Matrix Analysis

Yunwei Ren

Contents

1 Eigenvalues, eigenvectors, and similarity

1.0 Introduction

1.

Proof. Let $S = \{x \in \mathbb{R}^n : x^T x = 1\}$, which is clearly a compact subset of $\mathbb{R}^n$. Consider the function $f : x \mapsto x^T Ax$. Since,

$$\|f(x + \delta) - f(x)\| = \|(x^T A)\delta + \delta^T (Ax) + \delta^T A\delta\| \leq K\|\delta\|$$

for every $x \in \mathbb{R}$ and some fixed K, f is continuous. Hence, by Weierstrass's theorem, f attains its maximum value at some point $x \in S$. Namely, (1.0.3) has a solution x. Therefore, there exists some $\lambda \in \mathbb{R}$ such that $2(Ax - \lambda x) = 0$, implying that every real symmetric matrix has at least one real eigenvalue. $\qquad\square$

2.

Proof. Let $S = \{x \in \mathbb{R}^n : x^T x = 1\}$ and m be the maximum value of $x \mapsto x^T Ax$ in S. Suppose λ is an eigenvalue of A and $u \neq 0$ is its associated eigenvector, then

$$Au = \lambda u \quad \Rightarrow \quad u^T Au = \lambda \|u\|^2 \quad \Rightarrow \quad (u/\|u\|)^T A(u/\|u\|) = \lambda \quad \Rightarrow \quad m \geq \lambda.$$

Meanwhile, by the previous discussion, m itself is a eigenvalue of A. Hence, it is the largest real eigenvalue of A. $\qquad\square$

1.1 The eigenvalue-eigenvector equation

1.

Proof. It follows from

$$(A^{-1} - \lambda^{-1}I)x = (A^{-1} - \lambda^{-1}A^{-1}A)x = \lambda^{-1}A^{-1}(\lambda I - A)x = 0.$$

$\qquad\square$

3.

Proof. Since $A \in M_n(\mathbb{R})$, $u, v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$,

$$Ax = \lambda x \quad \Rightarrow \quad Au + iAv = \lambda u + i\lambda v$$

implies $Au = \lambda u$ and $Av = \lambda v$. As $x \neq 0$, at least one of u and v is nonzero and therefore A has a real eigenvector associated with λ. It can happen that only one of u and v is an eigenvector of A, because if $x \in \mathbb{R}^n$, which may happen as we discussed above, the imaginary part of x is 0. Finally, if x is a real eigenvector of A, then the eigenvalue λ it associated with must be real. Otherwise, at least one entry of λx is not real as $x \neq 0$, contradicting with the fact that Ax is real. $\qquad\square$

5.

Proof. Let $p(t) = t^2 - t$. Since A is idempotent, $p(A) = A^2 - A = 0$. Hence, 0 is the only eigenvalue of $p(A)$. By Theorem 1.1.6, the only values the eigenvalues of A can be are the zeros of p, namely, 0 and 1.

Suppose A is nonsingular, then multiplying A^{-1} on the both sides of $A^2 = A$ yields $A = I$. $\qquad\square$

7.

Proof. Suppose $\lambda \in \sigma(A)$ and x is its associated eigenvector, then
$$0 = (A - \lambda I)x = x^*(A^* - \bar{\lambda}I) = x^*(A - \bar{\lambda}I)$$
$$\Rightarrow \quad 0 = x^*(A - \bar{\lambda}I)x = x^*Ax - \bar{\lambda}x^*x = (\lambda - \bar{\lambda})\|x\|^2.$$

Hence, $\lambda = \bar{\lambda}$, implying all eigenvalues of A are real. $\qquad\square$

9.

Solution. Solve the equation $\det(A - \lambda I) = 0$ and we get $\lambda = \pm i$. $\qquad\square$

11.

Proof. If $\operatorname{rank}(A - \lambda I) < n - 1$, then $\operatorname{adj}(A - \lambda I) = 0$ by (0.8.2) and therefore we can always choose y to be the 0 and the other parts of the proposition clearly hold. Hence, in the following discussion, we assume that $\operatorname{rank}(A - \lambda I) = n - 1$.

Apply the full-rank factorization and we get $\operatorname{adj}(A - \lambda I) = \alpha x y^*$ for some nonzero $\alpha \in \mathbb{C}$ and $x, y \in \mathbb{C}^n$. Replacing x with αx and α with 1 proves the first part.

Suppose $\operatorname{adj}(A - \lambda I) = [\beta_1, \ldots, \beta_n]$, then
$$(A - \lambda I)\operatorname{adj}(A - \lambda I) \quad \Rightarrow \quad (A - \lambda I)\beta_k = 0 \quad (k = 1, 2, \ldots, n),$$

implying that β_k is an eigenvector of A associated with λ as long as it is nonzero. $\qquad\square$

13.

Proof. If $\operatorname{rank} A < n - 1$, then x is always an eigenvector of $\operatorname{adj} A$ associated with 0 as $\operatorname{adj} A = 0$. Hence, we may assume that $\operatorname{rank} = n - 1$. Then $\operatorname{adj} A = (\det A)A^{-1}$. By Exercise 1, x is an eigenvector of A^{-1} and therefore an eigenvector of $\operatorname{adj} A$. $\qquad\square$

1.2 The characteristic polynomial and algebraic multiplicity

2.

Proof. Suppose $A = [a_{ij}]_{m,n} = [\alpha_1, \ldots, \alpha_n]^T$ and $B = [b_{ij}]_{n,m} = [\beta_1, \ldots, \beta_n]$, then

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} \alpha_i \beta_i = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}b_{ji} = \sum_{j=1}^{m}\sum_{i=1}^{n} b_{ji}a_{ij} = \operatorname{tr}(BA).$$

Hence, for nonsingular $S \in M_n$, $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(S(S^{-1}A)) = \operatorname{tr}(A)$.

For $A \in M_n$, $\det(S^{-1}AS) = \det(S)\det(S^{-1})\det(A) = \det(A)$, which means the determinant function on M_n is similarity invariant. $\qquad\square$

4.

Proof. It follows immediately from the fact that $\sigma(A) \subset \{0, 1\}$ and $S_k(A)$ is the sum of some $\prod \lambda_{i_j}$. $\qquad\square$

6.

Proof. $\operatorname{rank}(A - \lambda I) = n - 1$ implies the matrix $A - \lambda I$ is singular, and therefore λ is an eigenvalue of A. However, it may not have multiplicity 1. For example[1], suppose

---

[1]Thanks to Zhihan Jin, one of my classmates.

$A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$. rank $A = 1$ but 0, the only eigenvalue of A is of multiplicity 2. $\qquad\square$

8.

Proof. $p_{A+\lambda I}(t) = \det(tI - (A + \lambda I)) = \det((t - \lambda)I - A) = p_A(t - \lambda)$ and hence the eigenvalues of $A + \lambda I$, the zeros of $p_{A+\lambda}(t)$, are $\lambda_1 + \lambda, \ldots, \lambda_n + \lambda$. $\qquad\square$

10.

Proof. Since $p_A(t)$ has n roots and non-real roots of a polynomial come in paris, at least one of the roots is real. Hence, A has at least one real eigenvalue. $\qquad\square$

12. TODO

14.

Proof. Suppose $C = \begin{bmatrix} \mu & 0 \\ * & B \end{bmatrix}$. By the exercise on p52,

$$p_A(t) = (t - \lambda)p_C(t) = (t - \lambda)p_{C^T}(t) = (t - \lambda)(t - \mu)p_B(t).$$

$\qquad\square$

16.

Proof. $f(t) = \det(A + (tx)y^T = \det A + y^T(\operatorname{adj} A)tx = \det A + t\beta$ where $\beta = y^T(\operatorname{adj} A)x$, a constant independent of t. Hence, for $t_1 \neq t_2$

$$\frac{t_2 f(t_1) - t_1 f(t_2)}{t_2 - t_1} = \frac{t_2(\det A + t_1\beta) - t_1(\det A + t_2\beta)}{t_2 - t_1} = \det A.$$

For the second part, we can get from calculation that

$$f(-b) = \det(A - b[1, \ldots, 1]^T[1, \ldots, 1]) = (d_1 - b) \cdots (d_n - b) = q(b)$$

and $f(-c) = q(-c)$. Hence, if $b \neq c$,

$$\det A = \frac{(-c)f(-b) - (-b)f(-c)}{(-c) - (-b)} = \frac{bq(c) - cq(b)}{b - c}.$$

Now suppose $b = c$. Note that $f(t)$ is a linear function of t, which is differentiable, implying that

$$\det A = \lim_{t_2 \to t_1} \frac{t_2 f(t_1) - t_1 f(t_2)}{t_2 - t_1} = f'(t_1)t_1 - f(t_1).$$

Meanwhile, since $q(t)$ is continuous, $q(t) \to f(-b)$ as $t \to b$. Thus,

$$\det A = \lim_{c \to b} \frac{(-c)f(-b) - (-b)f(-c)}{(-c) - (-b)} = q(b) - bq'(b).$$

Let

$$A_* = \lambda I - A = \begin{bmatrix} \lambda & -b & \cdots & -b \\ -c & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & -b \\ -c & \cdots & -c & \lambda \end{bmatrix}.$$

4

and $q_*(t) = (\lambda - t)^n$, then by the previous result,

$$p_A(\lambda) = \frac{-bq_*(-c) - (-c)q_*(-b)}{-c - (-b)} = \frac{b(\lambda + c)^n - c(\lambda + b)^n}{b - c}, \qquad \text{if } b \neq c,$$

$$p_A(\lambda) = q_*(-b) - (-b)q_*'(-b) = (\lambda + b)^{n-1}(\lambda - (n-1)b), \qquad \text{if } b = c.$$

$\square$

18.

Proof. The identity can be derived immediately from Observation 1.2.4 and the identity $a_1 = (-1)^{n-1} \operatorname{tr} \operatorname{adj}(A)$, the proof of which can be found on p53. $\square$

20.

Proof. By (1.2.13),

$$\det(I + A) = (-1)^n p_A(-1) = (-1)^n \left((-1)^n + \sum_{k=1}^n (-1)^{n-k} E_k(A)(-1)^k \right) = 1 + \sum_{k=1}^n E_k(A).$$

$\square$

22.

Proof. Suppose

$$A = \begin{bmatrix} t & -1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & t \end{bmatrix},$$

then

$$\begin{aligned} p_{C_n(\varepsilon)}(t) &= \det(A + [0, \ldots, 0, 1]^T[-\varepsilon, 0, \ldots, 0]) \\ &= \det A - \varepsilon[1, 0, \ldots, 0](\operatorname{adj} A)[0, \ldots, 0, 1]^T \\ &= \det A - \varepsilon((\operatorname{adj} A)[1, n]) \\ &= \det A - \varepsilon \det A[\{n\}^c, \{1\}^c] \\ &= t^n - \varepsilon. \end{aligned}$$

And its spectrum, namely the set of roots of $p_{C_n(\varepsilon)}$, is $\{\varepsilon^{1/n} e^{2\pi ik/n} : k = 0, 1, \ldots, n-1\}$. Hence,

$$\rho(I + C_n(\varepsilon)) = 1 + \rho(C_n(\varepsilon)) = 1 + \varepsilon^{1/n}.$$

$\square$

1.3 Similarity

1.

Proof. (a) Since A and B are diagonalizable and commute, by Theorem 1.3.21, they are simultaneously diagonalizable. Hence, there exists some nonsingular $S \in M_n$ such that

$$A + B = S^{-1} \operatorname{diag}(\lambda_1, \ldots, \lambda_n)S + S^{-1} \operatorname{diag}(\mu_{i_1}, \ldots, \mu_{i+n})S$$
$$= S^{-1} \operatorname{diag}(\lambda_1 + \mu_{i_1}, \ldots, \lambda_n + \mu_{i_n})S.$$

Therefore, $\sigma(A + B) = \{\lambda_1 + \mu_{i_1}, \ldots, \lambda_n + \mu_{i_n}\}$.
(b) By Exercise 1.1.6, $\sigma(B) = \{0\}$, completing the proof.
(c) $\sigma(AB) = \{\lambda_1 \mu_{i_1}, \ldots, \lambda_n \mu_{i_n}\}$, because

$$S^{-1}(AB)S = (S^{-1}AS)(S^{-1}BS) = \operatorname{diag}(\lambda_1 \mu_{i_1}, \ldots, \lambda_n \mu_{i_n}).$$

$\square$

3.

Proof.

$$\sum_{k=0}^{n} a_k A^k = \sum)k = 0^n S^{-1}(a_k \Lambda^k)S = S^{-1}p(\Lambda)S.$$

$\square$

5.

Proof. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = I$. Clear that A and B commute but are not simultaneously diagonalizable since A can not be diagonalized. This does not violate 1.3.12 because in 1.3.12, A and B are required to be diagonalizable. $\square$

7.

Proof. Suppose $B = \operatorname{diag}(b_1, \ldots, b_n)$, then clear that $A = \operatorname{diag}(\sqrt{b_1}, \ldots, \sqrt{b_n})$ is a square root of B.

Assume that such A exists, then by Theorem 1.1.6, $\sigma(A) = \{0\}$, implying that

$$A = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}.$$

Therefore,

$$A^2 = \begin{bmatrix} ab & 0 \\ 0 & ab \end{bmatrix} \neq B.$$

Contradiction. $\square$

9.

Proof. By Theorem 1.3.22, AB and BA have the same eigenvalues. And as similar matrices are of the same rank, AB and BA are not similar. $\square$

11.

Proof. Suppose that $A, B \in M_n$ commute and $Ax = \lambda x$ where $x \neq 0$ and k the is the smallest integer such that $B^k x \in \text{span}\{x, Bx, \dots, B^{k-1}x\} = \mathcal{S}$. For every $u = \sum_{i=0}^{k-1} x_i B^i x \in \mathcal{S}$, Bu is a linear combination of $\sum_{i=0}^{k-2} x_i B^{i+1} x$ and $x_{k-1} B^k x \in \mathcal{S}$. Hence $Bu \in \mathcal{S}$ and therefore $\mathcal{S}$ is B-invariant. By Observation 1.3.18, there exists some $0 \neq y \in \mathcal{S}$ which is an eigenvector of B. Meanwhile, since

$$A \sum_{i=0}^{k-1} x_i B^i x = \sum_{i=0}^{k-1} x_i (AB^i) x = \sum_{i=0}^{k-1} x_i B^i \lambda x = \lambda \sum_{i=0}^{k-1} x_i B^i x,$$

every nonzero vector in $\mathcal{S}$ is a eigenvector of A and so does y. Hence, A and B have a common eigenvector y.

Now we argue by induction on m, the size of the finite commuting family $\mathcal{F} = \{A_1, \dots, A_m\}$. Suppose that $y \neq 0$ is a common eigenvector of $A_1, \dots, A_{m-1}$ and let k be the smallest integer such that $A_m^k y \in \text{span}\{y, A_m y, \dots, A_m^{k-1} y\} = \mathcal{S}$. Then by some argument similar to the previous one, $\mathcal{S}$ is A_m-invariant and hence contains a eigenvector z of A_m. Meanwhile, since A_i and A_m commute, every nonzero vector in $\mathcal{S}$ is an eigenvector of A_i for $i = 1, \dots, m-1$ and so does z, concluding that matrices in a finite commuting families share a common eigenvector.

M_n is linear space of dimension n^2 and $\mathcal{F}$ is a subspace of M_n since for any $A, B, C \in \mathcal{F}$ and $a, b \in \mathbb{C}$,

$$(aA + bB)C = a(AC) + b(BC) = a(CA) + b(CB) = C(aA + bB).$$

Let $\mathcal{B} = \{B_1, \dots, B_k\}$ be a basis of $\mathcal{F}$. Since $\mathcal{B}$ is finite and commuting, (b) shows that the matrices in $\mathcal{B}$ have a common eigenvector x. Hence, supposing $B_i x = \lambda_i x$,

$$\left(\sum_{i=1}^{k} b_i B_i \right) x = \sum_{i=1}^{k} b_i (B_i x) = \left(\sum_{i=1}^{k} b_i \lambda_i \right) x$$

where b_i are some scalars. Thus, x is a eigenvector of every $A \in \mathcal{F}$. $\qquad\square$

13.

Proof. Since similar diagonalizable matrices have the same eigenvalues and multiplicities, their characteristic polynomials are therefore the same and vice versa.

However, this is not true for two matrices which are not both diagonalizable. For example, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $0_{2 \times 2}$ have the same characteristic polynomial but they are not similar. $\qquad\square$

15.

Proof. Suppose that $A = S \Lambda S^{-1}$ where Λ is a diagonal matrix, then

$$p(A) = \sum_{k=0}^{n} a_k (S \Lambda S^{-1})^k = \sum_{k=0}^{n} a_k S \Lambda^k S^{-1} = S p(\Lambda) S^{-1}.$$

Clear that $p(\Lambda)$ is again a diagonal matrix. Hence, $p(A)$ is also diagonalizable.

However the converse is not true. For example, $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not diagonalizable but $A^2 = 0$ itself is a diagonal matrix. $\qquad\square$

17.

Proof. Suppose that $A = TBT^{-1}$ where $T \in M_n(\mathbb{R})$ is nonsingular, then $\bar{A} = \overline{TBT^{-1}} = \bar{T}\bar{B}\bar{T}^{-1} = T\bar{B}T^{-1}$ since T is real. And the converse is obviously true. $\square$

19.

Proof. Clear that $Q = Q^T$ and $Q^2 = I$, implying $Q = Q^{-1}$.
(a) Suppose that $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$.

$$0 = K_{2n}A - AK_{2n} = \begin{bmatrix} A_{21} - A_{12} & A_{22} - A_{11} \\ A_{11} - A_{22} & A_{12} - A_{21} \end{bmatrix}.$$

Hence, A is 2-by-2 block centrosymmetric. And the proof of the converse is trivial. If A is nonsingular, then we have $A^{-1}K_{2n} = K_{2n}A^{-1}$, which implies A^{-1} is 2-by-2 block centrosymmetric. Meanwhile, since $K_{2n}^{-1} = K_{2n}$, $K_{2n}AK_{2n} = A$. Suppose B is a 2-by-2 block centrosymmetric matrix, then

$$K_{2n}AB = K_{2n}A(K_{2n}BK_{2n}) = (K_{2n}AK_{2n})B_{2n}K_{2n} = ABK_{2n}.$$

Therefore, AB is a 2-by-2 block centrosymmetric matrix as well.
(b)

$$\begin{aligned} Q^{-1}AQ &= \frac{1}{2}\begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix}\begin{bmatrix} B & C \\ C & B \end{bmatrix}\begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix} \\ &= \begin{bmatrix} B+C & 0 \\ 0 & B+C \end{bmatrix} = (B+C) \oplus (B-C). \end{aligned}$$

(c)

$$\det A = \det(Q^{-1}AQ) = \det\begin{bmatrix} B+C & 0 \\ 0 & B+C \end{bmatrix} = \det(B^2 + CB - BC - C^2)$$

and $\operatorname{rank} A = \operatorname{rank}(B+C) + \operatorname{rank}(B-C)$ follows immediately from $Q^{-1}AQ = (B+C) \oplus (B-C)$.
(d) $Q^{-1}\begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix} = C \oplus (-C)$. Since $p_{C\oplus(-C)}(t) = p_C(t)p_{-C}(t)$, the eigenvalues occur in $\pm$ pairs. $\square$

21. TODO

23.

Proof. Suppose that there exists some $X \in M_{n,m}$ such that $C = BX$ and let $S = \begin{bmatrix} I_n & X \\ 0 & I_m \end{bmatrix}$. Clear that S is nonsingular and $S^{-1} = \begin{bmatrix} I_n & -X \\ 0 & I_m \end{bmatrix}$. Since

$$S\begin{bmatrix} B & BX \\ 0_n & 0_m \end{bmatrix}S^{-1} = \begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix},$$

A is similar to $B \oplus 0_m$.

Now we suppose that A is similar to $B \oplus 0_m$. Since similar matrices have the same rank, $\operatorname{rank}[B\,C] = \operatorname{rank} B$. $\square$

25.

Proof. TODO $\qquad\square$

29.

Proof. Since $\det A = \sum_\sigma (\operatorname{sgn}\sigma \prod_{i=1}^n a_{i\sigma(i)})$, where σ is any permutation of $\{1, 2, \ldots, n\}$, the determinant of a matrix whose entries are integers is an integer.

Suppose that the a_{ij} is changed from -1 to 1 and denote the new matrix by $\tilde{A}$. Let $x, y \in \mathbb{C}^n$ be two vectors such that $x_i = 1$ and $y_j = 2$, then $\tilde{A} = A + xy^T$. By Cauchy's identity,
$$\det \tilde{A} = \det A + y^T(\operatorname{adj} A)x = \det A + 2\det A[\{i\}^c, \{j\}^c].$$

Hence, the parity of $\det A$ is unchanged.

[2] Since changing a -1 entry to 1 does not change the parity of the determinant, we can change all the entries to 1. Hence, the parity of $\det A$ is the same as the parity of $\det(J_n - I)$. Induction on n yields that it is opposite to the parity of n. Thus, if n is even, then $\det A$ is odd and therefore nonzero, implying A is nonsingular. $\qquad\square$

31.

Proof. The characteristic polynomial of the matrix is $p(t) = (t-a)^2 + b^2 = (t-a-ib)(t-a+ib)$. Hence its eigenvalues are $a \pm ib$. $\qquad\square$

33.

Proof.
(a) Since A is real, $A\bar{x} = \overline{\bar{A}x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}$.
(b) Since λ is not real, x and $\bar{x}$ are associated with different eigenvalues. Hence, x and $\bar{x}$ are linear independent. Suppose that $mu + nv = 0$. Then
$$0 = m(x + \bar{x}) - in(x - \bar{x}) = (m - in)x + (m + in)\bar{x}.$$

Since $m - in = 0$ and $m + in = 0$, $m = n = 0$. Thus, u and v are also linear independent.
(c)
$$Au = \frac{1}{2}A(x + \bar{x}) = \frac{1}{2}(\lambda x + \bar{\lambda}\bar{x}) = \frac{1}{2}[(a + ib)(u + iv) + (a - bi)(u - iv)] = au - bv.$$

Similarly, $Av = bu + av$. Hence,
$$A[u, v] = [Au, Av] = [au - bv, bu + av] = [u, v]B.$$

(d) Since $S\begin{bmatrix} I_2 \\ 0 \end{bmatrix} = [u\,v]$, $S^{-1}[u\,v] = \begin{bmatrix} I_2 \\ 0 \end{bmatrix}$ and the proof of the next result is trivial.
(e) Since $p_A(t) = p_B(t)p_{A_1}(t)$, the result amounts to the fact that λ and $\bar{\lambda}$ are two roots of $p_B(t)$. $\qquad\square$

---

[2]I don't know what J_n actually is here and assume it to be the matrix whose entries are all 1.