Eigenvalue, Rank and Diagonalizability

——-revision and complement to lecture

Chengjian Yao, ShanghaiTech University

December 1, 2018

Theorem 1. *Let A be a $m \times n$ and B be a $n \times m$ matrix. Then*

a) The number of nonzero eigenvalues of AB and BA are the same. In particular, their traces are the same.

b) The matrix $I_m - AB$ is invertible if and only if $I_n - BA$ is invertible. Moreover, $(I_n - BA)^{-1} = I_n + B(I_m - AB)^{-1}A$

c) The equality

$$rank(I_m - AB) - rank(I_n - BA) = m - n$$

holds always.

Proof. The block multiplication for matrices gives

$$\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \tag{1}$$

and

$$\begin{bmatrix} I_m & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} I_m - AB & 0 \\ -B & I_n \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ -B & I_n - BA \end{bmatrix} \tag{2}$$

Equation (1) tells that the characteristic polynomials are the same:

$$\left| \lambda I_{m+n} - \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \right| = \left| \lambda I_{m+n} - \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \right| \tag{3}$$

i.e.

$$\lambda^n \cdot p_{AB}(\lambda) = \lambda^m \cdot p_{BA}(\lambda)$$

Let $0, \lambda_1, \cdots, \lambda_k$ be the distinct eigenvalues of AB with multiplicities $m_0, m_1, \cdots, m_k$ respectively, and $0, \lambda_1', \cdots, \lambda_l'$ be the distinct eigenvalues of BA with multiplicities $m_0', m_1', \cdots, m_l'$ respectively. Then

$$\lambda^n \cdot \lambda^{m_0}(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k} = \lambda^m \cdot \lambda^{m_0'}(\lambda - \lambda_1')^{m_1'} \cdots (\lambda - \lambda_l')^{m_l'}$$

which implies

$$\begin{aligned} n + m_0 &= m + m_0' \\ m_1 + \cdots + m_k &= m_1' + \cdots + m_l'. \end{aligned}$$

And *trace* is the sum of all eigenvalues, therefore part a) is done.

Equation (2) tells that $I_m - AB$ is invertible iff $I_n - BA$ is invertible. Taking the inverse of both sides, we get

$$\begin{bmatrix} I_m & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} (I_m - AB)^{-1} & 0 \\ B(I_m - AB)^{-1} & I_n \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ (I_n - BA)^{-1}B & (I_n - BA)^{-1} \end{bmatrix}$$

Expanding out the block matrix multiplications on the LHS (left hand side), we get part b) of the theorem.

To prove part c), we need to make one observation:

$$rank \begin{bmatrix} A_1 & \mathbf{0} \\ A_2 & A_3 \end{bmatrix} = rank(A_1) + rank(A_3)$$

To validate this observation, let us assume A_1, A_2, A_3 have sizes $m \times k, n \times k, n \times l$ respectively, and write $\mathbb{R}^{m+n} = \mathbb{R}^m \oplus \mathbb{R}^n$ and any column vector $\boldsymbol{v} \in \mathbb{R}^{m+n}$ as

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$$

where $\boldsymbol{x} \in \mathbb{R}^m$, $\boldsymbol{y} \in \mathbb{R}^n$. It is easily seen that

$$\begin{bmatrix} A_1 & \mathbf{0} \\ A_2 & A_3 \end{bmatrix} \boldsymbol{v} = \mathbf{0}_{m+n}$$

is equivalent to

$$A_1 \boldsymbol{x} = \mathbf{0}_m, \quad \text{and} \quad A_3 \boldsymbol{y} = \mathbf{0}_n$$

It follows

$$N(\begin{bmatrix} A_1 & \mathbf{0} \\ A_2 & A_3 \end{bmatrix}) = N(A_1) \oplus N(A_3)$$

By *The Fundamental Theorem of Linear Algebra*,

$$\begin{aligned} rank(\begin{bmatrix} A_1 & \mathbf{0} \\ A_2 & A_3 \end{bmatrix}) &= m + n - dim\ N(\begin{bmatrix} A_1 & \mathbf{0} \\ A_2 & A_3 \end{bmatrix}) \\ &= m + n - (dim\ N(A_1) + dim\ N(A_3)) \\ &= rank(A_1) + rank(A_3) \end{aligned}$$

Applying the above observation to Equation (2) and using the fact that rank of similar matrices are the same, we know

$$rank(I - AB) + n = m + rank(I - BA)$$

which shows c). $\square$

Remark 2. For $n \neq m$, $\det(AB)$ and $\det(BA)$ cannot be both nonzero since either AB or BA will not have full rank. The reason is that $rank(AB) \leq \min\{rank(A), rank(B)\}$ and $rank(BA) \leq \min\{rank(A), rank(B)\}$.

Proposition 3. *a) If $A = A^T, B = B^T$, then $rank(AB) = rank(BA)$.*

b) If $A = A^T, B = B^T$ and AB is diagonalizable, then BA is also diagonalizable.

Proof. The rank of any matrix equals to the rank of its transpose (part of *The Fundamental Theorem of Linear Algebra*), thus $rank(AB) = rank((AB)^T) = rank(B^T A^T) = rank(BA)$.

To prove part b), let $AB = P\Lambda P^{-1}$ for some invertible matrix P, then

$$BA = B^T A^T = (AB)^T = (P\Lambda P^{-1})^T = (P^T)^{-1}\Lambda P^T$$

thus diagonalizable. $\square$

Remark 4. a) In general, $rank(AB) \neq rank(BA)$. For instance,

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

which gives nonzero rank one AB but zero BA. This counterexample also shows that assuming one of A and B symmetric is not enough to guarantee the rank equality stated in the proposition.

b) If we only assume A, B are symmetric, we could not conclude AB is diagonalizable. For instance,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

only has one eigenvector $[1\ 0]^T$ and thus not diagonalizable. Of course, BA is not diagonalizable either.

c) Without assuming A and B symmetric, AB diagonalizable is not necessarily guaranteeing BA diagonalizable. The same A, B as in part a) of this remark gives the counterexample. In this counterexample AB is not diagonalizable while zero matrix BA is of course diagonalizable.

Theorem 5. *For any diagonalizable matrix A, the number of nonzero eigenvalues of A is equal to $rank(A)$. In particular, the statement is true for real symmetric matrix.*

Proof. Let $\lambda_1, \cdots, \lambda_k$ be the nonzero eigenvalues of A, and 0 be a multiplicity $n-k$ eigenvalue. Assuming A is diagonalizable, then

$$A = P\Lambda P^{-1} = P \operatorname{diag}(\lambda_1, \cdots, \lambda_k, 0, \cdots, 0)\ P^{-1}$$

for some invertible matrix P. We've learned that multiplying by invertible matrices on the left and on the right neither change the rank of a matrix. Therefore,

$$rank(A) = \operatorname{rank}(\operatorname{diag}(\lambda_1, \cdots, \lambda_k, 0, \cdots, 0)) = k$$

The statement is true for real symmetric matrix because we've proved real symmetric matrix is always diagonalizable. □

Remark 6. For general square matrix, the number of nonzero eigenvalues is not necessarily equal to the rank. Consider the example

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

which has no nonzero eigenvalues but the rank is 2. It is actually true that the number of nonzero eigenvalues must be no bigger than the rank of the matrix. We will not prove it here since it must involve a deep theorem in linear algebra, *Jordan Decomposition Theorem*.