Linear Algebra and Matrix Methods
Chapters 1-2:
Introduction to Vectors,
Solving Linear Equations

A simple example: two equations in two variables

How do we solve this system of equations?

$$4x + 2y = 6$$
$$-10x + 5y = -20$$

Right at the outset, the question is tricky...

IS there a solution to this system of equations?

A simple example: two equations in two variables

How do we solve this system of equations?

$$\begin{aligned} 4x + 2y &= 6 \\ -10x + 5y &= -20 \end{aligned}$$

There are three possible cases:

- There are NO SOLUTIONS.
- There is EXACTLY ONE SOLUTION.
- There are INFINITELY MANY SOLUTIONS.

Vectors in two dimensions

We are going to analyze a lot of this kind of system.

An ordered pair (x, y) is a **vector** in the two-dimensional plane of real numbers $\mathbb{R}^2$.

When written like

$$\begin{pmatrix} x & y \end{pmatrix}$$

we call this a **row vector**.

When written like

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

we call this a **column vector**.

Considering each equation

Look at the first equation:

$$4x + 2y = 6$$

There are an infinite number of solutions (x, y) to this single equation: these solutions form a *line* in the plane.

$$L_1 = \{(x, y) : \ 4x + 2y = 6\}$$

Considering each equation

Likewise, there are an infinite number of solutions (x, y) to the second equation, all of which also form a line in the plane.

$$L_2 = \{(x, y) : \ -10x + 5y = -20\}$$

The solution of the system of equations is their intersection:

$$L_1 \cap L_2.$$

Considering each variable

We can consider all coefficients for one variable, and the right hand side, as column vectors.

This transforms the system of equations into one equation, where the variables are multiplied by vectors:

When written like this,

$$x \begin{pmatrix} 4 \\ -10 \end{pmatrix} + y \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ -20 \end{pmatrix},$$

our system is now one equation, with vector arithmetic.

What is a vector?

But what is this vector arithmetic we see here?

... wait, what even *is* a vector?

Vector arithmetic: scalar multiplication

A vector (x, y) in two dimensions is an arrow, starting from the origin, and pointing at the point (x, y) in the plane.

Scalar multiplication of a vector by a real number is *scaling* the vector: stretching or shrinking it to a different length. The real number doing the scaling is called (you guessed it) a **scalar**.

Vector arithmetic: scalar multiplication

To do scalar multiplication, multiply each coordinate by the scalar.

$$5.2 \begin{pmatrix} 4 \\ -10 \end{pmatrix} = \begin{pmatrix} 5.2 \cdot 4 \\ 5.2 \cdot -10 \end{pmatrix} = \begin{pmatrix} 20.8 \\ -52 \end{pmatrix}$$

If the scalar is negative, the vector “changes” “direction”.

$$-3.1 \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} -3.1 \cdot 5 \\ -3.1 \cdot 2 \end{pmatrix} = \begin{pmatrix} -15.5 \\ -6.2 \end{pmatrix}$$

Vector arithmetic: vector addition

Vector addition is done by adding the coordinates of each vector.

$$\begin{pmatrix} 4 \\ -10 \end{pmatrix} + \begin{pmatrix} 9 \\ 3 \end{pmatrix} = \begin{pmatrix} 4+9 \\ -10+3 \end{pmatrix} = \begin{pmatrix} 13 \\ -7 \end{pmatrix}$$

Graphically, we can draw the second vector as starting at the tip of the first vector; the sum is where the second vector now ends.

Vector arithmetic: vector addition

The sum of two vectors is called a **linear combination**.

Vector addition, like "regular" scalar addition, is **commutative**:

$$\begin{pmatrix} 4 \\ -10 \end{pmatrix} + \begin{pmatrix} 9 \\ 3 \end{pmatrix} = \begin{pmatrix} 4+9 \\ -10+3 \end{pmatrix} = \begin{pmatrix} 9+4 \\ 3+-10 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ -10 \end{pmatrix}.$$

Solving a system (column view): scalars = variables

To solve the column vector equation

$$xa_1 + ya_2 = b,$$

where

$$a_1 = \begin{pmatrix} 4 \\ -10 \end{pmatrix}, a_2 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, b = \begin{pmatrix} 6 \\ -20 \end{pmatrix},$$

we ask,

How do we scale a_1 and a_2, then add them, to get b?

A more "linear algebra"-type way of saying this is,

What **linear combination** of a_1 and a_2 is b?

Dot product of two vectors

The **dot product** of two vectors of the same size is the sum of the products of their coordinates, in sequence.

In two dimensions, consider two vectors $(1, 6)$ and $(-3, 5)$.

Their dot product is

$$(1, 6) \cdot (-3, 5) = (1 \cdot -3) + (6 \cdot 5) = -3 + 30 = 27.$$

Length (norm) of a vector

The **length** (**Euclidean norm**) of a vector is the square root of its dot product. We use doubled absolute value bars to denote length.

$$||(1,6)|| = \sqrt{(1,6)\cdot(1,6)} = \sqrt{1^2+6^2} = \sqrt{37}.$$

Dot product, angle between vectors

In general, if $v = \begin{pmatrix} v_1 & v_2 ... & v_n \end{pmatrix}$ and $w = \begin{pmatrix} w_1 & w_2 ... & w_n \end{pmatrix}$ are two vectors in $\mathbb{R}^n$, their dot product can be written in the summation notation

$$v \cdot w = \sum_{i=1}^{n} v_i w_i.$$

This number is related to the cosine of the angle θ between the two vectors v and w, regardless of the dimension of the space of their vectors:

$$\cos(\theta) = \frac{v \cdot w}{||v|| \cdot ||w||}.$$

Cauchy-Bunyakovsky-Schwarz inequality

In fact, since $-1 \leq \cos(\theta) \leq 1$ for any angle $0 \leq \theta < 2\pi$, there is a result called the **Cauchy-Bunyakovsky-Schwarz inequality**:

$$|v \cdot w| \leq ||v|| \cdot ||w||.$$

Orthogonal vectors

A special case of the dot product is when $v \cdot w = 0$; in this case $\cos(\theta) = 0$, and so v and w make a right angle. In this case, we call v and w **orthogonal** or **perpendicular**, and use the notation

$$v \perp w \iff v \cdot w = 0.$$

Note that the **zero vector**, the vector of all 0 (for any size), is orthogonal to any other vector of its size.

Unit, orthonormal, standard vectors

If the length of a vector v is $||v|| = 1$, we call it a **unit** vector.

Two orthogonal unit vectors v and w are called **orthonormal**.

The **standard unit vectors** in $\mathbb{R}^n$ are n vectors, each with a 1 in one coordinate, and 0 in every other coordinate. We denote them

$$\{e_1, e_2, ..., e_n\}.$$

Unit, orthonormal, standard vectors

For the plane $\mathbb{R}^2$, these are

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and in three-dimensional space $\mathbb{R}^3$, these are

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Length = Distance from 0 to end = Pythagorean Theorem

The length of a vector is the distance from the point $0 = (0,0)$ to the point at the end of the vector.

This is the hypotenuse of a right triangle that should be familiar from trigonometry: use the Pythagorean Theorem (also known as the Distance Formula) to find the length.

Length = Distance from 0 to end = Pythagorean Theorem

If two vectors v and w are orthogonal, then the Pythagorean Theorem can also be used to describe the length of their difference:

$$v \perp w \implies ||v||^2 + ||w||^2 = ||v - w||^2$$

Vector: direction and length, Triangle Inequality

For any nonzero vector v,

$$u = \frac{1}{||v||} v$$

is a unit vector.

Thus, a vector encodes the notions of "length" and "direction".

In fact, the difference in directions offers another inequality, since the difference of two vectors creates a triangle:

Triangle Inequality: $||v|| + ||w|| \geq ||v - w||$

Matrix view of the system of equations

A **matrix** is a row vector of column vectors, lined up together. (Equivalently, a matrix is a column vector of row vectors.)

We can write the columns of coefficient vectors from the system of equations as

$$A = \begin{pmatrix} a_1 & a_2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -10 & 5 \end{pmatrix}.$$

Matrix view of the system of equations

We can write the coefficients from each equation similarly as row vectors:

$$r_1 = \begin{pmatrix} 4 & 2 \end{pmatrix}, \; r_2 = \begin{pmatrix} -10 & 5 \end{pmatrix}$$

giving the same matrix

$$A = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -10 & 5 \end{pmatrix}.$$

Matrix view of the system of equations

Thus, the system of equations

$$\begin{aligned} 4x + 2y &= 6 \\ -10x + 5y &= -20 \end{aligned}$$

can be written, placing the variables in a column vector

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix},$$

as one matrix-vector equation

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

To understand this, we'll need to explain **matrix multiplication**.

Matrix multiplication

The **dimensions** of a matrix are

$$\#\text{ rows} \times \#\text{ columns},$$

which we will typically write as $m \times n$. We can multiply two matrices A and B, and get product AB (the order is important here), only if the # of columns of $A = \#$ of rows of B.

That is, if A is $m \times n$ and B is $r \times s$, then the matrix product AB only exists if $n = r$. The dimensions of AB are $m \times s$.

Thus, matrix multiplication is NOT commutative; in general, the products AB and BA are different matrices, and only both exist if $m = s$ and $n = r$.

Matrix multiplication

But what *is* the matrix product AB?

Let us assume A is $m \times n$ and B is $n \times p$, so that the matrix product $M = AB$ exists. M is $m \times p$.

Writing the rows of A with notation r_i, $i = 1, 2, ..., m$ and the columns of B as c_j, $j = 1, 2, ..., p$,

$$A = \begin{pmatrix} r_1 \\ r_2 \\ ... \\ r_m \end{pmatrix}, \; B = \begin{pmatrix} c_1 & c_2 & ... & c_p \end{pmatrix}$$

Matrix multiplication

... then the entry of M in row i, column j is the dot product

$$m_{ij} = r_i \cdot c_j = \sum_{k=1}^{n} a_{ik} b_{kj},$$

where a_{ik} is entry k in row i, and b_{kj} is entry k in column j.

We denote $M = (m_{ij})_{m \times p}$ to emphasize the (row, col) structure.

Matrix multiplication

A row vector in $\mathbb{R}^n$ is a $1 \times n$ matrix;
a column vector in $\mathbb{R}^n$ is a $n \times 1$ matrix.

If $M = AB$ as previously described, we can consider M to be a row vector of matrix-vector products:

$$M = \begin{pmatrix} Ac_1 & Ac_2 & \dots & Ac_p \end{pmatrix} = \begin{pmatrix} r_1 B \\ r_2 B \\ \dots \\ r_m B \end{pmatrix}.$$

Scalar multiplication

Now that we've addressed the hard part of matrix arithmetic, the easy parts are just like vector arithmetic:

- **scalar multiplication**: A matrix can be scaled by a real number by multiplying each element by that number.

  If $c \in \mathbb{R}$ and A is a matrix with entry a_{ij} in row i, column j, then cA is the matrix with entry ca_{ij} in that position.

  For example, if
  $$A = \begin{pmatrix} 2 & -4 & 5 \\ 6 & 6 & 7 \end{pmatrix},$$
  then
  $$-3A = \begin{pmatrix} -6 & 12 & -15 \\ -18 & -18 & -21 \end{pmatrix}.$$

Matrix addition

- **matrix addition**: Two matrices of the same dimensions can be added by adding their corresponding elements.

  For example, if

  $$A = \begin{pmatrix} 2 & -4 & 5 \\ 6 & 6 & 7 \end{pmatrix} \text{ and } B = \begin{pmatrix} 10 & 8 & 0 \\ -1 & -9 & 2 \end{pmatrix},$$

  then

  $$A + B = \begin{pmatrix} 12 & 4 & 5 \\ 5 & -3 & 9 \end{pmatrix}.$$

  Note that two matrices with different dimensions cannot be added together.

Matrix view of the system of linear equations

Back to the view of a system of equations

$$\begin{aligned} 4x + 2y &= 6 \\ -10x + 5y &= -20 \end{aligned}$$

as

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

Matrix view of the system of linear equations

In the row view, this product is

$$\begin{pmatrix} r_1 \cdot \vec{x} \\ r_2 \cdot \vec{x} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

In the column view, this product is

$$xa_1 + ya_2 = b.$$

These views all mean the same thing.

Solving a system of linear equations

For this simple 2×2 system, we are familiar with solving.

What if the system was much larger? 10×10, or 200×200?

We need a general technique that we will discover from the simplest case: a 1×1 "system".

Let $a, b \in \mathbb{R}$ with $a \neq 0$. What is the solution x of the equation

$$ax = b\ ?$$

This is easy:

$$x = \frac{b}{a}.$$

Solving a system of linear equations...

This fraction $\frac{b}{a}$ is the product $a^{-1}b$, where a^{-1} is the reciprocal **(multiplicative inverse)** of $a \in \mathbb{R}$: the number such that

$$a^{-1}a = aa^{-1} = 1.$$

1 is the **multiplicative identity** of the real numbers.

Solving a system of linear equations...

But be careful here: we know, for real numbers,

$$a^{-1}b = ba^{-1}$$

since regular real number multiplication is commutative.

Matrix multiplication is *not* commutative. (It *is* associative.)

We will stick to the order of terms with which we are presented.

... requires a matrix inverse.

The solution x to the equation

$$ax = b$$

is

$$a^{-1}ax = a^{-1}b \implies x = a^{-1}b.$$

What, then, is the solution to the system $A\vec{x} = b$?

... requires a matrix inverse.

If a **matrix inverse** A^{-1} exists, a matrix such that

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

with I the **identity matrix** (for that size matrix), then the solution to the system is, specifically using *left multiplication*,

$$A\vec{x} = b \implies A^{-1}A\vec{x} = A^{-1}b \implies \vec{x} = A^{-1}b.$$

When does a matrix inverse exist?

There are many requirements for a matrix inverse A^{-1} to exist.

First, A must be a **square** matrix: its dimensions must be $n \times n$.

Second, the column vectors must be **linearly independent**; we will define this more rigorously later.

When does a matrix inverse exist?

Let's see a 3×3 example of each of the three kinds of systems mentioned at the beginning:

- There are NO SOLUTIONS.
- There is EXACTLY ONE SOLUTION.
- There are INFINITELY MANY SOLUTIONS.

Dimension of a Space of Real Number-Scaled Vectors

The **dimension** of a set of vectors is the maximum number of vectors required to uniquely describe *any* point in the space *as a linear combination*.

The term "dimension" does not apply to the empty set $\emptyset$.

A set with only one vector has dimension 0: you cannot express any other vectors via linear combination. Geometrically, this is a point.

Dimension of a Space of Real Number-Scaled Vectors

If the dimension of the set of vectors is greater than 0, then the set has infinitely many vectors in it. Geometrically,

- dimension $= 1$: line
- dimension $= 2$: plane
- dimension $= 3$: space
- dimension > 3: while we still call this "space", we can't easily visualize this kind of space.

3 x 3 Systems

No solutions:

$$
\begin{aligned}
3x + 4y + 5z &= 6 \\
-6x - 8y - 8z &= -8 \\
3x + 4y + 5z &= 9
\end{aligned}
$$

Unique solution:

$$
\begin{aligned}
3x + 4y + 5z &= 6 \\
-6x - 8y - 8z &= -8 \\
x - y + 2z &= 5
\end{aligned}
$$

Infinitely many solutions:

$$
\begin{aligned}
3x + 4y + 5z &= 6 \\
-6x - 8y - 8z &= -8 \\
z &= 2
\end{aligned}
$$

One equation, three variables: a plane of solutions

One equation in three variables, such as

$$3x + 4y + 5z = 6,$$

requires two numbers to describe a point in its solution set.

That is, the set

$$\{(x, y, z) \in \mathbb{R}^3 : 3x + 4y + 5z = 6\}$$

has dimension 2.

Why? If you pick any $x, y \in \mathbb{R}$, then z is determined:

$$z = \frac{1}{5}(-3x - 4y + 6).$$

One equation, three variables: a plane of solutions

As a system, this one equation in three variables has infinitely many solutions: these solutions constitute a plane.

We can rewrite the solution set as

$$\left\{(x, y, z) : x, y \in \mathbb{R}, z = \frac{1}{5}(-3x - 4y + 6)\right\}.$$

Two equations, three variables: plane, line, or no solution

By adding a second equation in three variables, making a system of two equations, we may have any of the following scenarios:

- **redundancy**: If the new equation is a linear combination of previous equations in a system, then the new equation causes no change to the system solution. A shorthand for this scenario is "$0 = 0$" - no new information.
- **reduce the solution set dimension by 1**: Two intersecting planes have the common solution set of a line;
- **inconsistency**: If the new equation is a parallel plane to a previous equation, then there is no common solution for all equations. A shorthand is "$0 = 1$", a contradiction.

Two equations, three variables: plane $(0 = 0)$

$$\begin{aligned} 3x + 4y + 5z &= 6 \\ 9x + 12y + 15z &= 18 \end{aligned}$$

The second equation is the first equation multiplied by 3 on both sides. They are, basically, the same equation.

Adding -3 times "row (equation) 1" to row 2, the system "reduces":

$$\begin{aligned} 3x + 4y + 5z &= 6 \\ 0 &= 0 \end{aligned}$$

and we really only have the one equation system we started with.

Two equations, three variables: line (dim reduced by 1)

$$\begin{aligned} 3x + 4y + 5z &= 6 \\ -6x - 8y - 8z &= -8 \end{aligned}$$

These are two intersecting planes. The solution has dimension 1:

$$\{(x, y, z) : 3x + 4y + 5z = 6 \text{ and } -6x - 8y - 8z = -8\}$$

For each value of x, say $x = c$, the system reduces to a 2×2 system in y and z:

$$\begin{aligned} 4y + 5z &= 6 + -3c \\ -8y - 8z &= -8 + 6c \end{aligned}$$

which is a system that has a one-point solution $\{(y, z)\}$, where each is a function of c.

Two equations, three variables: inconsistency ($0 = 1$)

$$
\begin{aligned}
3x + 4y + 5z &= 6 \\
3x + 4y + 5z &= 9
\end{aligned}
$$

These two equations represent parallel planes.

If we subtract row 1 from row 2, then the system reduces to

$$
\begin{aligned}
3x + 4y + 5z &= 6 \\
0 &= 3
\end{aligned}
$$

which clearly is impossible.

Unique solution of a system of linear equations

What does it take to have a unique, one-point solution to a system of linear equations?

Each equation of n variables ideally reduces the solution set dimension by 1.

To have a unique solution:

- n equations in n variables
- all n equations are linearly independent
  (none is a linear combination of any subset of the others)

Unique solution of a system of linear equations

To have a unique solution:

- n equations in n variables
- all n equations are linearly independent (none is a linear combination of any subset of the others)

To *compute* a unique solution:

- reduce the system to a **triangular system** via the process of **elimination**, or **Gauss-Jordan reduction**
- use **back substitution** from the bottom to get the solution.

Unique solution $\iff$ inverse matrix exists

If a system of n equations in n variables can be written as the matrix-vector equation

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

then the system has a unique solution $\vec{x}$ iff A^{-1} exists; then,

$$\vec{x} = A^{-1}b.$$

If A^{-1} exists, we call A an **invertible** or **nonsingular** matrix.

If A^{-1} does not exist, we call A **non-invertible** or **singular**.

Identity matrix

If a square matrix A is invertible, then the product

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

the **identity matrix**, which generalizes the concept of the number "1" for square matrices.

I has 1 on the **diagonal** (row index = col index) and 0 elsewhere.

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

Multiplying a vector or matrix by I leaves it unchanged (much like $1 \cdot 5 = 5$).

$$IA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix}$$

Inverse, Identity matrices

Note that, in the special case of multplying by an inverse matrix,

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

multiplication is *commutative* - you can switch the order.

Another property of inverse matrices: just like in number division,

$$\frac{1}{(1/a)} = a,$$

we have that the inverse of the inverse matrix is the original matrix:

$$(A^{-1})^{-1} = A.$$

Elementary row operations

Solving a system involves **elementary row operations**.

(Recall, we refer to each equation as a row.)

There are three elementary row operations:

- swap two rows
- scale (multiply) a row
- add a scaled row to another row

For each of these operations, there is an **elementary matrix**, with one change of I, that does this operation via left multiplication on the system.

Elementary row operations: swap two rows

To **swap two rows**, use an elementary **permutation matrix**.

For example, to **swap two rows** 2 and 3, use

$$P_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

(In practice, we'll avoid this one by handling swaps outside of our solving process. Note that, intuitively, rerranging the equations in a system does not change the solution.) For example,

$$P_{2,3}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ -6 & -8 & -8 \\ 1 & -1 & 2 \end{pmatrix}.$$

Elementary row operations: scale one row

To **scale row** k **by the scalar** c, use the elementary matrix $E_{k,c}$, which is the modification of I where the row k diagonal is c instead of 1. For example,

$$E_{2,8} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ changes } A = \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix}$$

to

$$E_{2,8}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ 8 & -8 & 16 \\ -6 & -8 & -8 \end{pmatrix}.$$

Elementary row operations: scale row AND add to another

To **add the c-scaled row k to row j**, leaving only row j changed, use the elementary matrix $E_{k,j,c}$, which is the modification of I where the row j, column k entry is changed from 0 to c. For example, $E_{1,2,-3}$ causes the change

$$\text{row } 2 \to -3 \cdot \text{row } 1 + \text{row } 2:$$

$$E_{1,2,-3}A = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ -8 & -13 & -13 \\ -6 & -8 & -8 \end{pmatrix}.$$

Elementary matrices affect rows or columns

An elementary matrix changes the *rows* of a matrix if multiplying *on the left*.

An elementary matrix changes the *columns* of a matrix if multiplying *on the right*.

$$\text{row } 2 \to -3 \cdot \text{row } 1 + \text{row } 2:$$

$$E_{1,2,-3}A = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ -8 & -13 & -13 \\ -6 & -8 & -8 \end{pmatrix}.$$

$$\text{col } 1 \to -3 \cdot \text{col } 2 + \text{col } 1:$$

$$AE_{1,2,-3} = \begin{pmatrix} 3 & 4 & 5 \\ 1 & -1 & 2 \\ -6 & -8 & -8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -9 & 4 & 5 \\ 4 & -1 & 2 \\ 18 & -8 & -8 \end{pmatrix}.$$

Elementary matrix inverses: swap back

Each kind of elementary matrix has an inverse matrix that is also an elementary matrix of the same type.

To **swap two rows back**, use the **same elementary permutation matrix** - order of the row indices in the notation doesn't matter!

For example, to **swap two rows** 2 and 3 back to their original positions, just swap rows 3 and 2! Same action!

$$P_{j,k}^{-1} = P_{k,j} = P_{j,k}$$

$$P_{2,3}^{-1} = P_{3,2} = P_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Elementary matrix inverses: scale one row

To **undo scaling row** k **by the scalar** c, use the elementary matrix $E_{k,\frac{1}{c}}$, which scales row k by $\frac{1}{c}$.

$$E_{k,c}^{-1} = E_{k,\frac{1}{c}}$$

For example,

$$E_{2,8} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \; E_{2,8}^{-1} = E_{2,\frac{1}{8}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Elementary matrix inverses: scale and subtract

To **undo adding the** c**-scaled row** k **to row** j, leaving only row j changed, just subtract that scaled row:

$$E_{k,j,c}^{-1} = E_{k,j-c}$$

For example,

$$E_{1,2,-3} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};\ E_{1,2,-3}^{-1} = E_{1,2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solving a system of linear equations

Now we will outline the process to solve an arbitrarily large square system of linear equations (n equations in n variables).

If this process works, not only will we solve the system of equations

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

but we will have a **factorization** of the matrix A into the product of three matrices:

$$A = LDU,$$

where

- L is a **lower triangular** matrix with diagonal 1s and 0s above;
- D is a **diagonal** matrix, with nonzeros on the diagonal and 0s elsewhere;
- U is an **upper triangular** matrix with diagonal 1s and 0s below.

Solving a system of linear equations

First, we will assume this process works, and so there exists a factorization

$$A = LDU.$$

If this process fails at any point, then the rows of A must be rearranged, or there is not a unique solution to the system.

1. Commit forward, downward elimination of variables by left multiplying elementary matrices on the equation

   $$A\vec{x} = b, \quad \text{that is,} \quad LDU\vec{x} = b,$$

   until it is in the form

   $$DU\vec{x} = L^{-1}b.$$

Solving a system of linear equations

2. Scale each equation so that the leading variable on each is 1. This will change the system from

$$DU\vec{x} = L^{-1}c$$

into

$$U\vec{x} = D^{-1}L^{-1}b.$$

3. Upward, backward substitution of variables, starting at the bottom, yielding the solution

$$\vec{x} = U^{-1}D^{-1}L^{-1}b.$$

Solving a system of linear equations, factoring $A = LDU$

In other words, we compute the inverse matrix

$$A^{-1} = U^{-1}D^{-1}L^{-1}$$

and apply it to

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

to get

$$\vec{x} = A^{-1}b.$$

Since L^{-1}, D^{-1}, and U^{-1} are all products of elementary matrices, we can find their inverses L, D, and U to find the factorization

$$A = LDU.$$

Computing the solution, factorization of a linear system

We will use the technique of an **augmented matrix** to do our calculations by hand.

This allows us to do computations on A and b simultaneously.

We start with $A\vec{x} = b$ in the form

$$\left[\begin{array}{c|c} A & b \end{array}\right]$$

and attempt to end with the solution $\vec{x} = A^{-1}b$ in the form

$$\left[\begin{array}{c|c} I & A^{-1}b \end{array}\right]$$

by committing left multiplication of elementary matrices.

Computing using an augmented matrix

We will call the diagonal entries on the left side **pivots**.

Our goal is to make the pivots 1 (representing 1 of each variable) and 0 elsewhere on the left side, by eliminating variable coefficients below each pivot.

If ever a pivot is zeroed out, this process fails for the given row organization of A.

Computing using an augmented matrix

If ever a

$$\text{“}0 = 0\text{”}$$

(underdetermined system; infinitely many solutions)

or

$$\text{“}0 = 1\text{”}$$

(overdetermined system; no solutions)

occurs, then *no* reorganization of the rows of A will work, and the system does not have a unique solution.

Transposes

If v is a row vector, then its **transpose**, denoted v^t, is a column vector with the same entries.

If v is a column vector, then its **transpose** is a row vector.

The dot product of two row vectors v and w is the same value as the matrix product vw^t.

The dot product of two column vectors v and w is the same value as the matrix product $v^t w$.

In general, the **transpose** of the $m \times n$ matrix A is the $n \times m$ matrix A^t that changes rows to columns.

Properties of Transposes

- $(A^t)^t = A$
- $(A+B)^t = A^t + B^t$
- A is called a **symmetric matrix** if $A^t = A$.
- Diagonal matrices are clearly symmetric: $D^t = D$.
- If the product AB exists, then $(AB)^t = B^t A^t$.
- $A\vec{x} = b$ is the same system as $\vec{x}^t A^t = b^t$.

Properties of Inverses

- $(A^{-1})^{-1} = A$
- $(A^{-1})^t = (A^t)^{-1}$
- A is called an **involution** if $A^{-1} = A$. This means $A^2 = I$.
- If the product AB exists, and A and B are both invertible, then $(AB)^{-1} = B^{-1}A^{-1}$.
- $A\vec{x} = b$ is solved by $\vec{x} = A^{-1}b$, if A^{-1} exists.
- If $A = LDU$ and A is symmetric, then $U = L^t$, yielding the factorization $A = LDL^t$.
- Call an invertible square matrix A **orthogonal** if $A^t = A^{-1}$. Then the columns $a_1, a_2, ..., a_n$ of A are all orthonormal:

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \implies A^{-1}A = A^tA = I.$$

Properties of Permutations

- All permutation matrices P are square.
- Permutation matrices are orthogonal.
- If a square matrix A is invertible, but does not allow our elimination process to work to factor $A = LDU$, then some permutation of the system PA will.

  (We will not go through the process of seeing how to decide how to determine such a P here.)