4. Matrices, matrix multiplication, and row reduction

We have seen that we can make our lives simpler by dividing number systems into two parts: the particular part belonging to that number system and the general properties that make it a number system in the first place. Anything we prove using only the general properties of a ring, say, applies to *all* rings. In this way, we don't have to repeat ourselves in each ring, or have a profusion of concepts – one for each individual ring. This is the basic motivation for abstraction in mathematics.

There's a similar unification involving row reductions which and matrix multiplication. To see it, notice that we have multiple operations we can perform on matrices to get new matrices: we can add rows, or multiply rows by numbers, or interchange rows, but we can also multiply matrices. Right now, these are separate types of operations. However, all the row operations can actually be implemented using matrix multiplication. Thus, matrix multiplication is all we actually need.

First, we formally define a matrix:

Definition 11. *An $n\times m$ matrix with entries in a set S is a function $\mathbf{A}\colon \{1,\ldots,n\}\times\{1,\ldots,m\} \longrightarrow S$. The value of $\mathbf{A}(i,j)$ is called the* ijth*-entry of the matrix, and will usually be denoted as a_{ij}*[5]*. The set of all $n \times m$ matrices with values in S will be denoted $M_{n\times m}(S)$.*

The rectangular arrays of numbers we usually draw are *not* are visual depictions of these functions and the operations we perform on the matrix.

Important Aside: We will often think of elements of $M_{1\times m}(S)$ as ordered subsets of S. Recall that a subset $\{s_1,\ldots,s_m\} \subset S$ has no order and no repetitions. Thus, $\{2,3\} = \{3,2\} = \{3,2,2\}$. However, like matrices, $[s_1,\ldots,s_m]$ will mean an *ordered* set: i.e. where there is a first element, a second element, etc. and that location matters: thus, $[2,3] \neq [3,2]$ since the order is different, and $[3,2] \neq [3,2,2]$ since they have different lengths. We et both, by thinking of ordered sets as functions $O\colon \{1,\ldots k\} \longrightarrow S$[6]

When S is a ring, R, we can add algebraic structure to the matrices in $M_{n\times m}(R)$. For instance, we can add matrices of the same size, and define matrix multiplication. Let $\mathbf{A} \in M_{n\times m}(R)$ and $\mathbf{B} \in M_{m\times l}(R)$. Then the matrix product $\mathbf{AB}$ is the element $\mathbf{C} \in M_{n\times l}(R)$ defined by

$$c_{ij} = \sum_{r=1}^{m} a_{ir}b_{rj}$$

for each $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,l\}$.

(Question 4.1) Why do we only need a ring here, and not a field?

The depiction of matrix multiplication using rectangular arrays of numbers is the usual "row times column" operation of elementary linear algebra. That is, focusing just on the one element c_{ij}:

$$\begin{bmatrix} & & \\ & c_{ij} & \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ & \sum_{r=1}^{m} a_{ir}b_{rj} & \\ & & \end{bmatrix} = \begin{bmatrix} & & & \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ & & & \end{bmatrix} \begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{bmatrix}$$

[5]using the lowercase version of the same letter as the matrix, without bold

[6]Alternately, we can think of $M_{1\times m}(S)$ as $S\times S\times\cdots\times S$, where the product has m factors, but it will be more convenient to think of ordered lists as matrices.

(Question 4.2) Consider the following matrices:

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Now let $\mathbf{A}$ be a $3 \times m$ matrix over a field $\mathbb{F}$.

(1) Show that $\mathbf{DA}$ is the result of multiplying the third row of $\mathbf{A}$ by 2. In terms of the rows, what happens when you compute $\mathbf{SA}$ and $\mathbf{EA}$?

(2) Suppose $\mathbf{B}$ is an $n \times 3$ matrix. Can you say something similar about $\mathbf{BD}$, $\mathbf{BS}$, and $\mathbf{BE}$ (although not necessarily using rows)?

Let R be a ring. Then we can define several types of matrices in $M_{n\times n}(R)$:

I: For $u \in R^\times$, and $l \in \{1, \dots, n\}$, let $\mathbf{D}_l(u)$ be the matrix defined by

$$d_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \text{ and } i \neq l \\ u & i = j = l \end{cases}$$

For example, The matrix $\mathbf{D}$ in the question above is $\mathbf{D}_3(2)$ in $M_{3\times 3}(R)$.

II: For $l, m \in \{1, \dots, n\}$ with $l < m$, let $\mathbf{S}_{lm}$ be the matrix defined by

$$s_{ij} = \begin{cases} 1 & i = j \text{ and } i \neq l, m \\ 1 & \{i, j\} = \{l, m\} \\ 0 & \text{otherwise} \end{cases}$$

For example, the matrix $\mathbf{S}$ in the question above is $\mathbf{S}_{13}$ in $M_{3\times 3}(R)$.

III: For $l, m \in \{1, \dots, n\}$ with $l \neq m$, and $\lambda \in R$, let $\mathbf{E}_{lm}(\lambda)$ be the matrix defined by

$$e_{ij} = \begin{cases} 1 & i = j \\ \lambda & i = l, j = m \\ 0 & otherwise \end{cases}$$

For example, The matrix $\mathbf{E}$ in the question above is $\mathbf{E}_{31}(-2)$ in $M_{3\times 3}(R)$.

We will call the matrices $\mathbf{D}_l(u)$, $\mathbf{S}_{lm}$, and $\mathbf{E}_{lm}(\lambda)$ the *elementary matrices* in $M_{n\times n}(R)$. However, in some texts, only matrices of the form $\mathbf{E}_{lm}(\lambda)$ are called elementary.

If $\mathbf{M}$ is an $n \times m$-matrix, then

(1) $\mathbf{D}_l(u)\mathbf{M}$ is the matrix which results from multiplying the l^{th} row of $\mathbf{M}$ by u.

(2) $\mathbf{S}_{lm}\mathbf{M}$ is the matrix which results from swapping the l^{th} and m^{th} rows of $\mathbf{M}$.

(3) $\mathbf{E}_{lm}(\lambda)\mathbf{M}$ is the matrix which results from replacing the l^{th} row with the sum of the original l^{th} row with λ times the m^{th}.

(Question 4.3) Let $\mathbf{M}$ be an $n \times m$ matrix and suppose our elementary matrices are in $M_{m\times m}(R)$. How can we describe the result of computing $\mathbf{MD}_l(u)$, $\mathbf{MS}_{lm}$ and $\mathbf{ME}_{lm}(\lambda)$ in terms of $\mathbf{M}$?

Over any ring R, the elementary matrices are invertible since the corresponding row operation is reversible. In fact:

(1) $\mathbf{D}_l(u)^{-1} = \mathbf{D}_l(u^{-1})$
(2) $\mathbf{S}_{lm}^{-1} = \mathbf{S}_{lm}$, and
(3) $\mathbf{E}_{lm}(\lambda)^{-1} = \mathbf{E}_{lm}(-\lambda)$

Note that *the inverses are also elementary matrices*!

(Question 4.4) Explain these formulas in terms of row operations.

(Question 4.5) In $M_{n\times n}(R)$, show that for each $u \in R^\times$, $\mathbf{E}_{lm}(u) = \mathbf{D}_m(u^{-1})\mathbf{E}_{lm}(1)\mathbf{D}_m(u)$. Explain why, when $R = \mathbb{F}$ is a field, we only need to use the matrices $\mathbf{E}_{lm}(1)$'s to implement all the row operations.

We can us the correspondence between elementary matrices and row operations to understand Gaussian elimination: if $\mathbf{M}$ is an elementary matrix, the corresponding row operation on the augmented matrix for a linear system is

$$\mathbf{A}\mathbf{x} = \mathbf{b} \qquad \Longrightarrow \qquad \big(\mathbf{M}\mathbf{A}\big)\mathbf{x} = \mathbf{M}\mathbf{b}$$

Thus, we can repeatedly multiply by elementary matrices to try to get:

$$\mathbf{R}\mathbf{x} = \mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_k\mathbf{b}$$

where $\mathbf{R}$ is in RREF.

AS we have seen, row operations over a ring R do not necessarily suffice to get to $\mathbf{R}$ (we saw an example when $R = \mathbb{Z}$). However, when R is a field, Gaussian elimination allows us to reduce any $n \times m$ matrix $\mathbf{M}$ to an RREF matrix $\mathbf{R}$ (of the same shape). Therefore,

Theorem 12. *Let $\mathbb{F}$ be a field and $\mathbf{A} \in M_{n\times m}(\mathbb{F})$. Then there exist elementary matrices $\mathbf{M}_1, \ldots, \mathbf{M}_k \in M_{n\times n}(\mathbb{F})$ such that*

$$\mathbf{R} = \mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_k\mathbf{A}$$

is in row reduced echelon form.

Consequently, this theorem allows us to use matrix multiplication to describe, and reason about, Gaussian elimination. This theorem explains most of the computational algorithms in elementary linear algebra.

(Question 4.6) Suppose $\mathbf{A}, \mathbf{B} \in M_{n\times n}(R)$ and A and B are invertible. Prove that $\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the inverse for $\mathbf{AB}$.

(Question 4.7) Suppose $\mathbf{B}_i$ is invertible for all $i = 1, \ldots n$ for some $n \in \mathbb{N}$. Prove that $\mathbf{A} = \mathbf{B}_1\mathbf{B}_2\cdots\mathbf{B}_n$ is invertible and find its inverse. (*Hint:* Use induction!)

Definition 13. *Let R be a commutative ring. We will denote the set of invertible matrices in $M_{n\times n}(R)$ by $GL_n(R)$, the general linear group of $n \times n$ matrices over R.*

We have just seen that the product of matrices in $GL_n(R)$ is also in $GL_n(R)$, and if $\mathbf{A} \in GL_n(R)$ then $\mathbf{A}^{-1} \in GL_n(R)$. FUrthermore, all the elementary matrices are in $GL_n(R)$, and thus any matrix that is a product of elementary matrices is invertible over any ring R.

However, over a field $\mathbb{F}$, any $n \times n$ matrix that is invertible can be transformed via row operations into the $n \times n$ identity matrix. Thus, by the theorem above, there are elementary matrices $\mathbf{M}_i$ such that

$$\mathbf{M}_1\mathbf{M}_2 \cdots \mathbf{M}_k\mathbf{A} = \mathbf{I}_n$$

Left multiplying by $\mathbf{M}_1^{-1}$, then by $\mathbf{M}_2^{-1}$, etc. shows that

$$\mathbf{A} = \mathbf{M}_k^{-1}\mathbf{M}_{k-1}^{-1} \cdots \mathbf{M}_2^{-1}\mathbf{M}_1^{-1}$$

(Question 4.8) Verify this last identity. How can you handle not knowing what n is?

Each $\mathbf{M}_i^{-1}$ is an elementary matrix, so we see that an invertible matrix $\mathbf{A}$ is a product of elementary matrices. Together we get

Theorem 14. *A matrix $\mathbf{A} \in M_{n\times n}(\mathbb{F})$ is invertible if and only if it can be factored as a product of elementary matrices.*

(Question 4.9) Suppose $\mathbf{M}_1\mathbf{M}_2 \cdots \mathbf{M}_k\mathbf{A} = \mathbf{I}_n$ with $\mathbf{M}_i$ elementary.

(1) Explain why $\mathbf{A}^{-1} = \mathbf{M}_1\mathbf{M}_2 \cdots \mathbf{M}_k$.
(2) Technically, we have only found $\mathbf{A}^{-1}$ so that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$. Prove that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$ as well.
(3) Many elementary linear algebra courses present the following method for computing the inverse of an invertible $n \times n$ matrix $\mathbf{A}$:
  (a) Form the $n \times 2n$ matrix $\mathbf{G} = \left[\mathbf{A}\middle|\mathbf{I}_n\right]$
  (b) Perform row operations on $\mathbf{G}$ to get a matrix of the form $\left[\mathbf{I}_n\middle|\mathbf{B}\right]$ (which is possible since $\mathbf{A}$ is invertible)
  (c) Then $\mathbf{A}^{-1} = \mathbf{B}$

  Use the connection between row operations and elementary matrices to justify the conclusion that $\mathbf{B}$ is the inverse.

(Question 4.10) Suppose $\mathbf{M}$ is a product of elementary matrices of the form $\mathbf{E}_{ij}$ where $i > j$. Show that $\mathbf{M}$ is invertible, and $\mathbf{M}$ and $\mathbf{M}^{-1}$ are lower triangular with 1's on the diagonal. Now suppose that we have a matrix $\mathbf{A} \in M_{n\times n}(\mathbb{F})$ where we can obtain an upper triangular matrix $\mathbf{U}$ by performing row operations where we only ever add multiples of a row to a later row: that is $R_j + \lambda R_i \longrightarrow R_j$ for $j > i$. Show that $\mathbf{A} = \mathbf{L}\mathbf{U}$ where $\mathbf{L}$ is lower triangular.