □

The Perron-Frobenius Theorem.

We say that an $m \times n$ matrix is ***non-negative*** if all of its entries are real and non-negative, and ***strictly positive*** if all of the entries are strictly greater than zero. The following helps to show the special properties of such matrices.

Lemma A.5. *If the non-negative $n \times n$ matrix A has an eigenvector X which is strictly positive, then the corresponding eigenvalue λ (where $AX = \lambda X$) is equal to the spectral radius $|\lambda|_{\max}$. Furthermore $\lambda > 0$ unless A is the zero matrix.*

Proof. If A is not the zero matrix, then since X is strictly positive, it follows that the vector AX has at least one strictly positive component. The equation $AX = \lambda X$ then implies that $\lambda > 0$.

Now since $A^k X = \lambda^k X$, it follows that $\|A^k X\| = \lambda^k \|X\|$. If X has components $x_j \ge c > 0$, then $\|A^k X\| \ge c\|A^k\|$, hence

$$\lambda^k \|X\| \ \ge c\|A^k\| \ .$$

Taking the k-th root and passing to the limit as $k \to \infty$, using A.3, we get

$$\lambda \ \ge \ \lim \|A^k\|^{1/k} \ = \ |\lambda|_{\max} \ , \qquad \text{and hence} \qquad \lambda \ = \ |\lambda|_{\max} \ ,$$

as required. □

Definition. We will say that a non-negative $n \times n$ matrix A is *transitive* if for every i and j there exists an integer $k > 0$ so that the ij-th entry $a_{ij}^{(k)}$ of the k-th power matrix A^k is non-zero. (Some authors use the term "*irreducible*" for such matrices.) We can interpret this condition graphically as follows. Let $G(A)$ be the graph with one vertex v_i for eacch integer between 1 and n , and with a directed edge from v_i to v_j if and only if $a_{ij} > 0$. Then A is transitive if and only if we can get from any vertex to any other vertex by following these directed edges. Since such a path can always be chosen so as to pass through each vertex at most once, we may assume that its length is strictly less than n . Hence a completely equivalent condition would be that the matrix sum

$$A + A^2 + A^3 + \cdots + A^{n-1}$$

is strictly positive.

Theorem A.6. (O. Perron and G. Frobenius). *Every non-negative $n \times n$ matrix A has a non-negative eigenvector, $AY = \lambda Y$, with the property that the associated eigenvalue λ is equal to the spectral radius $|\lambda|_{\max}$. If the matrix A is transitive, then there is only one non-negative eigenvector Y up to multiplication by a positive constant, and this eigenvector is strictly positive: $y_i > 0$ for all i .*

Proof in the transitive case. We will make use of the *Brouwer Fixed Point Theorem*, which asserts that any continuous mapping $f : \Delta \to \Delta$ from a closed simplex to itself must have at least one fixed point, $X = f(X)$. In fact let Δ be the standard $(n-1)$-dimensional simplex consisting of all non-negative column vectors Y with $\|Y\| = y_1 + \cdots + y_n = 1$. If A is transitive and $Y \in \Delta$, note that $\|AY\| \neq 0$. For if AY were the zero vector, then $A^k Y$ would be zero for all $k > 0$. But choosing some component $y_j > 0$, for any i we can choose some power of A^k so that the (i,j)-th component of A^k is non-zero, and hence so that the i-th component of $A^k Y$ is non-zero. A similar argument shows that any eigenvalue $AY = \lambda Y$ must be strictly positive.

We can now define the required map $f : \Delta \to \Delta$ by the formula

$$f(Y) \ = \ AY/\|AY\| \ .$$

Thus there exists at least one fixed point

$$Y \ = \ f(Y) \ = \ AY/\|AY\| \ .$$

Taking $\lambda = \|AY\| > 0$, it follows that $AY = \lambda Y$ where, as noted above, Y must be strictly positive, with $\lambda > 0$. It then follows from A.5 that $\lambda = |\lambda|_{\max}$.

Now suppose that there were two distinct eigenvectors Y and Y' in Δ, both necessarily with eigenvalue equal to $|\lambda|\max$. Then every point on the straight line joining Y and Y' would also be an eigenvector. But such a line would contain boundary points of Δ, contradicting the assertion that any such eigenvector must be strictly positive. This proves A.6 in the transitive case.

Proof in the non-transitive case. We will give a constructive procedure for reducing the non-transitive case to the transitive one. The statement of A.6 is trivially true when $n = 1$, so we may assume that A is a non-transitive $n \times n$ matrix with $n \geq 2$. Then there exist indices $i \neq j$ so that the (i, j)-th entry $a_{ij}^{(k)}$ of A^k is zero for all k. Let us say that the index j is "accessible" from i if $a_{ij}^{(k)} > 0$ for some $k > 0$. Evidently this is a transitive relation between indices. After conjugating A by a permutation matrix, we may assume that $i = 1$, and that all of the indices which are accessible from i are listed first, with all of the remaining indices at the end. It follows that A can be written as a block matrix of the form

$$\begin{bmatrix} P & 0 \\ Q & R \end{bmatrix},$$

where P and R are square matrices. Here we may assume by induction on n that Theorem A.5 is true for P and R. Since the eigenvalues of such a block matrix are just the eigenvalues of P together with the eigenvalues of R, it certainly follows that the spectral radius $|\lambda|_{\max}$ is itself an eigenvalue. To complete the proof, we must solve the equation

$$\begin{bmatrix} P & 0 \\ Q & R \end{bmatrix} \begin{bmatrix} Y \\ Y' \end{bmatrix} = \lambda \begin{bmatrix} Y \\ Y' \end{bmatrix},$$

with $\lambda = |\lambda|_{\max}$, and with Y and Y' non-negative. If λ is an eigenvalue of R, we can simply take Y' to be the appropriate eigenvector for R, with $Y = 0$. Otherwise, we can assume that the spectral radius of R is strictly less than $\lambda = |\lambda|_{\max}$. Let $PY = \lambda Y$ be a non-negative eigenvalue for P. Then we must find a non-negative vector Y' which satisfies the equation $QY + RY' = \lambda Y'$, or in other words

$$(\lambda I - R)Y' = QY \qquad \text{or} \qquad Y' = (I - R/\lambda)^{-1} QY/\lambda$$

where I is the identity matrix of appropriate size. Since QY is certainly non-negative, it only remains to show that $(I - R/\lambda)^{-1}$ is non-negative. But according to (A : 3) this inverse matrix can be expressed as the sum of a convergentpower series

$$(I - R/\lambda)^{-1} = I + (R/\lambda) + (R/\lambda)^2 + \cdots,$$

where all summands on the right are non-negative. This completes the proof of A.6. □

We conclude with two problems.

Problem A-a. Iteration of $Y \mapsto AY$. For any real or complex $n \times n$ matrix A, show that there is a lower dimensional subspace V in the vector space of $n \times 1$ column vectors so that, for any fixed $Y \notin V$. we have

$$|\lambda|_{\max} = \lim_{k \to \infty} \|A^k Y\|^{1/k}.$$

(To prove this in the complex case, it is convenient to conjugate A to an appropriate normal form. It is then necessary to check that the real case follows.)

Problem A-b. Simplicity of the largest eigenvector, and computation. Suppose that A is an $n \times n$ matrix with a left eigenvector $\hat{X}A = \lambda\hat{X}$ and a right eigenvector $A\hat{Y} = \mu\hat{Y}$ such that $\hat{X}\hat{Y} = 1$. Show that $\lambda = \mu$. If $\lambda \neq 0$, then setting $B = A/\lambda$ we have

$$\hat{X}B \;=\; \hat{X} \qquad \text{and} \qquad B\hat{Y} \;=\; \hat{Y} \qquad \text{with} \qquad \hat{X}\hat{Y} = 1\ .$$

If $\hat{X}^{\perp}$ is the hyperplane consisting of all column vectors Y with $\hat{X}Y = 0$, show that the space of all column vectors splits as a dirct sum

$$^{n} \;=\; \hat{X}^{\perp} \,\oplus\, \hat{Y}\ .$$

Show that the linear map $Y \mapsto BY$ carries the hyperplane $\hat{X}^{\perp}$ into itself and fixes each point of the eigenspace $\hat{Y}$. Show that every orbit under this map $Y \mapsto BY$ converges towards some vector in $\hat{Y}$ if and only if all but one of the n eigenvalues of B belong to the open unit disk, $|\lambda_j| < 1$.

Now suppose that A is strictly positive. Let Δ be the simplex consisting of all non-negative Y with $\hat{X}Y = 1$ Show that the associated transformation

$$Y \;\mapsto\; BY \;=\; AY/|\lambda|_{\max}$$

carries Δ into its interior. Conclude that all orbits in Δ converge to $\hat{Y}$. For example we can prove this by constructing a kind of ‘norm’ function N on Δ which is linear on each line segment joining $\hat{Y}$ to the boundary of Δ , taking the value 0 at $\hat{Y}$ and the value $+1$ on $\partial\Delta$. It is then easy to check that $N(BY) \leq cN(Y)$ for some constant $c < 1$. Conclude that all but one of the n eigenvalues of A must satisfy $|\lambda_j| < |\lambda|_{\max}$.

Next suppose only that T is a non-negative transitive matrix. Applying the discussion above to the strictly positive matrix

$$A \;=\; T + T^2 + \cdots + T^{n-1}\ ,$$

we see that all but one of the eigenvalues of A satisfies $|\lambda_j| < |\lambda|_{\max}$. Conclude that the spectral radius of T must be a simple eigenvalue. That is, all but one of the eigenvalues of T must be different from the spectral radius of T .

Show that the positive eigenvector for a transitive matrix T can always be located by iterating the associated map

$$Y \;\mapsto\; AY/\|AY\|\ ,$$

or even the map $Y \mapsto (I+T)Y/\|(I+T)Y\|$.