Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

December 3, 2021 (unfinished!)

Contents

1. Positive and nonnegative matrices ([HorJoh, Chapter 8]) (cont'd) 1
1.1. The Perron theorems (cont'd) 1
1.1.1. Lemmas for the proofs 1

Math 504 Lecture 25

1. Positive and nonnegative matrices ([HorJoh, Chapter 8]) (cont'd)

1.1. The Perron theorems (cont'd)

1.1.1. Lemmas for the proofs

Before we prove the Perron and Perron–Frobenius theorems, we need to state a few lemmas. First, recall the corollary:

Corollary 1.1.1. Let $A \in \mathbb{R}^{n \times n}$ satisfy $A \geq 0$ and $n > 0$. Let $x_1, x_2, \ldots, x_n$ be any n positive reals. Then,

$$\min_{i \in [n]} \sum_{j=1}^{n} \frac{x_i}{x_j} A_{i,j} \leq \rho(A) \leq \max_{i \in [n]} \sum_{j=1}^{n} \frac{x_i}{x_j} A_{i,j}.$$

*Drexel University, Korman Center, 15 S 33rd Street, Philadelphia PA, 19104, USA

Corollary 1.1.2. Let $A \in \mathbb{R}^{n \times n}$ satisfy $A \geq 0$ and $n > 0$.
Let $x \in \mathbb{R}^n$ satisfy $x > 0$.
Let α be a nonnegative real. Then:
(a) If $Ax \geq \alpha x$, then $\rho(A) \geq \alpha$.
(b) If $Ax > \alpha x$, then $\rho(A) > \alpha$.
(c) If $Ax \leq \alpha x$, then $\rho(A) \leq \alpha$.
(d) If $Ax < \alpha x$, then $\rho(A) < \alpha$.

Proof. Write x as $x = \left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}\right)^T$; then, $x_1, x_2, \ldots, x_n > 0$.
(a) Assume that $Ax \geq \alpha x$. The old corollary yields

$$\min_{i \in [n]} \sum_{j=1}^{n} \frac{x_i}{x_j} A_{i,j} \leq \rho(A).$$

Thus,

$$\rho(A) \geq \min_{i \in [n]} \sum_{j=1}^{n} \frac{x_i}{x_j} A_{i,j} = \min_{i \in [n]} x_i \underbrace{\sum_{j=1}^{n} \frac{1}{x_j} A_{i,j}}_{\substack{= (\text{the } i\text{-th entry of } Ax) \geq \frac{\alpha}{x_i} \\ (\text{since } Ax \geq \alpha x)}} \geq \min_{i \in [n]} x_i \cdot \frac{\alpha}{x_i} = \alpha.$$

(b) Use the same argument as before, but with $>$ sign.
(c) Analogous, but now use the other half of the old corollary.
(d) Analogous. □

Corollary 1.1.3. Let $A \in \mathbb{R}^{n \times n}$ satisfy $A > 0$ and $n > 0$ and $\rho(A) = 1$. Let $w \in \mathbb{R}^n$ satisfy $w \geq 0$ and $w \neq 0$.
(a) We always have $Aw > 0$.
(b) If $Aw \geq w$, then $Aw = w > 0$.

Proof. **(a)** For each i, the i-th entry of Aw is $\sum_{j=1}^{n} A_{i,j} w_j$ (where w_j is the j-th entry of w). This is a sum of nonnegative addends, and at least one of these addends is actually positive (since $w \neq 0$ entails that $w_j > 0$ for some j, and then we also have $A_{i,j} > 0$ because $A > 0$). So this sum is positive. Thus we have shown that all entries of Aw are positive. In other words, $Aw > 0$.

(b) Assume that $Aw \geq w$. Let $z := Aw - w$. Then, $z \geq 0$. Hence, if we had $z \neq 0$, then part **(a)** (applied to z instead of w) would yield $Az > 0$, so that $A(Aw - w) > 0$.

Part **(a)** yields $Aw > 0$. If we had $AAw > Aw$, then part **(b)** of the preceding corollary (applied to $x = Aw$ and $\alpha = 1$) would yield $\rho(A) > 1$, which would contradict $\rho(A) = 1$. So we cannot have $AAw > Aw$. In other words, we cannot have $A(Aw - w) > 0$. Thus, we cannot have $z \neq 0$ (by the previous paragraph). So $z = 0$ and therefore $Aw = w$ (since $z = Aw - w$). This also entails $w = Aw > 0$ by part **(a)**. So part **(b)** is proved. □

Definition 1.1.4. Fix $n > 0$. Let $e = (1, 1, \ldots, 1)^T$.

Remark 1.1.5. **(a)** An $n \times n$-matrix A satisfies $Ae = e$ if and only if all row sums of A equal 1.

(b) An $n \times n$-matrix A satisfies $e^T A = e^T$ if and only if all column sums of A equal 1.

Lemma 1.1.6 (crucifix lemma, special case)**.** Let $A \in \mathbb{R}^{n \times n}$ satisfy $A > 0$ and $Ae = e$. Let $y \in \mathbb{R}^n$ satisfy $y^T A = y^T$ and $y \geq 0$ and $y^T e = 1$ (that is, the sum of all entries of y is 1). Then,

$$A^m \to ey^T \qquad \text{as } m \to \infty.$$

Proof. The entries of the vector y are nonnegative reals (since $y \geq 0$) and their sum is 1 (since $y^T e = 1$). Thus, all these entries lie in the interval $[0, 1]$.

From $Ae = e$, we conclude that all row sums of A equal 1. Since $A > 0$, this implies that all entries $A_{i,j}$ of A satisfy

$$0 < A_{i,j} \leq 1.$$

Let

$$\mu := 1 - \min \left\{ A_{i,j} \ \mid \ i, j \in [n] \right\}.$$

Then, $0 \leq \mu < 1$ (by the previous inequality).

Claim 1: For each $i \in [n]$ and each **proper** subset K of $[n]$, we have

$$\sum_{k \in K} A_{i,k} \leq \mu.$$

[*Proof of Claim 1:* Let $i \in [n]$. Let K be a proper subset of $[n]$. Then,

$$\sum_{k \in K} A_{i,k} = 1 - \underbrace{\sum_{k \notin K} A_{i,k}}_{\substack{\geq \min\left\{ A_{i,j} \ \mid \ i,j \in [n] \right\} \\ \text{(since there exists at least one } k \notin K)}} \leq 1 - \min \left\{ A_{i,j} \ \mid \ i, j \in [n] \right\} = \mu,$$

so Claim 1 is proved.]

Now, we claim:

Claim 2: For any $i, j \in [n]$ and any $m \in \mathbb{N}$, we have

$$\left| \left(A^m - ey^T \right)_{i,j} \right| \leq \mu^m.$$

Once Claim 2 is proved, it will follow easily that $\left(A^m - ey^T\right)_{i,j} \to 0$ as $m \to \infty$ (because $0 \leq \mu < 1$), so that $A^m \to ey^T$, and the lemma will thus follow.

[*Proof of Claim 2:* We induct on m:

Base case: We need to show that $\left|\left(I_n - ey^T\right)_{i,j}\right| \leq 1$ for all i, j. This follows from the fact that the entries of y lie in the interval $[0, 1]$ (because $\left(I_n - ey^T\right)_{i,j} = \delta_{i,j} - \underbrace{y_j}_{\in[0,1]} \in [-1, 1]$).

Induction step: Let $p \in \mathbb{N}$. Assume (as the induction hypothesis) that Claim 2 holds for $m = p$. We must now show that it also holds for $m = p + 1$.

Let $B := A^p - ey^T$ and $C := A^{p+1} - ey^T$. So our IH says that $\left|B_{i,j}\right| \leq \mu^p$ for all i, j. Our goal is to show that $\left|C_{i,j}\right| \leq \mu^{p+1}$ for all i, j.

Fix i, j. We have

$$AB = A\left(A^p - ey^T\right) = A^{p+1} - \underbrace{Ae}_{=e} y^T = A^{p+1} - ey^T = C.$$

Hence, $C = AB$, so that

$$C_{i,j} = \sum_{k=1}^{n} \underbrace{A_{i,k}}_{>0} B_{k,j} = \sum_{k \in P} A_{i,k} \left|B_{k,j}\right| - \sum_{k \in N} A_{i,k} \left|B_{k,j}\right|,$$

where

$$P := \left\{k \in [n] \mid B_{k,j} > 0\right\} \qquad \text{and} \qquad N := \left\{k \in [n] \mid B_{k,j} < 0\right\}.$$

However, the entries of the j-th column of B cannot all have the same sign (i.e., both subsets P and N of $[n]$ are proper). The reason for this is that

$$y^T B = y^T \left(A^p - ey^T\right) = \underbrace{y^T A^p}_{\substack{=y^T \\ (\text{since } y^T A = y^T)}} - \underbrace{y^T e}_{=1} y^T = y^T - y^T = 0$$

$$\Longrightarrow \qquad \text{(look at the } j\text{-th entry)}$$

$$\sum_{k=1}^{n} \underbrace{y_k}_{\geq 0} B_{k,j} = 0$$

and that y^T is a nonzero nonnegative vector, so there is a nontrivial linear combination of the entries of the j-th column of B with nonnegative coefficients that is 0.

So both subsets P and N of $[n]$ are proper.

Now, from

$$C_{i,j} = \sum_{k=1}^{n} \underbrace{A_{i,k}}_{>0} B_{k,j} = \sum_{k \in P} A_{i,k} \left|B_{k,j}\right| - \sum_{k \in N} A_{i,k} \left|B_{k,j}\right|,$$

we obtain

$$\left|C_{i,j}\right| = \left|\sum_{k\in P} A_{i,k}\left|B_{k,j}\right| - \sum_{k\in N} A_{i,k}\left|B_{k,j}\right|\right| \leq \max\left\{\sum_{k\in P} A_{i,k}\left|B_{k,j}\right|,\ \sum_{k\in N} A_{i,k}\left|B_{k,j}\right|\right\},$$

since any two nonnegative reals x and y satisfy $|x-y| \leq \max\{x,y\}$. Thus,

$$\left|C_{i,j}\right| \leq \sum_{k\in K} A_{i,k}\left|B_{k,j}\right|,$$

where K is either P or N. In either case, K is a proper subset of $[n]$. Therefore,

$$\left|C_{i,j}\right| \leq \sum_{k\in K} A_{i,k} \underbrace{\left|B_{k,j}\right|}_{\substack{\leq \mu^p \\ \text{(by IH)}}} \leq \mu^p \underbrace{\sum_{k\in K} A_{i,k}}_{\substack{\leq \mu \\ \text{(by Claim 1)}}} \leq \mu^p \mu = \mu^{p+1}.$$

This completes the induction step. Thus, Claim 2 is proved.] $\square$

> **Lemma 1.1.7** (crucifix lemma, general case)**.** Let $A \in \mathbb{R}^{n\times n}$ satisfy $A > 0$. Let $x \in \mathbb{R}^n$ satisfy $Ax = x$ and $x > 0$. Let $y \in \mathbb{R}^n$ satisfy $y^T A = y^T$ and $y \geq 0$ and $y^T x = 1$. Then,
> $$A^m \to xy^T \qquad \text{as } m \to \infty.$$

Proof. Let $x_1, x_2, \ldots, x_n$ be the entries of x; then $x_i > 0$.

Let $D = \operatorname{diag}(x_1, x_2, \ldots, x_n)$. Then, $De = x$.

Now, apply the previous lemma to $D^{-1}AD$ and Dy instead of A and y. Details are LTTR. $\square$

Proof of Perron's theorem. **(a)** Recall the corollary we had a while ago, which said that if $A \geq 0$ satisfies $A_{i,i} > 0$ for some $i \in [n]$, then $\rho(A) > 0$. This clearly applies, since $A > 0$. So part **(a)** is proved.

Next, knowing that $\rho(A) > 0$, we can replace A by $\dfrac{1}{\rho(A)}A$. This way, $\rho(A)$ becomes 1, but nothing else significantly changes. So we have $\rho(A) = 1$ now.

Next, we shall show that there is a positive 1-eigenvector of A. Indeed, from $\rho(A) = 1$, we see that A has an eigenvalue λ with $|\lambda| = 1$. Consider this λ. Pick any nonzero λ-eigenvector $z = (z_1, z_2, \ldots, z_n)^T \in \mathbb{C}^n$ of A. Then, $Az = \lambda z$. Hence,

$$\underbrace{A}_{=|A|}|z| = |A|\cdot|z| \geq |Az| = |\lambda z| = \underbrace{|\lambda|}_{=1}\cdot|z| = |z|.$$

Hence, applying part **(b)** to the last corollary, we conclude that $A|z| = |z| > 0$. Thus, $|z|$ is a positive 1-eigenvector of A.

We have thus constructed a positive 1-eigenvector of A. The same argument (applied to A^T instead of A) yields a positive 1-eigenvector of A^T, and thus (by transposing it) a positive left 1-eigenvector of A. Let these two eigenvectors be x and y. Thus, $x \in \mathbb{R}^n$ satisfies $Ax = x$ and $x > 0$, and $y \in \mathbb{R}^n$ satisfies $y^T A = y^T$ and $y > 0$. Moreover, by scaling y appropriately, we can achieve $y^T x = 1$. Thus, the crucifix lemma yields

$$A^m \to xy^T \qquad \text{as } m \to \infty.$$

This proves part **(e)**.

Remains to prove the uniqueness claims in parts **(c)** and **(d)**, and also parts **(b)** and **(f)**. We can do this in one fell swoop if we can show the following: If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A, then only one of these eigenvalues has absolute value ≥ 1.

This follows easily from $A^m \to xy^T$. Indeed, let (U, T) be the Schur triangularization of A. Then,

$$A = UTU^* = UTU^{-1},$$

and the diagonal entries of T are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A. Hence,

$$A^m = \left(UTU^{-1}\right)^m = UT^m U^{-1}.$$

Hence, taking the limit as $m \to \infty$, we get

$$xy^T = UT^\infty U^{-1},$$

where $T^\infty = U^{-1} xy^T U$ is a triangular matrix with diagonal entries $\lambda_1^\infty, \lambda_2^\infty, \ldots, \lambda_n^\infty$. This shows that all $\lambda_1, \lambda_2, \ldots, \lambda_n$ have absolute value ≤ 1, and the ones that have absolute value 1 must equal 1. Moreover, if more than one of the λ_is would equal 1, then $UT^\infty U^{-1}$ would have rank > 1, but then it could not equal xy^T (since rank $\left(xy^T\right) \leq 1$). Qed. $\square$

Theorem 1.1.8 (Perron theorem)**.** Let $A \in \mathbb{R}^{n \times n}$ satisfy $A > 0$ and $n > 0$. Then:

(a) We have $\rho(A) > 0$.

(b) The number $\rho(A)$ is an eigenvalue of A and has algebraic multiplicity 1 (and therefore geometric multiplicity 1 as well).

(c) There is a unique $\rho(A)$-eigenvector $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$ of A with $x_1 + x_2 + \cdots + x_n = 1$. This eigenvector x is furthermore positive. (It is called the *Perron vector* of A.)

(d) There is a unique vector $y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n$ such that $y^T A = \rho(A)\, y^T$ and $x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = 1$. This vector y is also positive.

(e) We have

$$\left(\frac{1}{\rho(A)} A\right)^m \to xy^T \qquad \text{as } m \to \infty.$$

(f) The only eigenvalue of A that has absolute value $\rho(A)$ is $\rho(A)$ itself.