In this handout we explore a striking relationship between two concepts: elementary row operations and matrix multiplication. This relationship yields a new way to interpret the process of solving a system of linear equations, and will be fundamental for understanding much of we will study in the rest of the course.

In this handout, you will learn:

- How to express an elementary row operation as a matrix multiplication.
- How to solve linear systems using matrix multiplication.
- The notion of inverse of a square matrix, and how to compute it.

1 Matrix Multiplication

The goal of this section is to develop and "algebraic" interpretation of elementary row operations. The main result we will get is that EROs are equivalent to multiplication by a certain kind of matrix. This fundamental insight will be key to understanding several concepts in the course.

We start with a review of matrix multiplication. Matrix multiplication can be reduced to the basic case of multiplying a *row vector* by a *column vector*, where both vectors must have the same number of entries:

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{j=1}^{n} v_j w_j$$

For example:

$$\begin{bmatrix} -2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = -2 \times 3 + 2 \times 1 + 3 \times 2 = 2$$

Notice that the result of the multiplication is a scalar (real number). This way of multiplying vectors is also known as *scalar multiplication* (because the result is a scalar) or *dot product* (because it is sometimes represented by a "dot").

General matrix multiplication can be interpreted in terms of this particular case as follows

- Let A and B be two matrices, where A is $m \times n$ and B is $n \times p$. Notice that *the number of columns of A is equal to the number of rows of B.*
- Let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_p$ denote the *rows* of A and $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_p$ denote the *columns* of B:

$$A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} \qquad B = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_p \end{bmatrix}$$

- Each entry of the product AB is computed by multiplying the corresponding row of A and column of B:

$$AB = \begin{bmatrix} \mathbf{r}_1\mathbf{c}_1 & \mathbf{r}_1\mathbf{c}_2 & \cdots & \mathbf{r}_1\mathbf{c}_p \\ \mathbf{r}_2\mathbf{c}_1 & \mathbf{r}_2\mathbf{c}_2 & \cdots & \mathbf{r}_2\mathbf{c}_p \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_m\mathbf{c}_1 & \mathbf{r}_m\mathbf{c}_2 & \cdots & \mathbf{r}_m\mathbf{c}_p \end{bmatrix}$$

  Notice that matrix AB has dimensions $m \times p$

For example, lets consider the product:

$$AB = \begin{bmatrix} 2 & -3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 4 & -2 & 5 \\ -3 & 2 & 0 \end{bmatrix}$$

To compute entry $(1,2)$ of the product we do:

$$(\text{Row 1 of } A) \times (\text{Column 2 of } B) = \begin{bmatrix} 2 & -3 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = 2 \times (-2) + (-3) \times 2 = -10$$

Doing the analogous computation for all entries we get:

$$\begin{aligned} \begin{bmatrix} 2 & -3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 4 & -2 & 5 \\ -3 & 2 & 0 \end{bmatrix} &= \begin{bmatrix} 2 \times 4 + (-3) \times (-3) & 2 \times (-2) + (-3) \times 2 & 2 \times 5 + (-3) \times 0 \\ 4 \times 4 + 2 \times (-3) & 4 \times (-2) + 2 \times 2 & 4 \times 5 + 2 \times 0 \end{bmatrix} \\ &= \begin{bmatrix} 17 & -10 & 10 \\ 10 & -4 & 20 \end{bmatrix} \end{aligned}$$

Recall that the $n \times n$ *identity matrix* is the $n \times n$ matrix with ones on the diagonal and zeros everywhere else:

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots 0 & \\ 0 & 1 & 0 & \cdots 0 & \\ 0 & 0 & 1 & \cdots 0 & \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots 1 & \end{bmatrix}$$

2 Matrix Multiplication and EROs

The goal of this section is to explore the relationship between elementary row operations and matrix multiplication. The examples are in terms of a 3×3 matrix, but the results are valid for matrices with arbitrary dimensions.

Let's consider the matrix:

$$A = \begin{bmatrix} a & b & c \\ d & f & g \\ h & i & j \end{bmatrix}$$

Now, suppose that we perform the ERO **R1** $*\ (\lambda)$ + **R2** => **R2** to this matrix:

$$\begin{bmatrix} a & b & c \\ d & f & g \\ h & i & j \end{bmatrix} \overset{\mathbf{R1}*(\lambda)+\mathbf{R2}=>\mathbf{R2}}{\sim} \begin{bmatrix} a & b & c \\ \lambda a+d & \lambda b+f & \lambda c+g \\ h & i & j \end{bmatrix} \tag{2.1}$$

Let's now consider the following sequence of operations:

1. Perform the ROP **R1** $*\ (\lambda)$ + **R2** => **R2** on the 3×3 identity matrix, calling the result E:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{\mathbf{R1}*(\lambda)+\mathbf{R2}=>\mathbf{R2}}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Multiply matrix E by matrix A (in this order!):

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & f & g \\ h & i & j \end{bmatrix} = \begin{bmatrix} a & b & c \\ \lambda a+d & \lambda b+f & \lambda c+g \\ h & i & j \end{bmatrix} \tag{2.2}$$

Notice that the resulting matrices in (2.1) and (2.2) are the same! By doing more experiments, we can see that the same is true for other EROs. Based on this, we make the following definition:

Definition 2.1. An *elementary matrix* is a matrix obtained by applying an elementary row operation to an identity matrix. We say that the resulting matrix *corresponds* to the applied ERO.

We then have the following basic principle:

Proposition 2.2. *Suppose that A is a $m \times n$ matrix and we are given an ERO. Let E be the elementary matrix corresponding to this ERO. Then:*

$$\textit{Result of applying the ERO to } A = EA$$

We do not present a proof of the proposition, but, before we continue, let's observe the following alternative characterization of the elementary matrices:

ERO	**Elementary Matrix**
Ri $*\ (\lambda)$ + **Rj** => **Rj**	Identity matrix with the 0 at position (j,i) replaced by λ.
Ri $*\ (\lambda)$ => **Ri**	Identity matrix with the 1 at position (i,i) replaced by λ.
Ri <=> **Rj**	Identity matrix with rows i and j swapped.

3 Solving Systems and Matrix Multiplication

Let's now put together the observation of the previous section with the process of Gaussian Elimination. Recall that general system of linear equations can always be written as

$$A\mathbf{x} = \mathbf{v}$$

In this section, we will consider the important case where A is a *square matrix*, that is, it has the same number of rows and columns. For concreteness, lets consider the case where A is the 3×3 matrix of of the system:

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

For the sake of generality, we will let $\mathbf{v}$ be a generic vector:

$$\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Let's now perform a sequence of row operation to find the RREF of A. Of course, algorithmically this must be done step by step, but we summarize the whole process in a single display as follows:

$$\begin{bmatrix} 2 & 1 & -2 \\ 2 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \overset{\substack{\mathbf{R1*(1/2)=>R1} \\ \mathbf{R1*(-2)+R2=>R2} \\ \mathbf{R1*(-1)+R3=>R3} \\ \mathbf{R2*(-1/2)=>R2} \\ \mathbf{R2*(-1/2)+R1=>R1} \\ \mathbf{R2*(-1/2)+R3=>R3} \\ \mathbf{R3*(1/3)=>R3} \\ \mathbf{R3*(2)+R2=>R2}}}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Notice that the RREF of matrix A is the identity matrix. This is important, because, in this case the solution of the system can be found by *applying the same sequence of row operations to the vector* $\mathbf{v}$:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \overset{\substack{\mathbf{R1*(1/2)=>R1} \\ \mathbf{R1*(-2)+R2=>R2} \\ \mathbf{R1*(-1)+R3=>R3} \\ \mathbf{R2*(-1/2)=>R2} \\ \mathbf{R2*(-1/2)+R1=>R1} \\ \mathbf{R2*(-1/2)+R3=>R3} \\ \mathbf{R3*(1/3)=>R3} \\ \mathbf{R3*(2)+R2=>R2}}}{\sim} \begin{bmatrix} \frac{a}{4} + \frac{b}{4} \\ -\frac{b}{3} + \frac{2c}{3} \\ -\frac{a}{4} + \frac{b}{12} + \frac{c}{3} \end{bmatrix}$$

Notice that the solution can be written as a matrix multiplication:

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{4} & \frac{1}{12} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{a}{4} + \frac{b}{4} \\ -\frac{b}{3} + \frac{2c}{3} \\ -\frac{a}{4} + \frac{b}{12} + \frac{c}{3} \end{bmatrix}$$

Let's now interpret what we have done in terms of matrix multiplication. Each row operation can be associated to an elementary matrix:

ERO	**Elementary Matrix**
R1 ∗ (1/2) => R1	$E_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
R1 ∗ (−2) + R2 => R2	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
R1 ∗ (−1) + R3 => R3	$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
R2 ∗ (−1/2) => R2	$E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$
R2 ∗ (−1/2) + R1 => R1	$E_5 = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
R2 ∗ (−1/2) + R3 => R3	$E_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$
R3 ∗ (1/3) => R3	$E_7 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$
R3 ∗ (2) + R2 => R2	$E_8 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

Then, the process of applying row operations can be interpreted as multiplication by a succession of matrices:

$$E_8E_7E_6E_5E_4E_3E_2E_1A = I_3$$

The solution of the system is the result of the application of the same matrix multiplication to $\mathbf{v}$

$$\mathbf{x} = E_8E_7E_6E_5E_4E_3E_2E_1\mathbf{v}$$

This whole process can be greatly simplified with the following observation:

The matrix

$$E = E_8E_7E_6E_5E_4E_3E_2E_1$$

L. Felipe Martins (l.martins@csuohio.edu)

can be obtained by applying the sequence corresponding row operations to the identity matrix I_3.

We then have the following recipe to solve the system

$$A\mathbf{x} = \mathbf{b}$$

in the case where A is $n \times n$:

1. Find a sequence of EROs that reduces A to the identity matrix (if this is not possible, the method can't be used).
2. Let E be the matrix obtained by performing the same sequence of row operations to the identity matrix, I_n.
3. Then, the system has a unique solution, given by:

$$\mathbf{x} = E\mathbf{v}$$

4 The Inverse of a Matrix

We can reinterpret the procedure from the previous section in an even more efficient way. To describe the new algorithm, let's consider the 4×4 matrix:

$$\begin{bmatrix} 0 & 4 & -1 & 0 \\ 0 & -2 & -2 & \frac{1}{3} \\ -1 & 1 & 0 & \frac{1}{12} \\ 2 & -1 & 3 & 0 \end{bmatrix}$$

Let's now augment the matrix A by attaching to it a copy of the 4×4 identity matrix, I_4:

$$\begin{bmatrix} 0 & 4 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & -2 & \frac{1}{3} & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & \frac{1}{12} & 0 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

We now perform a sequence of EROs to A, and find its RREF:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ 0 & 1 & 0 & 0 & \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ 0 & 0 & 1 & 0 & -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ 0 & 0 & 0 & 1 & \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix}$$

The RREF of A is the identity matrix. Thus, if we let E be the matrix resulting from applying the same EROs to the identity we get:

$$E = \begin{bmatrix} \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix}$$

Multiplying this matrix by A gives the identity:

$$EA = \begin{bmatrix} \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix} \begin{bmatrix} 0 & 4 & -1 & 0 \\ 0 & -2 & -2 & \frac{1}{3} \\ -1 & 1 & 0 & \frac{1}{12} \\ 2 & -1 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Notice that if we compute the product in reversed order we also get the identity:

$$AE = \begin{bmatrix} 0 & 4 & -1 & 0 \\ 0 & -2 & -2 & \frac{1}{3} \\ -1 & 1 & 0 & \frac{1}{12} \\ 2 & -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The fact that the reversed product AE is also the identity will be proved later in the course. We say that the matrix E obtained above is the *inverse* of the matrix A, characterized algebraically by:

$$EA = I = AE.$$

We can summarize the process of finding the inverse of a square matrix A as follows:

1. Augment the matrix A by appending the identity matrix with the same dimensions:

$$M = [A|I_n]$$

2. Perform EROs to bring A to RREF.
3. If the RREF of A is the identity matrix, the inverse of A will appear on the left of the resulting matrix, in the place where the identity.

If the RREF of A is not the identity, A is *not invertible*, that is it does not have an inverse.

5 Solving Systems Using the Inverse

Let's now go back to the system:

$$A\mathbf{x} = \mathbf{v}$$

Suppose that A is invertible, and let E be the inverse of A. Recall that we have the following dual interpretation for E:

- E is obtained by performing to the identity matrix the same sequence of EROs that transform A into the identity.
- $EA = I = AE$

The solution of the system is then:

$$\mathbf{x} = E\mathbf{v}$$

There are two ways to verify this:

- To find the solution, we must perform on $\mathbf{v}$ the same sequence of EROs that reduce A to the identity. This is equivalent to computing $E\mathbf{v}$
- Multiplying $A\mathbf{x} = \mathbf{v}$ by E (to the left) gives:

$$EA\mathbf{x} = E\mathbf{v}$$

Now, $EA\mathbf{x} = I\mathbf{x} = \mathbf{x}$, so we get:

$$\mathbf{x} = E\mathbf{v}$$

For example, let's solve the system:

$$\begin{bmatrix} 0 & 4 & -1 & 0 \\ 0 & -2 & -2 & \frac{1}{3} \\ -1 & 1 & 0 & \frac{1}{12} \\ 2 & -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 5 \\ -1 \end{bmatrix}$$

We compute the inverse of the matrix above in the previous section:

$$E = \begin{bmatrix} \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix}$$

So, we obtain the solution by computing $E\mathbf{v}$:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{5}{18} & \frac{11}{72} & -\frac{11}{18} & \frac{7}{36} \\ \frac{2}{9} & -\frac{1}{36} & \frac{1}{9} & \frac{1}{18} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{2}{3} & \frac{13}{6} & \frac{10}{3} & \frac{5}{3} \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 5 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{227}{72} \\ \frac{37}{36} \\ \frac{19}{9} \\ \frac{59}{6} \end{bmatrix}$$

As usual, we can verify that this is a solution by computing:

$$Ax = \begin{bmatrix} 0 & 4 & -1 & 0 \\ 0 & -2 & -2 & \frac{1}{3} \\ -1 & 1 & 0 & \frac{1}{12} \\ 2 & -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} -\frac{227}{72} \\ \frac{37}{36} \\ \frac{19}{9} \\ \frac{59}{6} \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 5 \\ -1 \end{bmatrix} = \mathbf{v}$$

6 Notation

We have been using the letter E to denote the inverse of a matrix A. The most common notation for the inverse is A^{-1}. So, we can write the defining algebraic property of the inverse of A by:

$$AA^{-1} = I = A^{-1}A$$

Also notice that not every matrix is invertible. In this case we say that "A^{-1} does not exist".

7 Summary

Here is what we learned in this handout:

- To every ERO we associate an elementary matrix, obtained by applying the ERO to the identity matrix.
- Applying a sequence of EROs to a matrix A is equivalent to left-multiplying the matrix A by the product of the corresponding elementary matrices.
- To find the inverse of a square matrix A, we form the matrix

$$M = [A|I]$$

  Then, apply a sequence of EROs to M to reduce A to the identity matrix. If this is possible, then the matrix A^{-1} appears in the place originally occupied by I in M.
- The inverse of a matrix A is characterized by:

$$AA^{-1} = I = A^{-1}A$$

- If A is a square matrix and is invertible, the solution of the system

$$A\mathbf{x} = \mathbf{v}$$

  is unique, and is given by:

$$\mathbf{x} = A^{-1}\mathbf{v}.$$