Week 2
Gaussian
Elimination

> **Exercise.**
>
> If C is an $m \times n$ matrix corresponding to a linear map $\mathcal{C} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, and D is a $n \times l$ matrix corresponding to a linear map $\mathcal{D} : \mathbb{R}^l \longrightarrow \mathbb{R}^n$, then the product CD is the $m \times l$ matrix corresponding to the composition of linear maps:
>
> $$\mathcal{C} \circ \mathcal{D} : \mathbb{R}^l \longrightarrow \mathbb{R}^m.$$

Matrix multiplication is associative, namely:

$$(AB)C = A(BC).$$

However, in general $AB \neq BA$, even if both products are defined. This failure of commutativity is unsurprising if one views matrix multiplication in terms of the composition of linear maps (Why should $\mathcal{C} \circ \mathcal{D}$ be equal to $\mathcal{D} \circ \mathcal{C}$? They might not even have the same domain.)
The $n \times n$ matrix $I_n = I = (I_{ij})$, defined by:

$$I_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

is called the **identity matrix**. Its only nonzero entries are 1's along the diagonal:

$$I = \begin{pmatrix} 1 & 0 & 0 \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

It corresponds to the identity map from $\mathbb{R}^n$ to itself, namely, $I\vec{v} = \vec{v}$ for all $\vec{v} \in \mathbb{R}^n$. Moreover, for all $n \times m$ matrices A and $m \times n$ matrices B, we have:

$$I_n A = A, \quad BI_n = B.$$

Systems of Linear Equations

Let $x_1, x_2, \ldots x_n$ be variables. A system of linear equations in $x_1, \ldots, x_n$ is a set of equations of the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad \vdots \quad &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where the a_{ij}'s and b_i's are constants. Solving the above system means finding all values of $x_1, \ldots, x_n$ which simultaneously satisfy all the equations. In the language of matrices, the above system is equivalent to the following **matrix equation**:

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\vec{b}}$$

Hence, solving a system of linear equations is equivalent to solving the associated matrix equation:

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

for the vector $\vec{x}$.

Row Echelon Form

A matrix is said to be in **row echelon form** if:

1. The first (counting from the left) nonzero entry of each row is 1.
2. For the first nonzero entry of each row, the entry right below it is zero.
3. The number of leading zeroes (from the left) on each row is greater than or equal to the number of leading zeros in the row above.

We call the first nonzero entry of a row of a matrix in row echelon form a **pivot.** For a matrix in row echelon form, all the entries below a pivot are zero.

Example.

The following matrices are in row echelon form:

$$\begin{pmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The following matrices are *not* in row echelon form:

$$\begin{pmatrix} 2 & 3 & 4 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -5 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

If a matrix A is in row echelon form, then a matrix equation of the form $A\vec{x} = \vec{b}$ is relatively easy to solve.

Example.

Consider the matrix equation

$$\begin{pmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 7 \end{pmatrix}$$

Computing the matrix multiplication on the left-hand side, we have:

$$\begin{pmatrix} x_1 + 4x_2 + 2x_3 \\ x_2 + 3x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 7 \end{pmatrix}.$$

We see right away that $x_3 = 7$. Comparing the second row of the above vector and that of the vector on the right-hand side of the equation, we have

$$x_2 + 3x_3 = -1.$$

Since we already know that $x_3 = 7$, we deduce that :

$$x_2 = -1 - 3(7) = -22.$$

Finally, it follows from the first row that:

$$x_1 = 5 - 4x_2 - 2x_3 = 5 - 4(-22) - 2(7) = 79.$$

Gaussian Elimination

Definition.

Two matrices are said to be **row equivalent** if one can be obtained from the other via the following row operations:

I. Interchange two rows.
II. Multiply a row by a nonzero number.
III. Replace a row with its sum with a scalar multiple of another row.

Fact.

Every matrix is row equivalent to a matrix in row echelon form.

Augmented Matrix

Given a matrix equation $A\vec{x} = \vec{b}$, we define the associated **augmented matrix** as follows:

$$\left(\begin{array}{c|c} A & \vec{b} \end{array} \right) = \left(\begin{array}{cccc|c} A_{11} & A_{12} & \cdots & A_{1n} & b_1 \\ A_{21} & A_{22} & \cdots & A_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} & b_m \end{array} \right)$$

Theorem.

If two augmented matrices $\left(A \,|\, \vec{b}\right)$ and $\left(A' \,|\, \vec{b}'\right)$ are row equivalent, then $\vec{x}$ is a solution to $A\vec{x} = \vec{b}$ if and only if it is a solution to $A'\vec{x} = \vec{b}'$. In other words, the systems of linear equations associated with the two augmented matrices have the same solution set.

Hence, we may solve a matrix equation $A\vec{x} = \vec{b}$ as follows:

1. Form the augmented matrix $\left(A \,|\, \vec{b}\right)$.
2. Perform a sequence of row operations which turn $\left(A \,|\, \vec{b}\right)$ into a matrix $\left(A' \,|\, \vec{b}'\right)$ which is in row echelon form.
3. Solve the matrix equation $A'\vec{x} = \vec{b}'$ by examining $\left(A' | \vec{b}'\right)$.

The solutions to $A'\vec{x} = \vec{b}'$ are precisely those to $A\vec{x} = \vec{b}$.

Example.

Solve the matrix equation:

$$\begin{pmatrix} 0 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 7 \end{pmatrix}$$

We want to row reduce the following augmented matrix to row echelon form:

$$\left(\begin{array}{ccc|c} 0 & 2 & 1 & 2 \\ 3 & -1 & -3 & -2 \\ 2 & 3 & 1 & 7 \end{array}\right).$$

Step 1. Rearranging rows if necessary, make sure that the first nonzero entry of the first row is no further right than any of the nonzero entries in the rows below. We call this nonzero entry a **pivot.** Then, we scale the row containing the pivot so that the pivot is equal to 1.

$$\left(\begin{array}{ccc|c} 0 & 2 & 1 & 2 \\ 3 & -1 & -3 & -2 \\ 2 & 3 & 1 & 7 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & 3 & 1 & 7 \\ 3 & -1 & -3 & -2 \\ 0 & 2 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 3/2 & 1/2 & 7/2 \\ 3 & -1 & -3 & -2 \\ 0 & 2 & 1 & 2 \end{array}\right)$$

Step 2. Make all entries below the pivot zero by performing row operation III:

$$\left(\begin{array}{ccc|c} \boxed{1} & 3/2 & 1/2 & 7/2 \\ 0 & -11/2 & -9/2 & -25/2 \\ 0 & 2 & 1 & 2 \end{array}\right)$$

Here, we have added -3 times the first row to the second row. Once we have zeroed out the entries below the pivot, we ignore the first row and first column and apply Steps 1 and 2 to the smaller remaining matrix. Since $-11/2$ is nonzero, we may use it as a pivot, and scale the second row by $-2/11$:

$$\left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 7/2 \\ 0 & \boxed{1} & 9/11 & 25/11 \\ 0 & 2 & 1 & 2 \end{array}\right)$$

Then, we add -2 times the second row to the third row to obtain:

$$\left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 7/2 \\ 0 & \boxed{1} & 9/11 & 25/11 \\ 0 & 0 & -7/11 & -28/11 \end{array}\right)$$

Finally, scaling the third row by $-11/7$, we obtain the following matrix, which is in row echelon form:

$$\left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 7/2 \\ 0 & 1 & 9/11 & 25/11 \\ 0 & 0 & 1 & 4 \end{array}\right)$$

This is the augmented matrix corresponding to the matrix equation:

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ 0 & 1 & 9/11 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7/2 \\ 25/11 \\ 4 \end{pmatrix}$$

It can be easily solved: $x_3 = 4$, $x_2 = 25/11 - (9/11)4 = -1$, $x_1 = 7/2 - (3/2)(-1) - (1/2)(4) = 3$.

The process of reducing a matrix to row echelon form using row operations I, II, and III is commonly known as **Gaussian Elimination.**
In the previous example, the matrix A is row equivalent to a matrix which is in **strict triangular form**: It is a matrix in row echelon form such that its last row contains a single 1 at the right-most position. In this case, any matrix equation of the form $A\vec{x} = \vec{b}$ has a unique solution. But there are matrices in row echelon form which are not strictly triangular.

Example.

Solve the following system of linear equations:

$$\begin{aligned} 2x + 4y + 2z &= 8 \\ -3x - 6y + 2z &= -7 \\ -4x - 8y + 6z &= -6 \end{aligned}$$

This is equivalent to solving the matrix equation:

$$\begin{pmatrix} 2 & 4 & 2 \\ -3 & -6 & 2 \\ -4 & -8 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ -7 \\ -6 \end{pmatrix}$$

The associated augmented matrix is:

$$\left(\begin{array}{ccc|c} 2 & 4 & 2 & 8 \\ -3 & -6 & 2 & -7 \\ -4 & -8 & 6 & -6 \end{array}\right)$$

Which is row equivalent to:

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Notice that the entries of the last row are all zero. What does that mean? Reinterpreting the augmented matrix in terms of a matrix equation, we have:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$

This implies that $z = 1$, and $x + 2y = 3$, with no further conditions on x and y. If we introduce a free parameter $t \in \mathbb{R}$ and let $y = t$, we have:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3-2t \\ t \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Hence, the system of equations have infinitely many solutions.

Example.

Now, consider the linear system with corresponding augmented matrix:

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 2 & 2 \end{array}\right)$$

This augmented matrix is row equivalent to:

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 6 \end{array}\right)$$

Viewed in terms of the corresponding matrix equation, we have:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 6 \end{pmatrix},$$

which stipulates that $0 \cdot x_2 = 6$. We conclude that the linear system has no solution.

From the previous examples, and considering all possible row echelon forms, we see that a linear system can either have:

- **no solution,**
- **a unique solution,**
- **or infinitely many solutions.**

You will never encounter a linear system which has, say, exactly 2 solutions.

Invertible Matrices

An $n \times n$ matrix A is said to be **invertible** if there exists an $n \times n$ matrix B such that:

$$BA = AB = I_n,$$

where I_n is the $n \times n$ identity matrix. We call B an inverse of A.

> **Exercise.**
>
> If two $n \times n$ matrices B, C are both inverses of A, then $B = C$.

In other words, if A is invertible, then its inverse is unique. We denote the inverse of A by A^{-1}.

> **Exercise.**
>
> If A, B are invertible $n \times n$ matrices, then AB is invertible, with:
>
> $$(AB)^{-1} = B^{-1}A^{-1}.$$

> **Exercise.**
>
> Suppose $\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map corresponding to a matrix A. Then, $\mathcal{A}$ is invertible (as a map) if and only if A is an invertible matrix. Moreover, if $\mathcal{A}$ is invertible, then $\mathcal{A}^{-1}$ is also linear, corresponding to the matrix A^{-1}.

Why does Gaussian Elimination Give the Right Solutions?

Elementary Matrices

There are 3 types of elementary matrices.

Type I. This type is obtained by interchanging two rows of an $n \times n$ identity matrix.

Example.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

If E is an elementary matrix obtained from an identity matrix by interchanging the i-th and j-th rows, then the multiplication of a vector by E interchanges the i-th and j-th entries of the vector:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_4 \\ x_3 \\ x_2 \end{pmatrix}$$

And since a matrix can be viewed as an array of column vectors, it is easy to see that multiplication of a matrix by E from the left switches the i-th and j-th rows of the matrix.

Type II An elementary matrix of type II is obtained by multiplying a single row of the identity matrix by a *nonzero* scalar.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

If E is a obtained from the identity matrix I by multiplying the i-th row of I by λ, then multiplication of a matrix by E from the left multiplies the i-th row of the matrix by λ, and leaves all other entries unchanged.

Example.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 7x_3 \end{pmatrix}$$

Type III An elementary matrix of type **III** is obtained from the identity matrix by adding a scalar multiple of one row to another row:

Example.

$$\begin{pmatrix} 1 & 0 & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

If E is obtained from I by adding λ times the i-th row of I to the j-th row of I, then multiplying a matrix A by E from the left adds (λ times the i-th row of A) to the (j-th row of A).

$$\begin{pmatrix} 1 & 0 & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1+6(7) & 2+6(8) \\ 4 & 5 \\ 7 & 8 \end{pmatrix}$$

Exercise.

A matrix B is row equivalent to a matrix A if and only if there is a sequence of elementary matrices $E_1, E_2, \ldots E_k$, such that:

$$B = E_k \cdots E_2 E_1 A$$

Let A be an $m \times n$ matrix, and $\vec{b} \in \mathbb{R}^m$. We want to solve the matrix equation $A\vec{x} = \vec{b}$.

Exercise.

If C an $m \times m$ invertible matrix, then: $\vec{x}_0 \in \mathbb{R}^n$ is a solution to:

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

if and only if it is a solution to:

$$(CA)\vec{x} = C\vec{b}.$$

The idea here is that, if we could find an invertible C such that CA is in row echelon form, then solving $A\vec{x} = \vec{b}$ becomes much easier.

Exercise.

For any finite set of elementary matrices $E_1, E_2, \ldots, E_k$, the product $E_k \cdots E_2 E_1$ is invertible.

Suppose $E_1, E_2, \ldots E_k$ are elementary matrices such that $E_k \cdots E_2 E_1 A$ is in row echelon form. Multiplying both sides of the original equation $A\vec{x} = \vec{b}$ with $E_k \cdots E_1$ from the left, we have:

$$E_k \cdots E_2 E_1 A\vec{x} = E_k \cdots E_2 E_1 \vec{b}.$$

This matrix equation has the same solution set as the original equation $A\vec{x} = \vec{b}$, since $E_k \cdots E_2 E_1$ is invertible.
The associated augmented matrix is:

$$\left(E_k \cdots E_2 E_1 A \,|\, E_k \cdots E_2 E_1 \vec{b}\right) = E_k \cdots E_2 E_1 \left(A \,|\, \vec{b}\right)$$

which is precisely the matrix obtained when we perform Gaussian Elimination on $\left(A \,|\, \vec{b}\right)$ using the row operations associated with $E_1, E_2, \ldots, E_k$.
