Matrix Analysis

Yunwei Ren

Contents

1 Eigenvalues, eigenvectors, and similarity **2**

1.0 Introduction 2

1.1 The eigenvalue-eigenvector equation 2

1.2 The characteristic polynomial and algebraic multiplicity 3

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,

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

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. $\square$

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

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

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

1.1 The eigenvalue-eigenvector equation

1.

Proof. It follows from

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

$\square$

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. $\square$

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 $A = I$. $\square$

7.

Proof. Suppose $\lambda \in \sigma(A)$ and x is its associated eigenvector, then

$$\begin{aligned} & 0 = (A - \lambda I)x = x^*(A^* - \bar{\lambda} I) = x^*(A - \bar{\lambda} I) \\ \Rightarrow\quad & 0 = x^*(A - \bar{\lambda} I)x = x^*Ax - \bar{\lambda} x^*x = (\lambda - \bar{\lambda})\|x\|^2. \end{aligned}$$

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

9.

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

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. □

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 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 $\operatorname{adj} A$. □

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. □

4.

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

6.

Proof. $\text{rank}(A - \lambda I) = n - 1$ implies the matrix $A - \lambda I$ is singular, and therefore λ is an eigenvalue of A. TODO $\square$

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 $\lambda_1 + \lambda, \ldots, \lambda_n + \lambda$. $\square$

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. $\square$

12. TODO

14.

Proof. Suppose $C = \left[\begin{smallmatrix} \mu & 0 \\ * & B \end{smallmatrix}\right]$. By the exercise on p52,

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

$\square$

16.

Proof. $f(t) = \det(A + (tx)y^T = \det A + y^T(\text{adj}\, A)tx = \det A + t\beta$ where $\beta = y^T(\text{adj}\, A)x$, 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 liner 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

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

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

$$\begin{aligned} p_A(\lambda) &= \frac{-bq_*(-c) - (-c)q_*(-b)}{-c-(-b)} = \frac{b(\lambda+c)^n - c(\lambda+b)^n}{b-c}, && \text{if } b \neq c, \\ p_A(\lambda) &= q_*(-b) - (-b)q_*'(-b) = (\lambda+b)^{n-1}(\lambda - (n-1)b), && \text{if } b = c. \end{aligned}$$

□

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. □

20.

Proof. By (1.2.13),

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

□

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 $\{\varepsilon^{1/n} e^{2\pi i k/n} : k = 0, 1, \ldots, n-1\}$. Hence,

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

□