

Problem Set 9

D. Zack Garza

November 26, 2019

Contents

1	Problem 1	1
1.1	Part 1	1
1.2	Part 2	3
2	Problem 2	3
2.1	Part 1	3
2.2	Part 2	3
3	Problem 3	4
4	Problem 4	4
5	Problem 5	6
5.1	Part 1	6
5.2	Part 2	7
5.3	Part 3	7
5.4	Part 4	8
6	Problem 6	8

Note: I use the convention that \mathbf{a} denotes a column vector and \mathbf{a}^t a row vector, and if A is a matrix, then $(A)_{ij} = a_{ij}$ denotes the entry in the i th row and j th column.

1 Problem 1

1.1 Part 1

Let $A = (a_{ij})$ and consider ϵ_{ij} , the matrix with a 1 in the i th row and j th column and zeros elsewhere.

Then, for a fixed (i, j) , if we write $A = [\mathbf{a}_1^t, \mathbf{a}_2^t, \dots, \mathbf{a}_n^t]$ as a block matrix of column vectors, we have

$$A\mathbf{e}_{ij} = [0, 0, \dots, \mathbf{a}_i^t, 0, \dots, 0]$$

as a block matrix where \mathbf{a}_i^t occurs as the j th column.

In other words, right-multiplication by \mathbf{e}_{ij} selects column i from A , placing it in column j of a matrix of zeros.

For example, for $(i, j) = (3, 2)$ we have

$$A\mathbf{e}_{32} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a_{13} & 0 \\ 0 & a_{23} & 0 \\ 0 & a_{33} & 0 \end{pmatrix},$$

which is a matrix that contains column 3 of A (the i value) as its 2nd column (the j value).

On the other hand, *left* multiplication by \mathbf{e}_{ij} selects the j th **row** of A and places it the i th **row** of a zero matrix, so for example we have

$$\mathbf{e}_{32}A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

In general, these two products will not be equal, since the first has a nontrivial column and the latter has a nontrivial row. If $A \in Z(M_n(R))$, these two must be equal, so we can equate corresponding entries to find that

- $a_{21} = 0$, from comparing entries in row 3, column 1,
- $a_{23} = 0$, from comparing entries in row 3, column 3
- $a_{22} = a_{33}$ by comparing entries in row 3, column 2.

Letting the multiplication run over all possibilities for \mathbf{e}_{ij} yields $a_{ii} = a_{jj}$ for every pair i, j and $a_{ij} = 0$ whenever $i \neq j$. Setting $r = a_{ii} = a_{jj}$ for all $1 \leq i, j \leq n$ forces A to be a matrix of the form

$$A = \begin{pmatrix} r & 0 & 0 & \cdots & 0 \\ 0 & r & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r \end{pmatrix} := rI_n.$$

To see that we must have $r \in Z(R)$, let $sI_n \in Z(M_n(R))$ be arbitrary, where s is not assumed to be in $Z(R)$. Then $(rI_n)(sI_n) = (sI_n)(rI_n)$ by assumption, since these are matrices in the center of $M_n(R)$. But $M_n(R)$ is an R -module, and so the scalars r, s commute with the module elements I_n . This means that we in fact have

$$\begin{aligned} (rI_n)(sI_n) &= (rs)I_n^2 = (rs)I_n, \\ (sI_n)(rI_n) &= (sr)I_n^2 = (sr)I_n \\ &\implies (rs)I_n = (sr)I_n \\ &\implies (rs - sr)I_n = 0_n, \end{aligned}$$

the $n \times n$ zero matrix.

But then by equating (for example) the 1, 1 entry of the matrix $(rs - sr)I_n$ with the corresponding entry in 0_n , we find $rs - sr = 0_R$, which means $rs = sr \in R$.

Now since $s \in R$ was arbitrary, we find that $r \in Z(R)$ as desired.

1.2 Part 2

Define a map

$$\begin{aligned}\phi : Z(R) &\rightarrow Z(M_n(R)) \\ r &\mapsto rI_n.\end{aligned}$$

By part 1, this map is surjective. To see that it is also injective, we can consider $\ker \phi = \{r \in Z(R) \mid rI_n = 0_n\}$, which clearly forces $r = 0_R$. It is also a homomorphism of R -modules, since $\phi(rx + y) = (rx + y)I_n = r(xI_n) + yI_n$.

Thus by the first isomorphism theorem, we have $Z(R) \cong Z(M_n(R))$.

2 Problem 2

2.1 Part 1

If A, B are (skew)-symmetric, then $A^t = \pm A$ and $B^t = \pm B$ respectively. But then

$$(A + B)^t = A^t + B^t = \pm A + \pm B = \pm(A + B),$$

which shows that $A + B$ is (skew)-symmetric.

2.2 Part 2

\implies : Suppose that whenever A, B are symmetric then AB is symmetric as well.

We then have $(AB)^t = AB$ by assumption, and then by calculation we have $(AB^t) = B^t A^t = BA$, so $AB = BA$.

\impliedby : Suppose that $AB = BA$ and A, B are symmetric. We want to show that AB is also symmetric, so we compute

$$(AB)^t = B^t A^t = BA = AB.$$

□

Now let $B \in M_n(R)$ be arbitrary. We have

- $(BB^t)^t = (B^t)^t B^t = BB^t$, so BB^t is symmetric,
- $(B + B^t)^t = B^t + (B^t)^t = B^t + B = B + B^t$, so $B + B^t$ is symmetric,
- $(B - B^t)^t = B^t - B = -(B + B^t)$, so $B - B^t$ is skew-symmetric

3 Problem 3

Definition: We say $A \sim B$ in $M_n(R)$ \iff there exists an invertible P such that $B = PAP^{-1}$.

- Reflexive, $A \sim A$:

Take $P = I_n$ the identity matrix.

- Symmetric, $A \sim B \implies B \sim A$:

$B = PAP^{-1} \implies BP = PA \implies P^{-1}BP = A$, so we can take $Q = P^{-1}$ to yield $A = QBQ^{-1}$.

- Transitive, $A \sim B \& B \sim C \implies A \sim C$:

If $B = PAP^{-1}, C = QBQ^{-1}$, then $C = Q(PAP^{-1})Q^{-1} = (QP)A(QP)^{-1}$, so take $L = QP$ to yield $C = LAL^{-1}$.

Definition: We say $A \sim B$ in $M(n \times n, R)$ $\iff B = PAQ$ with $P \in GL(n, R), Q \in GL(m, R)$.

- Reflexive, $A \sim A$:

Take $P = I_{m,n}$ the matrix with 1s on the diagonal and zeros elsewhere, and $Q = P^t$.

- Symmetric, $A \sim B \implies B \sim A$:

$B = PAQ \implies BQ^{-1} = PA \implies P^{-1}BQ^{-1} = A$, so we can take $S = P^{-1}, T = Q^{-1}$ to yield $A = QBT$.

- Transitive, $A \sim B \& B \sim C \implies A \sim C$:

If $B = PAQ, C = RBS$, then $C = R(PAQ)S = (RP)A(QS)$, so take $L = RP, M = QS$ to yield $C = LAM$.

4 Problem 4

Lemma: The rank-nullity theorem holds over division rings.

Proof: A linear map $\phi : D^m \rightarrow D^n$ induces a short exact sequence:

$$0 \rightarrow \ker \phi \rightarrow D^m \xrightarrow{\phi} \text{im } \phi \rightarrow 0$$

But every module over a division ring is free; in particular, $\text{im } \phi \leq D^n$ is a module over D and is thus free. So by a lemma in class, since the right-most term is a free module, this sequence splits and we have

$$D^m \cong \ker \phi \oplus \text{im } \phi$$

and taking dimensions yields

$$m = \dim \ker(\phi) + \text{rank}(\phi).$$

□

1. $A \in M(n \times m, D)$ has a left inverse $B \iff \text{rank}(A) = m$:

\implies : Suppose toward the contrapositive that $\text{rank}(A) < m$, so A has at least one pair of linearly dependent columns. So wlog write

$$A = [\mathbf{a}_1^t, \mathbf{a}_2^t, \dots, \mathbf{a}_m^t]$$

in block form with each \mathbf{a}_i a column vector, and we can assume that $\mathbf{a}_1, \mathbf{a}_2$ are linearly dependent.

Now suppose such a left inverse B were to exist. Write it in block form as

$$B = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]^t,$$

so each \mathbf{b}_i is a row of B .

Now if $BA = I_m$ is to hold, noting that $(BA)_{ij} = \langle \mathbf{b}_i, \mathbf{a}_j \rangle$, we must have

$$I_{1,1} = \langle \mathbf{b}_1, \mathbf{a}_1 \rangle = 1$$

$$I_{1,2} = \langle \mathbf{b}_1, \mathbf{a}_2 \rangle = 0$$

$$I_{1,3} = \langle \mathbf{b}_1, \mathbf{a}_3 \rangle = 0$$

$$\vdots$$

$$I_{2,1} = \langle \mathbf{b}_2, \mathbf{a}_1 \rangle = 0$$

$$I_{2,2} = \langle \mathbf{b}_2, \mathbf{a}_2 \rangle = 1$$

$$I_{2,3} = \langle \mathbf{b}_2, \mathbf{a}_3 \rangle = 0$$

$$\vdots$$

But the claim is that this can *not* happen if $\mathbf{a}_1, \mathbf{a}_2$ are linearly dependent. To see why, note that the linear dependence supplies elements $d_1, d_2 \neq 0 \in D$ such that $d_1\mathbf{a}_1 + d_2\mathbf{a}_2 = \mathbf{0}$. But then taking inner products against, e.g. \mathbf{b}_1 (that is, applying $\langle \mathbf{b}_1, \cdot \rangle$ to everything in sight), we obtain

$$\begin{aligned} d_1\mathbf{a}_1 + d_2\mathbf{a}_2 &= \mathbf{0} \\ \implies \langle \mathbf{b}_1, d_1\mathbf{a}_1 \rangle + \langle \mathbf{b}_1, d_2\mathbf{a}_2 \rangle &= \langle \mathbf{b}_1, \mathbf{0} \rangle = 0 \\ \implies d_1\langle \mathbf{b}_1, \mathbf{a}_1 \rangle + d_2\langle \mathbf{b}_1, \mathbf{a}_2 \rangle &= \langle \mathbf{b}_1, \mathbf{0} \rangle = 0 \\ \implies d_1\langle \mathbf{b}_1, \mathbf{a}_1 \rangle + d_2\langle \mathbf{b}_1, \mathbf{a}_2 \rangle &= 0 \\ \implies d_1 + d_2\langle \mathbf{b}_1, \mathbf{a}_2 \rangle &= 0 \\ \implies \langle \mathbf{b}_1, \mathbf{a}_2 \rangle &= -\frac{d_1}{d_2} \neq 0, \end{aligned}$$

which contradicts $\langle \mathbf{b}_1, \mathbf{a}_2 \rangle = 0$ as required by the previous equations.

\Leftarrow : Suppose $\text{rank}(A) = m$, so A has m linearly independent columns – note that this is *all* of its columns.

Note: since row rank equals column rank, this also says that A has m linearly independent rows, so $n \geq m$.

Viewing A as a representative of a map $\phi : D^m \rightarrow D^n$, we find that $\dim \operatorname{im} \phi = m \leq n$. In particular, from the rank nullity theorem, we have

$$m = \dim \ker \phi + \operatorname{rank}(\phi) = \dim \ker \phi + m \implies \dim \ker \phi = 0.$$

So $\ker A = \{\mathbf{0}\}$, and A represents an injective map $f_A : D^m \rightarrow D^n$.

But any injective *set* map $f : S_1 \rightarrow S_2$ has a left-inverse g such that $g \circ f = \operatorname{id}_{S_1}$. So $f_A : D^m \rightarrow D^n$ as a *set* map has a left inverse $g_B : D^n \rightarrow D^m$ satisfying $g_B \circ f_A = \operatorname{id}_{D^m}$. But then taking the matrix associated to g_B yields a matrix $B \in M(m \times n, D)$ such that $BA = I_m$ as desired. \square

2. A has a right inverse $B \iff \operatorname{rank}(A) = n$:

\implies : By a similar argument, supposing that $\operatorname{rank} A < n$ but $AB = I_n$ for some B , we find that A has at least two linearly dependent *rows* this time, say $\mathbf{a}_1, \mathbf{a}_2$, whereas we obtain a system of equations of the form $\langle \mathbf{a}_i, \mathbf{b}_k \rangle = \delta_{ik}$ where \mathbf{b}_i are now the columns of B .

In a similar manner, the linear dependence forces, say, $\langle \mathbf{a}_2, \mathbf{b}_1 \rangle \neq 0$, which is a contradiction.

\impliedby : By another similar argument, we find that A represents a map $f_A : D^m \rightarrow D^n$, and since $\operatorname{rank} A = \dim \operatorname{im} A = n$, we find that A represents a surjective map f_A . Surjective set maps have *right* inverses, so there is some $g_B : D^n \rightarrow D^m$ such that $f_A \circ g_B = \operatorname{id}_{D^n}$, and when translated to matrices this yields $AB = I_n$. \square

5 Problem 5

5.1 Part 1

\impliedby : Suppose that $A\mathbf{x} = \mathbf{b}$ has a solution \mathbf{x} .

Write $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m]^t$ in block form with each \mathbf{a}_i a row of A . By definition, a solution to this equation is a $\mathbf{x} = (x_i)$ such that for each i , we have $\langle \mathbf{a}_i, \mathbf{x} \rangle = b_i$ (by carrying out the matrix multiplication).

But

$$\begin{aligned} \langle \mathbf{a}_i, \mathbf{x} \rangle &= b_i \\ \implies \sum_{j=1}^m a_{ij}x_j &= b_i, \end{aligned}$$

which says that the collection x_1, \dots, x_n solves the equation

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{im}x_m = b_i$$

for every i , which is exactly the statement that the x_i simultaneously solve the given system.

\implies : Suppose that the given system has a simultaneous solutions x_1, x_2, \dots, x_n , and consider the matrix equation $A\mathbf{x} = \mathbf{b}$.

Letting $\mathbf{x} = [x_1, x_2, \dots, x_n]$, we can rewrite

$$b_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{im}x_m = \langle \mathbf{a}_i, \mathbf{x} \rangle,$$

where $\mathbf{a}_i = [a_{i1}, a_{i2}, \dots, a_{im}]$.

But then \mathbf{a}_i is the i th row of A , and $A\mathbf{x} = \mathbf{b}$ has a solution iff there is a \mathbf{x} such that $\langle \mathbf{a}_i, \mathbf{x} \rangle = b_i$ for all i , which is exactly what we've constructed.

5.2 Part 2

Noting that applying a row operation to A is the same as taking the product EA for some elementary matrix E , we can write $A_1 = \left(\prod_{i=1}^{\ell} E_i\right) A$ and $B_1 = \left(\prod_{i=1}^{\ell} E_i\right) B$,

thus

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \implies E_{\ell}A\mathbf{x} &= E_{\ell}\mathbf{b} \\ \implies E_{\ell-1}E_{\ell}A\mathbf{x} &= E_{\ell-1}E_{\ell}\mathbf{b} \\ &\vdots \\ \implies E_1E_2 \cdots E_{\ell}A\mathbf{x} &= E_1E_2 \cdots E_{\ell}\mathbf{b} \\ \implies A_1\mathbf{x} &= B_1 \end{aligned}$$

5.3 Part 3

1. $AX = B$ has a solution $\iff \text{rank}(A) = \text{rank}(C)$:

Note that we can only have $\text{rank } C \geq \text{rank } A$.

\implies :

Suppose that $AX = B$ has a solution; then \mathbf{b} is in the column space of A . But this says that

$$\text{span}(\{\mathbf{a}_i\}) = \text{span}(\{\mathbf{a}_i\} \cup \{\mathbf{b}\}),$$

where \mathbf{a}_i are the columns of A . But then taking dimensions on both sides yields $\text{rank } A = \text{rank } C$, since the rank of the dimension of the column space.

\Leftarrow :

Suppose $\text{rank } A = \text{rank } C$; then the

$$\dim \text{span}(\{\mathbf{a}_i\}) = \dim \text{span}(\{\mathbf{a}_i\} \cup \{\mathbf{b}\}),$$

which says that \mathbf{b}_i is in the column space of A , and thus $AX = B$ has a solution. \square

2. The solution is unique $\iff \text{rank}(A) = m$.

\implies : To the contrapositive, Suppose $\text{rank}(A) < m$. Then by rank-nullity, $\dim \ker A > 0$, so there is a vector $\mathbf{v} \neq \mathbf{0}$ such that $A\mathbf{v} = \mathbf{0}$. But noting that $\mathbf{x} = \mathbf{0}$ is always a solution to $A\mathbf{x} = \mathbf{0}$, this yields two distinct solutions.

\Longleftarrow :

Suppose that $\text{rank}(A) = m$. Then by rank-nullity, $\dim \ker A = 0$, so $\ker A = \{\mathbf{0}\}$. Now suppose $\mathbf{v}_1, \mathbf{v}_2$ are potentially distinct solutions to $A\mathbf{x} = \mathbf{b}$.

Then,

$$\begin{aligned} A\mathbf{v}_1 &= A\mathbf{v}_2 = \mathbf{b} \\ \implies A\mathbf{v}_1 - A\mathbf{v}_2 &= \mathbf{b} - \mathbf{b} = \mathbf{0} \\ \implies A(\mathbf{v}_1 - \mathbf{v}_2) &= \mathbf{0} \\ \implies \mathbf{v}_1 - \mathbf{v}_2 &\in \ker A \\ \implies \mathbf{v}_1 - \mathbf{v}_2 &= \mathbf{0} \\ \implies \mathbf{v}_1 &= \mathbf{v}_2, \end{aligned}$$

which shows that any solution is unique.

5.4 Part 4

We want to show that $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\iff \text{rank}(A) < m$.

\implies : Suppose $A\mathbf{v} = \mathbf{0}$ for some $\mathbf{v} \neq \mathbf{0}$. Then $\dim \ker A \geq 1$, and by rank nullity we must have $m = \dim \ker A + \text{rank}(A)$. But this immediately forces $\text{rank}(A) \leq m - 1$.

\Longleftarrow : Suppose $\text{rank}(A) < m$. Then again by rank nullity, this forces $\dim \ker A \geq 1$, so A has a nontrivial kernel and thus there is a nontrivial solution to $A\mathbf{x} = \mathbf{0}$.

6 Problem 6

Proof following <http://sierra.nmsu.edu/morandi/notes/SmithNormalForm.pdf>

The goal is to show that any matrix $A \in M(m \times n, R)$ is *equivalent* to a matrix D of the described form, so $A = PDQ$ for some matrices P, Q . Since S is in fact the set of Smith Normal Forms for such matrices, it suffices to show that $SNF(A)$ can be obtained by left and right multiplication by invertible matrices. Moreover, since row operations can be performed by left-multiplication of elementary matrices, and column operations by right-multiplication.

We proceed by induction on $m + n$.

For the base case $m + n = 2$, this can only yield a 1×1 matrix, and the result holds vacuously.

For the inductive step, we will proceed by considering the top-left 2×2 block, say $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

and showing it can be reduced to a block of the form $M' = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ where $d_1 \mid d_2$. Then the sub-matrix obtained by deleting the row and column containing d_1 is a strictly smaller matrix, allowing the inductive hypothesis to be applied.

Moreover, note that if we are able to perform this reduction by a series of left and right multiplications, this will yield $A_1 = P_1 A Q_1$, and inductively we will have $A_r = (P_r \cdots P_2 P_1) A (Q_1 Q_2 \cdots Q_r)$, so each matrix will remain equivalent at every step.

Note: since R is a PID, it is also a Euclidean domain, so we can compute greatest common divisors.

We'll first reduce the top-left entry and eliminate the bottom-left entry.

Let $d = \gcd(a, c)$, so we can write $d = sa + tc$ for some $s, t \in R$. We would like to construct an operation that replaces a in M with d .

So let ℓ_1, ℓ_2 be parameters to be determined; we can then compute

$$P_1 A = \begin{bmatrix} s & t \\ \ell_1 & \ell_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} d & sb + td \\ \ell_1 a + \ell_2 c & \ell_1 b + \ell_2 d \end{bmatrix},$$

where we now only have to choose ℓ_1, ℓ_2 so that P_1 is invertible.

This lets us engineer an inverse matrix

$$\begin{aligned} P_1^{-1} &:= \begin{bmatrix} \ell_2 & -t \\ -\ell_1 & s \end{bmatrix} \\ \implies P_1 P_1^{-1} &= \begin{bmatrix} s & t \\ \ell_1 & \ell_2 \end{bmatrix} \begin{bmatrix} \ell_2 & -t \\ -\ell_1 & s \end{bmatrix} \\ &= \begin{bmatrix} s\ell_2 - t\ell_1 & -ts + st \\ \ell_1\ell_2 - \ell_2\ell_1 & -t\ell_1 + s\ell_2 \end{bmatrix}, \end{aligned}$$

which just says that we need to pick ℓ_1, ℓ_2 such that $s\ell_1 - t\ell_2 = 1$, since the off-diagonal entries vanish because R is commutative.

But this can be done by writing $a = dk_1$ and $c = dk_2$, since d was their gcd, then

$$d = sa + tc = sdk_1 + tdk_2 \implies 1 = sk_1 + tk_2,$$

so just choose $\ell_1 = k_1, \ell_2 = -k_2$ to yield $P_1 P_1^{-1} = I_2$.

We can observe that in the matrix $P_1 A$, since d divides a and c , d also divides $\ell_1 a + \ell_2 c$. So write $k_1 d = \ell_1 a + \ell_2 c$, we can then perform a row operation by left-multiplying:

$$Q_1 P_1 A := \begin{bmatrix} 1 & 0 \\ -k & 1 \end{bmatrix} \begin{bmatrix} d & sb + td \\ \ell_1 a + \ell_2 c & \ell_1 b + \ell_2 d \end{bmatrix} = \begin{bmatrix} d & sb + td \\ 0 & -k(sb + td) + \ell_1 b + \ell_2 d \end{bmatrix}.$$

We now carry out the same process with the top *row* instead of the first *column*. This begins by computing $d_1 = \gcd(d, sb + td)$, where we can immediately note that d_1 divides d .

We then write

$$d_1 = ds' + (sb + td)t',$$

then perform column operations (i.e. right-multiplying by some R_1) to obtain a matrix of the form

$$Q_1 P_1 A R_1 := \begin{bmatrix} d & sb + td \\ 0 & -k(sb + td) + \ell_1 b + \ell_1 d \end{bmatrix} \begin{bmatrix} s' & \ell_3 \\ t' & \ell_4 \end{bmatrix} = \begin{bmatrix} d_1 & d\ell_3 + (sb + td)\ell_4 \\ ? & ? \end{bmatrix}$$

where again ℓ_3, ℓ_4 are parameters that can be chosen to make R_1 invertible.

We can again observe that d_1 divides the top-left and (now) the top-right entry, which means we can find a k' such that

$$Q_1 P_1 A R_1 S_1 := \begin{bmatrix} d_1 & d\ell_3 + (sb + td)\ell_4 \\ ? & ? \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -k' & 1 \end{bmatrix} = \begin{bmatrix} d_1 & 0 \\ ? & ? \end{bmatrix},$$

which puts us back in the original situation.

We can then continue by obtaining a d_2 that divides d_1 , doing row operations, and obtaining a matrix of the form

$$P_2 Q_1 P_1 A R_1 S_1 := \begin{bmatrix} d_2 & ? \\ 0 & ? \end{bmatrix},$$

and so on.

In a PID, “to divide is to contain” for ideals, so this generates a sequence of ideals

$$(d) \subseteq (d_1) \subseteq (d_2) \subseteq \cdots$$

and since every PID is Noetherian, this increasing chain of ideals eventually stabilizes.

This means that after finitely many steps, we find $d_{N+1} := \gcd(d_N, \dots) = d_N$,

obtain a matrix

$$N := \left(\prod_i Q_i P_i \right) A \left(\prod_i R_i S_i \right) = \begin{bmatrix} d_N & x \\ y & z \end{bmatrix}$$

where either

- $x = 0$ and y divides d_N , or
- $y = 0$ and x divides d_N .

Without loss of generality, supposing the first case holds, we can write $d_N = \alpha y$; then

$$EN := \begin{bmatrix} 1 & 0 \\ 1 & -\alpha \end{bmatrix} \begin{bmatrix} d_N & 0 \\ y & z \end{bmatrix} = \begin{bmatrix} d_N & 0 \\ 0 & z \end{bmatrix},$$

where E is again invertible, yielding a diagonal matrix.

Note: in the general case of an $m \times n$ matrix, this eliminates entries 1, 2 and 2, 1. Eliminating the remaining entries in row 1 and column 1 proceed similarly, and never perturb entries that were made zero in a previous step.

Since it is not necessarily the case that d_N divides z here, a small additional modification is needed. This is accomplished by a series of row operations, as described here:

Moreover, write $a = d\alpha$ and $b = d\beta$ for some $\alpha, \beta \in R$. We then perform the following row and column operations, yielding

$$\begin{aligned} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} &\longrightarrow \begin{pmatrix} a & 0 \\ ax & b \end{pmatrix} \longrightarrow \begin{pmatrix} a & 0 \\ ax + by & b \end{pmatrix} = \begin{pmatrix} a & 0 \\ d & b \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} 0 & -b\alpha \\ d & b \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -b\alpha \\ d & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} d & 0 \\ 0 & -b\alpha \end{pmatrix}, \end{aligned}$$

a diagonal matrix in Smith normal form since d divides $-b\alpha$. □

This yields the desired form in the top-left 2×2 block, zeroing out the first column and row, so the inductive hypothesis applies to the remaining block. □