Fall 2019 Final Answer Key

Problem 1. We claim that the following formula is true.

$$A^n = \begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$$

Proof. Use induction.

(1) When $n = 1$, the left hand side is $A^1 = A$ and the right hand side equals A also. So the formula is true.

(2) Assume the formula is true for n, we prove it for $n+1$. We assume

$$A^n = \begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$$

and compute

$$\begin{aligned} A^{n+1} &= A^n A \\ &= \begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1+n & n+\frac{n(n-1)}{2} \\ 0 & 1 & 1+n \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & n+1 & \frac{n(n+1)}{2} \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Hence, the formula is also true for $n+1$. By the Principle of Mathematical Induction, the formula is true for all n.

Problem 2.

(1) The reduced echelon form is

$$\begin{bmatrix} 1 & 0 & \frac{5}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & -\frac{1}{12} & -\frac{7}{12} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(2) The system corresponding to the reduced echelon form is

$$\begin{array}{rcrcrcrcl} x_1 & & + & \frac{5}{3}x_3 & + & \frac{2}{3}x_4 & + & \frac{1}{3}x_5 & = 0 \\ & x_2 & - & \frac{1}{12}x_3 & - & \frac{7}{12}x_4 & + & \frac{1}{3}x_5 & = 0 \end{array}$$

Note that x_1, x_2 are basic variables and x_3, x_4, x_5 are free variables. Solving the system, we get

$$\begin{array}{rcrcrcrcr} x_1 & = & - & \frac{5}{3}x_3 & - & \frac{2}{3}x_4 & - & \frac{1}{3}x_5 \\ x_2 & = & & \frac{1}{12}x_3 & + & \frac{7}{12}x_4 & - & \frac{1}{3}x_5 \end{array}$$

In vector form, the solutions are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -\frac{5}{3} \\ \frac{1}{12} \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -\frac{2}{3} \\ \frac{7}{12} \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Remark: Recall that the constant vectors on the right hand side form a basis for the null space of A. So,

$$\{ \begin{bmatrix} -\frac{5}{3} \\ \frac{1}{12} \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{2}{3} \\ \frac{7}{12} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \\ 0 \\ 1 \end{bmatrix} \}$$

is a basis for the null space of A. Hence the dimension of the null space of A is 3 (also see below).

(3) No. In order for $A\mathbf{x} = \mathbf{b}$ to have a solution for every choice of $\mathbf{b} = [\mathbf{b_1}, \mathbf{b_2}, \mathbf{b_3}]^\top$, A has to have a pivot position in every row. But A has no pivot position in the 3rd row.

(4) Since A has two pivot positions, the rank of A is 2. Hence the dimension of the row space of $A =$ the dimension of the column space of $A =$ rank of $A = 2$. By the rank theorem, the dimension of the null space of A is $5 - 2 = 3$. [Rank Theorem: Let A be an $m \times n$ matrix. Then, the rank $(A) +$ dimNull $(A) = n$.] Since $A^\top$ is a 5×3 matrix, and rank$(A^\top) =$ rank$(A) = 2$, again by the rank theorem, dimNull$(A^\top) = 3-$rank$(A^\top) = 3 - 2 = 1$.

bf Remark: Note that the dimension of the null space of $A^\top$ can also be calculated by using the Fundamental Theorem of Linear Algebra:

$$\mathrm{Col}(A)^\perp = \mathrm{Null}(A^\top), \quad \mathrm{Row}(A)^\perp = \mathrm{Null}(A)$$

In this problem Col(A) is a subspace of $\mathbb{R}^3$, so $\dim \mathrm{Col}(A)^\perp = 3 - \dim \mathrm{Col}(A) = 3 - \mathrm{rank}(A) = 3 - 2 = 1$. Hence, $\dim \mathrm{Null}(A^\top) = \dim \mathrm{Col}(A)^\perp = 1$.

Remark: Recall that the pivot columns in the original A form a basis for the column space of A. In this problem, the first and second column are pivot columns, hence the following is a basis for the column space of A.

$$\{ \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -12 \\ -4 \end{bmatrix} \}$$

Recall also that the rows in an **row echelon form of** A form a basis for the row space of A. Hence, the following is a basis for the row space of A.

$$\{(1,0,\frac{5}{3},\frac{2}{3},\frac{1}{3}),(0,1,-\frac{1}{12},-\frac{1}{12},\frac{1}{3})\}$$

Problem 3.

(1) $(2,-6,1)$ is a normal vector to the plane. Normalizing it, we get an unit normal vector

$$\mathbf{n}=\frac{(2,-6,1)}{||(2,-6,1)||}=\frac{(2,-6,1)}{\sqrt{41}}=(\frac{2}{\sqrt{41}},-\frac{6}{\sqrt{41}},\frac{1}{\sqrt{41}})$$

(2) Let $T(x_0,y_0,z_0)$ be any point on the plane. The distance from P to H is the norm of the orthogonal projection of the vector $\vec{PT}=(1,2,3)-(x_0,y_0,z_0)$ onto n. The orthogonal projection of $\vec{PT}$ onto n is

$$\begin{aligned}\frac{\vec{PT}\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n} &= \frac{(1,2,3)\cdot\mathbf{n}-(x_0,y_0,z_0)\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n}\\ &= \frac{-7+2x_0-6y_0+z_0}{\sqrt{41}}\mathbf{n}\\ &= \frac{-7+0}{\sqrt{41}}\mathbf{n}=-\frac{7}{\sqrt{41}}\mathbf{n},\end{aligned}$$

where we have used the fact that $2x_0-6y_0+z_0=0$, since $T(x_0,y_0,z_0)$ is a point on the plane. Since $\mathbf{n}$ is a unit vector, the norm of the projection is $\frac{7}{\sqrt{41}}$. Thus, the distance from P to H is $\frac{7}{\sqrt{41}}$.

(3) In the parametric form for L, $\mathbf{x_0}+t\mathbf{v}$, $\mathbf{x_0}$ is a vector whose head is on the line, and $\mathbf{v}$ is a vector that is parallel to the line. Since Q is a point on the line, we can choose $\mathbf{x_0}=\vec{OQ}=(3,1,0)$. Now we need to find a vector that is parallel to L. By assumption, L is contained in the plane H, so L is orthogonal to a normal vector of H, hence L is orthogonal to $(2,-6,1)$. Also by assumption, L is orthogonal to $(1,2,3)$. Since the cross product of two vectors is orthogonal to both of these vectors, L is parallel to the cross product of $(2,-6,1)$ and $(1,2,3)$. Let's now calculate the cross product:

$$\begin{bmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 2 & -6 & 1\\ 1 & 2 & 3\end{bmatrix}=-20\mathbf{i}-5\mathbf{j}+10\mathbf{k}$$

So we can choose $\mathbf{v}=(-20,-5,10)$. Hence the equation of the line is $(3,1,0)+t(-20,-5,10)$.

Problem 4.

(1)

$$\begin{aligned}q(x_1,x_2,x_3) &= 2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_1x_3-2x_2x_3\\ &= [x_1,x_2,x_3]\begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}\\ &= \mathbf{x}^\top A\mathbf{x},\end{aligned}$$

where the symmetric matrix

$$A=\begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{bmatrix}$$

(2) To find eigenvalues, we first compute the characteristic polynomial.

$$\begin{aligned}
\det(\lambda I - A) &= \begin{vmatrix} \lambda-2 & 1 & 1 \\ 1 & \lambda-2 & 1 \\ 1 & 1 & \lambda-2 \end{vmatrix} \\
&= \begin{vmatrix} 0 & 1-(\lambda-2)^2 & 3-\lambda \\ 1 & \lambda-2 & 1 \\ 0 & 3-\lambda & \lambda-3 \end{vmatrix} ((2-\lambda) \text{ Row 2 added to Row 1})(\text{ - Row 2 added to Row 3}) \\
&= -\begin{vmatrix} 1-(\lambda-2)^2 & 3-\lambda \\ 3-\lambda & \lambda-3 \end{vmatrix} (\text{cofactor expansion along first column}) \\
&= [(\lambda-2)^2-1](\lambda-3)+(\lambda-3)^2 = \lambda(\lambda-3)^2
\end{aligned}$$

Solving $\det(\lambda I - A) = 0$, we get three eigenvalues, in descending order,

$$\lambda_1 = 3, \quad \lambda_2 = 0, \quad \lambda_3 = 0.$$

(3) We now find P whose columns consist of orthonormal eigenvectors. First, we find a basis for the eigenspace corresponding to $\lambda_1 = 3$. Solving $(3I - A)\mathbf{x} = \mathbf{0}$, we get

$$\mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

So $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ form a basis for the eigenspace corresponding to the eigenvalue $\lambda_1 = 3$. Note that $\mathbf{u}$ and $\mathbf{v}$ are not orthogonal, so we need to use Gram-Schmidt process to get an orthogonal basis:

$$\begin{aligned}
\mathbf{u_1} &= \mathbf{u} \\
\mathbf{u_2} &= \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u_1}}{\mathbf{u_1} \cdot \mathbf{u_1}} \mathbf{u_1} \\
&= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}
\end{aligned}$$

So,

$$\mathbf{u_1} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u_2} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

form an orthogonal basis for the eigenspace. Now, we normalize it to get an orthonormal basis for the eigenspace corresponding to $\lambda_1 = 3$:

$$\left\{ \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{bmatrix} \right\}$$

Now, we find a basis for the eigenspace corresponding to $\lambda_2 = \lambda_3 = 0$. Solving $(0I - A)\mathbf{x} = \mathbf{0}$, we find

$$\mathbf{x} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Normalizing the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we get an orthonormal basis for the eigenspace corresponding to the eigenvalue 0:$\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \}$ We now combine the two bases to get an orthonormal set of 3 eigenvectors:

$$\left\{ \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \right\}$$

Let

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

The P is an orthogonal matrix and if we let $\mathbf{y} = P^{\top}\mathbf{x}$, or equivalently, $\mathbf{x} = P\mathbf{y}$, the quadratic form becomes

$$q(x_1, x_2, x_3) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 = 3y_1^2,$$

the standard form for the quadratic.

Problem 5.

(1) We only verify the positivity, and leave the rest to you.

$$< p, p >= \int_0^1 [p'(x)]^2 \, dx \; + [p(0)]^2$$

So, $<p,p> \geq 0$ and

$$\begin{aligned}
<p,p>=0 \iff & p'(x)=0 \text{ for all } x \in [0,1] \text{ and } p(0)=0 \\
\iff & p(x) \text{ is a constant function for } \ x \in [0,1] \text{ and } p(0)=0 \\
\iff & p(x)=0 \text{ for } \ x \in [0,1] \\
\iff & p^{(k)}(0)=0 \text{ for } \ k=0,1,\ldots,n \ \text{ since } p \text{ is a polynomial of degree at most } n. \\
\iff & p(x)=0 \text{ for all } \ x \in (-\infty,\infty) \\
\iff & p \ \text{ is the zero vector in } \mathbb{P}_n
\end{aligned}$$

(2) Let

$$p_0(x)=1,\ p_1(x)=x,\ p_2(x)=x^2,\ p_3(x)=x^3.$$

One can check that $\{p_0,p_1,p_2,p_3\}$ is a basis for $\mathbb{P}_3$. We now use the Gram-Schmidt to get an orthogonal basis.

$$\begin{aligned}
q_1(x) &= p_0(x)=1 \\
q_2(x) &= p_1(x)-\frac{<p_1,q_1>}{<q_1,q_1>}q_1(x)=p_1(x)=x \\
q_3(x) &= p_2(x)-\frac{<p_2,q_1>}{<q_1,q_1>}q_1(x)-\frac{<p_2,q_2>}{<q_2,q_2>}q_2(x) \\
&= p_2(x)-q_2(x)=x^2-x \\
q_4(x) &= p_3(x)-\frac{<p_3,q_1>}{<q_1,q_1>}q_1(x)-\frac{<p_3,q_2>}{<q_2,q_2>}q_2(x)-\frac{<p_3,q_3>}{<q_3,q_3>}q_3(x) \\
&= p_3(x)-q_2(x)-\frac{3}{2}q_3(x) \\
&= x^3-x-\frac{3}{2}(x^2-x)=x^3-\frac{3}{2}x^2+\frac{1}{2}x
\end{aligned}$$

So,

$$\{1,x,x^2-x,x^3-\frac{3}{2}x^2+\frac{1}{2}x\}$$

is an orthogonal basis for $\mathbb{P}^3$. We now normalize it to get an orthonormal basis:

$$\{1,x,\sqrt{3}(x^2-x),\sqrt{20}(x^3-\frac{3}{2}x^2+\frac{1}{2}x)\}$$

That is,

$$\{1,x,\sqrt{3}(x^2-x),\sqrt{5}(2x^3-3x^2+x)\}$$

Problem 6

(1) We need to verify that T preserves addition and scalar multiplication. We compute directly from the definition of T.

$$T(X+Y)=X+Y+(X+Y)^\top=X+Y+X^\top+Y^\top=X+X^\top+Y+Y^\top=T(X)+T(Y)$$

$$T(kX)=kX+(kX)^\top=kX+kX^\top=k(X+X^\top)=kT(X)$$

(2) We need to express the image of each of the basis vectors under T. Let's introduce some notations:

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

One can think of E_1 as $e_1 = [1, 0, 0, 0]^\top$, and E_2 as $e_2 = [0, 1, 0, 0]^\top$, E_3 as $e_3 = [0, 0, 1, 0]^\top$ and E_4 as $e_1 = [0, 0, 0, 1]^\top$

$$T(E_1) = T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^\top = 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 2E_1$$

$$T(E_2) = T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^\top = E_2 + E_3$$

$$T(E_3) = T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}^\top = E_2 + E_3$$

$$T(E_4) = T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}^\top = 2E_4$$

So the matrix for T is

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

(3) We will omit the details. The characteristic polynomial of A is $\lambda(\lambda - 2)^3$. So the eigenvalues of A are $2, 0$. The basis for the eigenspace of A belonging to 2 is

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

So the corresponding basis for the eigenspace of T belonging to 2 is

$$\{E_1, E_2 + E_3, E_4\}$$

That is

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

The basis for the eigenspace of A belonging to 0 is

$$\left\{\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}\right\}$$

So the corresponding basis for the eigenspace of T belonging to 2 is

$$\{E_2 - E_3\}$$

That is

$$\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\}$$

(3) We combine the bases vectors for the eigenspaces to get a basis that diagonalizes A:

$$\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}\right\}$$

One checks easily that this is already orthogonal. Hence, all we need to do is to normalize it to get an orthonormal basis:

$$\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}\right\}$$

The corresponding vectors in $M_{2\times 2}$ are

$$B' = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}\right\}$$

Problem 7

(a) One can use the Principle of Mathematical Induction to prove it. We omit the details.

(b) We first diagonalize

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

The eigenvalues of A are $\alpha = \frac{1-\sqrt{5}}{2}$, $\beta = \frac{1+\sqrt{5}}{2}$. A basis for the eigenspace corresponding to α is

$$\left\{\begin{bmatrix} \alpha \\ 1 \end{bmatrix}\right\}$$

and a basis for the eigenspace corresponding to β is

$$\left\{\begin{bmatrix} \beta \\ 1 \end{bmatrix}\right\}$$

Let

$$P = \begin{bmatrix} \alpha & \beta \\ 1 & 1 \end{bmatrix}$$

Then P is invertible and $A = P \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} P^{-1}$. So $A^{n+1} = P \begin{bmatrix} \alpha^{n+1} & 0 \\ 0 & \beta^{n+1} \end{bmatrix} P^{-1}$ Since $P^{-1} = \frac{1}{\alpha-\beta} \begin{bmatrix} 1 & -\beta \\ -1 & \alpha \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & \beta \\ 1 & -\alpha \end{bmatrix}$. Hence,

$$\begin{aligned} A^{n+1} &= \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha & \beta \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^{n+1} & 0 \\ 0 & \beta^{n+1} \end{bmatrix} \begin{bmatrix} -1 & \beta \\ 1 & -\alpha \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \beta^{n+2} - \alpha^{n+2} & \beta\alpha^{n+2} - \alpha\beta^{n+2} \\ \beta^{n+1} - \alpha^{n+1} & \beta\alpha^{n+1} - \alpha\beta^{n+1} \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \beta^{n+2} - \alpha^{n+2} & \beta^{n+1} - \alpha^{n+1} \\ \beta^{n+1} - \alpha^{n+1} & \beta^{n} - \alpha^{n} \end{bmatrix}, \end{aligned}$$

where we have used the fact that $\alpha\beta = -1$. Hence, we have

$$\begin{bmatrix} a_{n+2} & a_{n+1} \\ a_{n+1} & a_n \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \beta^{n+2} - \alpha^{n+2} & \beta^{n+1} - \alpha^{n+1} \\ \beta^{n+1} - \alpha^{n+1} & \beta^{n} - \alpha^{n} \end{bmatrix},$$

We now express α^n in terms of β. Recall that $\alpha = \frac{1-\sqrt{5}}{2}$ and $\beta = \frac{1+\sqrt{5}}{2}$. So, $\alpha = 1 - \beta$. Hence, $\beta^n - \alpha^n = \beta^n - (1-\beta)^n$. Therefore,

$$a_n = \frac{1}{\sqrt{5}}[\beta^n - (1-\beta)^n]$$

In particular, $a_{1000} = \frac{1}{\sqrt{5}}[\beta^{1000} - (1-\beta)^{1000}]$.

Problem 8

Before we begin, we observe that the factors in the product $(\lambda_1 I - A) \cdots (\lambda_r I - A)$ are commutative. For example, $(\lambda_1 I - A)(\lambda_2 I - A) = (\lambda_2 I - A)(\lambda_1 I - A)$ (Check it!).

(1) Since A is diagonalizable, there are n linearly independent eigenvectors which we denote by

$$\mathbf{v_1}, \ldots, \mathbf{v_n}$$

To simplify notations, we put

$$B = (\lambda_1 I - A) \cdots (\lambda_r I - A)$$

In order to show $B = 0$, all we have to do is to show that for any vector $\mathbf{v_i}$, $B\mathbf{v_i} = \mathbf{0}$. [Here is an argument: Let $P = [\mathbf{v_1}, \ldots, \mathbf{v_n}]$. Since the columns of P are linearly independent, P is invertible. If we can show that $B\mathbf{v_i} = \mathbf{0}$ for any i, then $BP = 0$. But P is invertible, multiplying each side from right by P^{-1}, we get $BPP^{-1} = 0$. So, $B = 0$.] Since $\mathbf{v_i}$ is an eigenvector, it has to belong some eigenvalue, say λ_j. In other words, $A\mathbf{v_i} = \lambda_j \mathbf{v_i}$. Since factors in B are commutative, we move, if necessary, the factor $(\lambda_j I - A)$ to the end, so B is expressed as

$$B = \cdots (\lambda_j I - A)$$

So,

$$B\mathbf{v_i} = \cdots(\lambda_j I - A)\mathbf{v_i} = \cdots(\lambda_j \mathbf{v_i} - A\mathbf{v_i}) = \cdots \mathbf{0} = \mathbf{0}.$$

Thus $B = 0$ as we noted earlier.

(2) Before we begin, let's review some basic facts about linear transformation (matrix transformation). Let K be an $m \times n$ matrix. Define a map L from $\mathbb{R}^n$ to $\mathbb{R}^m$ by $L(\mathbf{x}) = K\mathbf{x}$. L is a matrix transformation and is also linear transformation. The following facts are important.

$$\begin{aligned} L \text{ is onto (surjective)} \quad &\Leftrightarrow \quad K\mathbf{x} = \mathbf{b} \text{ has a solution for each } \mathbf{b} \in \mathbb{R}^m \\ &\Leftrightarrow \quad \text{columns of } K \text{ span } \mathbb{R}^m \\ &\Leftrightarrow \quad \text{Every row of } K \text{ has a pivot position} \end{aligned}$$

$$\begin{aligned} L \text{ is one-to-one (injective)} \quad &\Leftrightarrow \quad K\mathbf{x} = \mathbf{0} \text{ has only the trivial solution} \\ &\Leftrightarrow \quad \text{columns of } K \text{ are linearly independent} \\ &\Leftrightarrow \quad \text{Every column of } K \text{ has a pivot position} \end{aligned}$$

As a consequence,

$$\begin{aligned} L \text{ is onto (surjective)} \quad &\Rightarrow \quad \text{Every row of } K \text{ has a pivot position} \\ &\Rightarrow \quad m \leq n. \end{aligned}$$

$$\begin{aligned} L \text{ is one-to-one (injective)} \quad &\Rightarrow \quad \text{Every column of } K \text{ has a pivot position} \\ &\Rightarrow \quad n \leq m. \end{aligned}$$

Now, we prove the following fact: Let C and D be $n \times n$ matrices. Then

$$\boxed{\dim \mathrm{Null}(CD) \leq \dim \mathrm{Null}(C) + \dim \mathrm{Null}(D)}$$

Assume $\dim \mathrm{Null}(CD) = k$. If $k = 0$, the equality holds automatically. So we assume $k > 0$. Let $\{\mathbf{v_1}, \ldots, \mathbf{v_k}\}$ denote a basis for $\mathrm{Null}(CD)$. Then $CD\mathbf{v_i} = 0$ for all i. Let $M = [\mathbf{v_1}, \ldots, \mathbf{v_k}]$. Then $CDM = 0$. This implies that for every vector $\mathbf{w} \in \mathbb{R}^k$, $CDM\mathbf{w} = \mathbf{0}$, i.e., $C(DM\mathbf{w}) = \mathbf{0}$. In other words, $DM\mathbf{w} \in \mathrm{Null}(C)$. But$\mathbf{w}$ is an arbitrary vector, this means that $\mathrm{Col}(DM) \subset \mathrm{Null}C$. So $\dim \mathrm{Col}(DM) \leq \dim \mathrm{Null}(C)$. Note that DM is an $n \times k$ matrix, by the Rank Theorem, $k - \dim \mathrm{Null}(DM) \leq \dim \mathrm{Null}(C)$, that is , $k \leq \dim \mathrm{Null}(C) + \mathrm{Null}(DM)$. Since $k = \dim \mathrm{Null}(CD)$, we have

$$\dim \mathrm{Null}(CD) \leq \dim \mathrm{Null}(C) + \dim \mathrm{Null}(DM)$$

What remains is to show that $\dim \mathrm{Null}(DM) \leq \dim \mathrm{Null}(D)$. To show this, we use the above observation that if there is a linear transformation from an $n-$dimensional vector space to an $m-$dimensional vector space that is one-to-one, then $n \leq m$. So, we construct a linear transformation L from $\mathrm{Null}(DM)$ to $\mathrm{Null}(D)$ by defining $L(\mathbf{w}) = M\mathbf{w}$, for $\mathbf{w} \in \mathrm{Null}(DM)$. We want to make sure that $M\mathbf{w} \in \mathrm{Null}(D)$. But this is easy to see. In fact, since $\mathbf{w} \in \mathrm{Null}(DM)$, $(DM)\mathbf{w} = \mathbf{0}$. So, $D(M\mathbf{w}) = \mathbf{0}$. This means that $M\mathbf{w} \in \mathrm{Null}(D)$. Now we notice that the columns of M are linearly independent, since those columns are basis vectors. So L is one-to-one. Hence $\dim \mathrm{Null}(DM) \leq \dim \mathrm{Null}(D)$.. This completes the proof of the boxed formula.

The formula we have just proved, when applied to product of three matrices, gives

$$\dim \mathrm{Null}(CDE) \leq \dim \mathrm{Null}(CD) + \dim \mathrm{Null}(E) \leq \dim \mathrm{Null}(C) + \dim \mathrm{Null}(D) + \dim \mathrm{Null}(E)$$

So, by applying the formula repeatedly, we have

$$\boxed{\dim \mathrm{Null}(C_1 \cdots C_r) \leq \dim \mathrm{Null}(C_1) + \ldots + \dim \mathrm{Null}(C_r)}$$

We now apply the above formula to

$$(\lambda_1 I - A) \cdots (\lambda_r I - A):$$

$$\dim \mathrm{Null}[(\lambda_1 I - A) \cdots (\lambda_r I - A)] \leq \dim \mathrm{Null}(\lambda_1 I - A) + \ldots + \dim \mathrm{Null}(\lambda_r I - A)$$

By assumption, $(\lambda_1 I - A) \cdots (\lambda_r I - A) = 0$, hence $\dim \mathrm{Null}[(\lambda_1 I - A) \cdots (\lambda_r I - A)] = n$. Thus, we have

$$\dim \mathrm{Null}(\lambda_1 I - A) + \ldots + \dim \mathrm{Null}(\lambda_r I - A) \geq n$$

Since $\lambda_1, \ldots, \lambda_r$ are distinct eigenvalues, the left hand side represents the number of linearly independent eigenvectors. So one has at least n linearly independent eigenvectors. But of course, the number of linearly independent vectors cannot be lager than n. Hence A must have exactly n linearly independent eigenvectors. Therefore, A is diagonalizable.

Remark: Combining the rank theorem with

$$\dim \mathrm{Null}(CD) \leq \dim \mathrm{Null}(C) + \dim \mathrm{Null}(D),$$

one gets

$$\boxed{\mathrm{rank}(CD) \geq \mathrm{rank}(C) + \mathrm{rank}(D) - n}$$