Sampad B Mohanty —————————————— November 4, 2021

1.

a) First we are going to prove the forward direction, i.e similar matrices have the same characteristics and minimal polynomials.

Let A and B are similar, i.e $A \sim B$, then $A = PBP^{-1}$ for some invertible matrix P. Let the characteristic polynomials of A and B be p_A and p_B.

Theorem 1: Similar matrices have the same characteristics polynomial.

Proof -

$$\begin{aligned} p_A(x) = \det(A - xI) &= \det(PBP^{-1} - xI) \\ &= \det(PBP^{-1} - xPIP^{-1}) \quad \because I\ commutes \\ = \det(P\ (B - xI)\ P) = \ det(P)\ &\det(B - xI)\det(P^{-1}) \\ &= \det(B - xI) \\ &= p_B(x) \quad \blacksquare \end{aligned}$$

Since characteristic polynomials for similar matrices are the same, they have the same eigenvalues.

Lemma 1: For any polynomial $f(x)$, $S \sim T \implies f(S) \sim f(T)$. In other words, if $f(x)$ be any polynomial and S and T be similar matrices, then, $f(STS^{-1}) = Sf(T)S^{-1}$

Proof -

$$\begin{aligned} Let,\ f(x) &= a_0 + a_1x + a_2x^2 + ... \\ So,\ f(STS^{-1}) &= a_0 + a_1STS^{-1} + a_2(STS^{-1})^2 + ... \\ &= a_0 + a_1STS^{-1} + a_2STS^{-1}STS^{-1} + ... \\ &= a_0 + a_1STS^{-1} + a_2ST^2S^{-1} + ... \\ &= S(a_0 + a_1T + a_2T^2 + ...)S^{-1} \\ &= Sf(T)S^{-1} \quad \blacksquare \end{aligned}$$

Lemma 2: For similar matrices, not only is the characteristics polynomial the same, but all annihilating polynomials coincide. More explicitly, if $A = PBP^{-1}$ and $q_A(x)$ is an annihilating polynomial of A, then it is an annihilating polynomial of B too. Similarly, any annihilating polynomial q_B of B also annihilates A.

Proof -

$$\begin{aligned} q_A(A) &= 0 \\ \implies q_A(PBP^{-1}) &= 0 \\ \implies Pq_A(B)P^{-1} = 0 \quad &\because \textit{Lemma 1} \\ \implies q_A(B) = 0 \quad \because \det(P) \neq 0 \quad &\blacksquare \end{aligned}$$

Let the minimal polynomials for A and B be m_A and m_B. By definition, minimal polynomials are annihilating polynomials of the learst degree. Hence $m_A(A) = 0$ and $m_B(B) = 0$.

By Lemma 2, $m_A(B) = 0$ and $m_B(A) = 0$. What does this tell us? Well, atleast from the surface, we can see that the minimal polynomials of A and B must be of the same degree.

Proof (by contradiction) -

Without loss of generality (WLOG), assume and $degree(m_A) < degree(m_B)$ and hence m_A and m_B are distinct.

By Lemma 2, $m_A(B) = 0$; i.e m_A not only annihilates A but also annihilates B, we cannot have m_B as the minimal polynomial of B since m_A has lesser degree and is also an annihilator of B.

Hence, $degree(m_A) = degree(m_B)$.

But what more can we say? On more close inspection, we can actually prove that the two are the same polynomial. We will prove this in *Theorem 2* (by contradiction again) but we need the following lemma first.

<u>Lemma 3</u>: The minimal polynomial $m_T(x)$ of any matrix T divides all annihilating polynomials $q_T(x)$ of T.

Since, $m_T(x)$ is an annihilating polynomial of T with the least degree, $degree(m_T) \leq degree(q_T)$. Note that we haven't proved yet that minimal polynomial (m_T here) of any matrix (T here) is unique.

Proof-

From polynomial algebra, we know that we can divide q_T by m_T to get a quotient n(x) and a remainder r(x) i.e $q_T(x) = m_T(x) \cdot n(x) + r(x)$ such that $degree(r) < degree(m_T)$.

Now,

$$\begin{aligned} q_T(T) &= 0 \\ \implies m_T(T)\cdot n(T) + r(T) &= 0 \\ \implies 0\cdot n(T) + r(T) = 0 \quad &\because m_T(T) = 0 \\ \implies r(T) &= 0 \end{aligned}$$

Since $r(T) = 0$ and $degree(r) < degree(m_T)$, $m_T(x)$ cannot be the minimal polynomial as assumed since $r(x)$ also annihilates T.

Theorem 2: Similar matrices have the same minimal polynomial. In notations, $m_A = m_B$ when A and B are similar, i.e when $A = PBP^{-1}$ for some P with $\det(P) \neq 0$.

Proof- Assume that m_A and m_B are distinct.

By lemma 2, m_B annihilates A and by lemma 3 m_A divides m_B. By the same argument m_B divides m_A.

Hence, $m_A|m_B$ and $m_B|m_A \iff m_A|m_B|m_A. \iff m_A = m_B$ ■

Lemma 4: Minimal polynomial $m_T(x)$ of any matrix T is unique.

Proof- Let there be two distinct minimal polynomials ${}_1m_T(x)$ and ${}_2m_T(x)$. Due to the definition of minimal polynomial being the annihilating polynomial of the least degree, we must have $degree({}_1m_T) = degree({}_2m_T)$ as otherwise only one of them will be called the minimal polynomial.

Since both are minimal, both must by definition also be annihilating polynomials. Hence by lemma 3,

${}_1m_T \mid {}_2m_T$ and ${}_2m_T \mid {}_1m_T \iff {}_1m_T \mid {}_2m_T \mid {}_1m_T \iff {}_1m_T = {}_2m_T$ ■

So far, we have been able to show that similar matrices have the same -

0. Annihilating Polynomials (lemma 2)
1. Characteristics polynomial (theorem 1)
2. Minimal polynomials (theorem 2)

Additionally, along the way, we also showed that -

0. $A \sim B \implies f(A) \sim f(B)$ for all polynomials $f(x)$ (lemma 1)
1. Minimal polynomial $m_T(x)$ of T divides all annihilating polynomials $q_T(x)$ of T (lemma 3)
2. Minimal polynomial $m_T(x)$ of a matrix T is unique (lemma 4)

So we have shown the forward direction for all matrices of any size. Now we have to show that for m=3, A and B (of shape $m \times m$) over a field F are similar if they have the same characteristics and minimal polynomial.

For m=3, the characteristics polynomial is of order 3 and hence has 3 roots (counting multiplicities) - $\lambda_1, \lambda_2, \lambda_3$.

The multiplicities of $\lambda_1, \lambda_2, \lambda_3$ can be represented as a 3-tuple and the possibilities are of the type (WLOG to rearranging the roots) - $(3,0,0), (2,1,0)$ and $(1,1,1)$.

The Jordan normal forms for these must have to match since the minimal and characteristics polynomial together define the Jordan normal form and there is no ambiguity for m=3 (or less).

b) However, for the m=4, the above does not hold and here is a counter-example -

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

Characteristics polynomial: $p_A(x) = p_B(x) = (x-1)^4$
Minimal Polynomial: $m_A(x) = m_B(x) = (x-1)^2$

Although the characteristics and the minimal polynomials are the same, $A \not\sim B$, i.e $\nexists P, det(P) \neq 0 \; s.t \; A = PBP^{-1}$.

<u>References</u> -

`https://www.statlect.com/matrix-algebra/minimal-polynomial`

`http://www.mathcounterexamples.net/two-non-similar-matrices-having-same-minimal-and-characteristic-polynomials/`

2. Given - $A^m \sim B^m$ for sufficiently large m.

   To Prove: $det(A) \neq 0 \implies A \sim B$

   and

   if A is non-invertible, A and B might not be similar.

   Proof -

   DO NOT KNOW HOW TO

   My guess is we will need to use Cayley-Hamilton Theorem somewhere to show that $A = f(A)$ for some polynomial $f(x)$ and the polynomial can be multiplied by a large enough power of A to make all the terms in the polynomial greater than m which is the power after with $A^m \sim B^m$.

$\det(A^m) = \det(A)^m \neq 0$ and hence A^m is invertible. Since $A^m \sim B^m$, hence $A^m = QB^mQ^{-1} \implies A^mQ = QB^m$.

$$A = f(A) = a_0 + a_1A + ...a_n - 1A^{n-1}$$
$$\implies AA^m = f(A)A^m = a_0A^m + a_1A^{m+1} + ...a_nA^{m+n-1}$$
$$(incomplete...)$$

When A is not invertible, we need not have $A \sim B$ even if $A^m \sim B^m$ for sufficiently large m. Here is an example - let A, B be 2×2 matrices. Let A represent a rotation by 90 degrees anti-clockwise and then a projection onto x-axis. Let B be the 0 matrix.

So,

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

Clearly, $A \nsim B$. However, as can be easily guessed, after first operation of A, we have x-axis as the range, and doing a second operation will kill the x-axis since it is also the kernel of A.

Computationally, $A^2 = AA = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Hence, $A^m = 0 \; \forall \; m \geq 2$. Similarly $B^m = 0 \; \forall \; m$.

Hence for sufficiently large m, $A^m \sim B^m$ although $A \nsim B$.

3. Given - A, B are $m \times m$ matrices over a field F.

a) If A is invertible, show that $AB \sim BA$.

Proof -

$$BA = IBA = A^{-1}ABA = A^{-1}(AB)A \quad - eq1$$

Similarly,

$$AB = ABI = ABAA^{-1} = A(BA)A^{-1} \quad - eq2$$

Hence $AB \sim BA$ $\blacksquare$

b)

c) To prove - AB and BA have the same characteristics polynomial i.e $p_{AB}(x) = p_{BA}(x)$.

Proof -

CASE 1: If A (or B) is invertible, then by part(a), $AB \sim BA$ and then by theorem-1, both have the same characteristics polynomial.

Otherwise, we give the following proof -

CASE 2: If neither of A and B is invertible (otherwise we would be in the above case), with elementary row and column operations, we can reduce A to the block matrix equivalence form

$$A_{eq} = \begin{bmatrix} I_{k\times k} & 0 \\ 0 & 0 \end{bmatrix}$$

where $k < m$ is the rank of the matrix A.

Let us put these elementary row and column operations into R and C matrices such that $R = \prod R_i$ and $C = \prod C_i$ where R_i and C_i are single row and column operations. Since all the row/column operations can be undone one by one, R and C are invertible. Hence,

$$A_{eq} = RAC$$

Now, we can also do the following notation jugglery,

$$B = R(R^{-1}BC^{-1})C = R\ B_{eq}\ C$$

where

$$B_{eq} = R^{-1}BC^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

for some block matrix with blocks $B_{11}, B_{12}, B_{21}, B22$ of shapes of B_{11} and B_{22} being $k \times k$ and $(m-k) \times (m-k)$ respectively.

Since matrix multiplication can also be done block-wise, we have

$$AB = \begin{bmatrix} I_{k\times k} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & 0 \end{bmatrix}$$
$$BA = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} I_{k\times k} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_{11} & 0 \\ B_{21} & 0 \end{bmatrix}$$

Now, clearly

$$AB - xI = \begin{bmatrix} B_{11} - xI_{k\times k} & B_{12} \\ 0 & -xI_{(m-k)\times(m-k)} \end{bmatrix}$$

$$BA - xI = \begin{bmatrix} B_{11} - xI_{k\times k} & 0 \\ B_{21} & -xI_{(m-k)\times(m-k)} \end{bmatrix}$$

and

$$p_{AB}(x) = \det(AB - xI) = x^{m-k} det(B_{11} - xI_{k\times k})$$
$$p_{BA}(x) = \det(BA - xI) = x^{m-k} det(B_{11} - xI_{k\times k})$$

$$\implies p_{AB}(x) = p_{BA}(x) \quad \blacksquare$$

This proof is attributed to Paul Halmos.

Reference: `https://people.math.sc.edu/howard/Classes/700/charAB.pdf`

Alternate Proof:

<u>Theorem 3</u> (stronger than what the question asks):

Let A be an $n \times m$ matrix and B be an $m \times n$ matrix. Then -

$$x^n p_{AB}(x) = x^m p_{BA}(x)$$

Proof -

Let C and D be $(m+n)\times(m+n)$ matrices such that -

$$C = \begin{bmatrix} xI_n & A \\ B & I_m \end{bmatrix}, D = \begin{bmatrix} I_n & 0 \\ -B & xI_m \end{bmatrix}$$

Then,

$$CD = \begin{bmatrix} xI_n - AB & xA \\ 0 & xI_m \end{bmatrix}, DC = \begin{bmatrix} xI_n & A \\ 0 & xI_m - BA \end{bmatrix}$$

Now,

$$\det(CD) = x^m \det(xI_n - AB) = x^m p_{AB}(x)$$

and

$$\det(DC) = x^n \det(xI_m - BA) = x^m p_{BA}(x)$$

We are done since

$$\det(CD) = \det(DC) \quad \blacksquare$$

Now, using Theorem-3 and setting $m = n$, we get

$$\begin{aligned} x^n p_{AB}(x) &= x^n p_{BA}(x) \\ \iff x^n(p_{AB}(x) - p_{BA}(x)) &= 0 \; \textit{(identically zero)} \\ \iff p_{AB}(x) &= p_{BA}(x) \quad \blacksquare \end{aligned}$$

Reference: `http://www2.math.ou.edu/~dmcculloough/teaching/slides/maa2010.pdf` This proof was attributed to *J. Schmid, A remark on characteristic polynomials, Am. Math. Monthly,77 (1970), 998-999*

Other Proof References -

`https://math.berkeley.edu/~chenhi/math54-u15/charABeqBA.pdf`

`https://www.cambridge.org/core/services/aop-cambridge-core/content/view/S0950184300003104`

4. J is a matrix over C that is a single Jordan block with eigenvalue 2. Compute e^J.

Let $N = J - 2I$. Then -

$$J = 2I + N \implies e^J = e^{2I}\, e^N = (e^2 I)\, e^N = e^2\, e^N$$

Now,

$$e^N = I + N + \frac{N^2}{2!} + ...$$

Since N is always nilpotent, we just need to take the power upto $k-1$ where J has shape $k \times k$ since $N^k = 0$. Turns out that e^N is a $k \times k$ upper triangular matrix with diagonal and the upper triangular entries all $1/(r-1)!$ where r corresponds to the r^{th} upper diagonal.

For example, if J is of shape 3×3, then we have

$$e^J = e^2 \begin{bmatrix} 1 & 1 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

For example, if J is of shape 4×4, then we have

$$e^J = e^2 \begin{bmatrix} 1 & 1 & 1/2 & 1/6 \\ 0 & 1 & 1 & 1/2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

—

5. Let V be a vector space over $\mathbb{R}$ spanned by the functions $= 1, t, e^t, te^t, t^2e^t, t^3e^t$ Let D denote differentiation. Show that D maps V into V. Determine the Jordan canonical form D on V.

$$\begin{aligned} D(1) &= 0 \\ D(t) &= 1 \\ D(e^t) &= e^t \\ D(te^t) &= (t+1)e^t \\ D(t^2e^t) &= (t^2+2t)e^t \\ D(t^3e^t) &= (t^3+3t^2)e^t \end{aligned}$$

In matrix form,

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t \\ e^t \\ te^t \\ t^2e^t \\ t^3e^t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ e^t \\ (t+1)e^t \\ (t^2+2t)e^t \\ (t^3+3t^2)e^t \end{bmatrix}$$

Hence in this basis, D is represented by the matrix

$$\text{D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{bmatrix}$$

The jordan form of this matrix turns out to be

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