

# Lecture 4 Matrix Operations

Yiping Lu Based on Dr. Ralph Chikhany's Slide



# Strang Sections 2.3 – Elimination Using Matrices and 2.4 – Rules for Matrix Operations

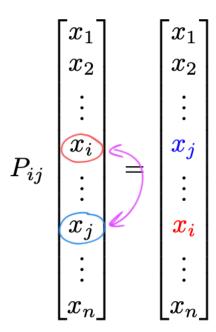
Course notes adapted from *Introduction to Linear Algebra* by Strang (5<sup>th</sup> ed), N. Hammoud's NYU lecture notes, and *Interactive Linear Algebra* by Margalit and Rabinoff, in addition to our text



#### Recall

$$I = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \xrightarrow{I\vec{x}} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

$$P_{ij} = egin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \ dots & & & & & & \ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \ dots & & & & & & \ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \ dots & & & & & & \ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$



$$P_{ij} = egin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \ dots & & & & & & & \ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \ dots & & & & & & & \ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \ dots & & & & & & \ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$P_{ij}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$$

$$P_{31} = egin{bmatrix} 0 & 0 & 1 & \dots & 0 \ 0 & 1 & 0 & \dots & 0 \ 1 & 0 & 0 & \dots & 0 \ dots & \ddots & & & \ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$P_{ij} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$P_{ij} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix}$$

$$P_{ij} egin{bmatrix} x_1 \ x_2 \ dots \ x_i \ dots \ x_j \ dots \ x_n \end{bmatrix} egin{bmatrix} x_1 \ x_2 \ dots \ x_j \ dots \ x_i \ dots \ x_n \end{bmatrix}$$

$$P_{31} = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \xrightarrow{P_{31} \vec{x}} \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$



# **Matrix Operations**

#### Recall

Let A be an  $m \times n$  matrix.

We write  $a_{ij}$  for the entry in the *i*th row and the *j*th column. It is called the *ij*th entry of the matrix.

The entries  $a_{11}, a_{22}, a_{33}, \ldots$  are the **diagonal entries**; they form the **main diagonal** of the matrix.

A diagonal matrix is a *square* matrix whose only nonzero entries are on the main diagonal.

The  $n \times n$  identity matrix  $I_n$  is the diagonal matrix with all diagonal entries equal to 1. It is special because  $I_n v = v$  for all v in  $\mathbf{R}^n$ .

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

$$jth \ column$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

$$\begin{pmatrix}
a_{11} & 0 & 0 \\
0 & a_{22} & 0 \\
0 & 0 & a_{33}
\end{pmatrix}$$

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

#### Recall

The **zero matrix** (of size  $m \times n$ ) is the  $m \times n$  matrix 0 with all zero entries.

The **transpose** of an  $m \times n$  matrix A is the  $n \times m$  matrix  $A^T$  whose rows are the columns of A. In other words, the ij entry of  $A^T$  is  $a_{ji}$ .

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

$$A^{T}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \qquad \qquad \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}$$

#### Matrix Addition and Scalar Multiplication

You add two matrices component by component, like with vectors.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{pmatrix}$$

Note you can only add two matrices of the same size.

You multiply a matrix by a scalar by multiplying each component, like with vectors.

$$c \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{pmatrix}.$$

These satisfy the expected rules, like with vectors:

$$A + B = B + A$$

$$c(A + B) = cA + cB$$

$$(cd)A = c(dA)$$

$$(A + B) + C = A + (B + C)$$

$$(c + d)A = cA + dA$$

$$A + 0 = A$$

$$A = egin{bmatrix} a_{11} & \dots & a_{1n} \ a_{21} & \dots & a_{2n} \ dots & & & \ \vdots & & & \ a_{m1} & \dots & a_{mn} \end{bmatrix} \quad B = egin{bmatrix} b_{11} & \dots & b_{1m} \ b_{21} & \dots & b_{2m} \ dots & & \ \vdots & & \ b_{l1} & \dots & b_{lm} \end{bmatrix} \quad C = egin{bmatrix} c_{11} & \dots & c_{1k} \ c_{21} & \dots & c_{2k} \ dots & & \ \vdots & & \ c_{n1} & \dots & c_{nk} \end{bmatrix}$$

Beware: matrix multiplication is more subtle than addition and scalar multiplication.

must be equal

Let A be an  $m \times n$  matrix and let B be an  $n \times p$  matrix with columns  $v_1, v_2, \ldots, v_p$ :

$$B = \begin{pmatrix} | & | & | \\ v_1 & v_2 & \cdots & v_p \\ | & | & | \end{pmatrix}.$$

The **product** AB is the  $m \times p$  matrix with columns  $Av_1, Av_2, \ldots, Av_p$ :

The equality is a definition 
$$AB \stackrel{\text{def}}{=} \begin{pmatrix} | & | & | \\ Av_1 & Av_2 & \cdots & Av_p \\ | & | & | \end{pmatrix}$$
.

In order for  $Av_1, Av_2, \ldots, Av_p$  to make sense, the number of columns of A has to be the same as the number of rows of B.

Example
$$\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}
\begin{pmatrix}
1 & -3 \\
2 & -2 \\
3 & -1
\end{pmatrix} = \begin{pmatrix}
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}
\cdot
\begin{pmatrix}
1 \\
2 \\
3
\end{pmatrix}
\quad
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}
\cdot
\begin{pmatrix}
-3 \\
-2 \\
-1
\end{pmatrix}
\end{pmatrix}$$

A row vector of length n times a column vector of length n is a scalar:

$$(a_1 \quad \cdots \quad a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \cdots + a_nb_n.$$

A row vector of length n times a column vector of length n is a scalar:

$$(a_1 \quad \cdots \quad a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \cdots + a_nb_n.$$

Another way of multiplying a matrix by a vector is:

$$Ax = \begin{pmatrix} -r_1 - \\ \vdots \\ -r_m - \end{pmatrix} x = \begin{pmatrix} r_1 x \\ \vdots \\ r_m x \end{pmatrix}.$$

A row vector of length n times a column vector of length n is a scalar:

$$(a_1 \quad \cdots \quad a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \cdots + a_nb_n.$$

Another way of multiplying a matrix by a vector is:

$$Ax = \begin{pmatrix} -r_1 - \\ \vdots \\ -r_m - \end{pmatrix} x = \begin{pmatrix} r_1 x \\ \vdots \\ r_m x \end{pmatrix}.$$

On the other hand, you multiply two matrices by

$$AB = A \begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ Ac_1 & \cdots & Ac_p \\ | & & | \end{pmatrix}.$$

A row vector of length n times a column vector of length n is a scalar:

$$(a_1 \quad \cdots \quad a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \cdots + a_nb_n.$$

Another way of multiplying a matrix by a vector is:

$$Ax = \begin{pmatrix} -r_1 - \\ \vdots \\ -r_m - \end{pmatrix} x = \begin{pmatrix} r_1 x \\ \vdots \\ r_m x \end{pmatrix}.$$

On the other hand, you multiply two matrices by

$$AB = A \begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ Ac_1 & \cdots & Ac_p \\ | & & | \end{pmatrix}.$$

It follows that

$$AB = \begin{pmatrix} -r_1 - \\ \vdots \\ -r_m - \end{pmatrix} \begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} = \begin{pmatrix} r_1c_1 & r_1c_2 & \cdots & r_1c_p \\ r_2c_1 & r_2c_2 & \cdots & r_2c_p \\ \vdots & \vdots & & \vdots \\ r_mc_1 & r_mc_2 & \cdots & r_mc_p \end{pmatrix}$$

The ij entry of C = AB is the ith row of A times the jth column of B:

$$c_{ij} = (AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$

This is how everybody on the planet actually computes AB. Diagram (AB = C):

$$\begin{pmatrix}
a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots \\
a_{i1} & \cdots & a_{ik} & \cdots & a_{in}
\end{vmatrix} \cdot \begin{pmatrix}
b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
b_{k1} & \cdots & b_{kj} & \cdots & b_{kp} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
b_{n1} & \cdots & b_{nj} & \cdots & b_{np}
\end{pmatrix} = \begin{pmatrix}
c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
c_{m1} & \cdots & c_{mj} & \cdots & c_{mp}
\end{pmatrix}$$

$$jth column$$

$$ij \text{ entry}$$

The ij entry of C = AB is the ith row of A times the jth column of B:

$$c_{ij} = (AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$

This is how everybody on the planet actually computes AB. Diagram (AB = C):

$$\begin{pmatrix}
a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots \\
a_{i1} & \cdots & a_{ik} & \cdots & a_{in}
\end{vmatrix} \cdot \begin{pmatrix}
b_{11} & \cdots & b_{1j} \\
\vdots & \vdots & \vdots \\
b_{k1} & \cdots & b_{kj} \\
\vdots & \vdots & \vdots \\
b_{n1} & \cdots & b_{nj}
\end{pmatrix} \cdot \begin{pmatrix}
c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\
\vdots & \vdots & \vdots & \vdots \\
c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\
\vdots & \vdots & \vdots & \vdots \\
c_{m1} & \cdots & c_{mj} & \cdots & c_{mp}
\end{pmatrix}$$

$$jth column$$

$$jth column$$

$$ij entry$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 & \square \\ \square & \square \end{pmatrix} = \begin{pmatrix} 14 & \square \\ \square & \square \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} \Box & \Box \\ 4 \cdot 1 + 5 \cdot 2 + 6 \cdot 3 & \Box \end{pmatrix} = \begin{pmatrix} \Box & \Box \\ 32 & \Box \end{pmatrix}$$

#### Matrix-Matrix and Matrix-Vector

Matrix vector multiplication is a Matrix Matrix multiplication

$$A[\vec{v}_1, \cdots, \vec{v}_k] = [A\vec{v}_1, \cdots, A\vec{v}_k]$$

$$egin{bmatrix} ec{r}_1^{ op} \ ec{r}_2^{ op} \ dots \ ec{r}_k^{ op} \end{bmatrix} A = egin{bmatrix} ec{r}_1^{ op} A \ ec{r}_2^{ op} A \ dots \ ec{r}_k^{ op} A \end{bmatrix}$$

Let 
$$A = \begin{bmatrix} 1 & 2 & 5 & 3 \\ 3 & -1 & 0 & -3 \\ 2 & -2 & 1 & 0 \end{bmatrix}$$
 and  $B = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 4 & 1 & 3 \\ 0 & -2 & 1 \end{bmatrix}$ .

Compute AB and BA (if possible).

Mostly matrix multiplication works like you'd expect. Suppose A has size  $m \times n$ , and that the other matrices below have the right size to make multiplication work.

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

$$(B+C)A = BA + CA$$

$$c(AB) = A(cB)$$

$$A(B+C) = (AB + AC)$$

$$c(AB) = (cA)B$$

$$I_nA = A$$

$$AI_m = A$$

Most of these are easy to verify.

#### Warnings!

ightharpoonup AB is usually not equal to BA.

In fact, AB may be defined when BA is not.

▶ AB = AC does not imply B = C, even if  $A \neq 0$ .

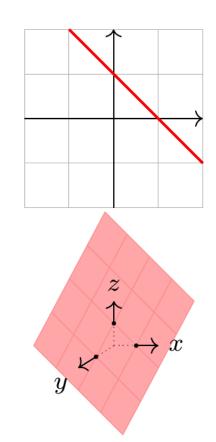
▶ AB = 0 does not imply A = 0 or B = 0.



What does the solution set of a linear equation look like?

► 
$$x + y = 1$$
  
 $x + y = 1$   
 $y = 1 - x$ 

$$x + y + z = 1$$
  
 $x + y + z = 1$   
 $x + y + z = 1$ 

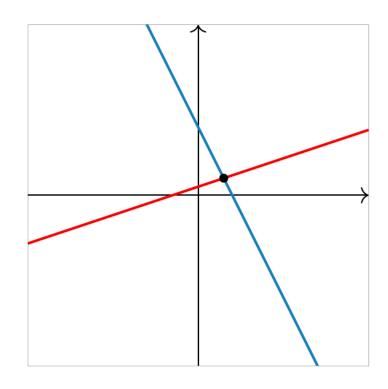


[not pictured here]

What does the solution set of a *system* of more than one linear equation look like?

$$x - 3y = -3$$
$$2x + y = 8$$

... is the *intersection* of two lines, which is a *point* in this case.

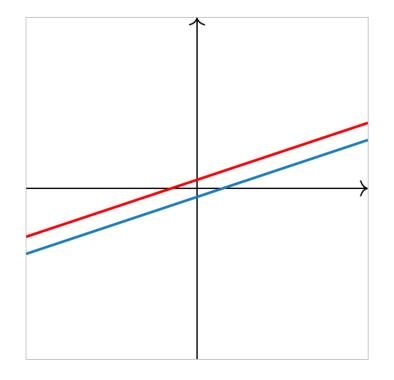


In general it's an intersection of lines, planes, etc.

In what other ways can two lines intersect?

$$x - 3y = -3$$
$$x - 3y = 3$$

has no solution: the lines are parallel.

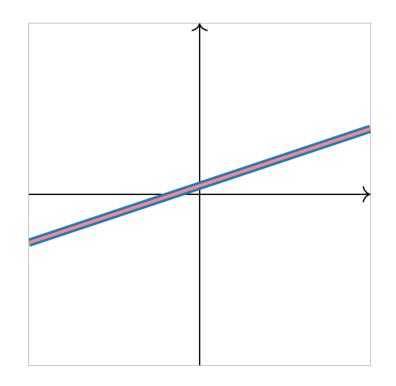


A system of equations with no solutions is called **inconsistent**.

In what other ways can two lines intersect?

$$x - 3y = -3$$
$$2x - 6y = -6$$

has infinitely many solutions: they are the *same line*.



Note that multiplying an equation by a nonzero number gives the *same* solution set. In other words, they are equivalent (systems of) equations.

#### Example

Solve the system of equations

$$x + 2y + 3z = 6$$
  
 $2x - 3y + 2z = 14$   
 $3x + y - z = -2$ 

This is the kind of problem we'll talk about for a good portion of the course.

- ▶ A **solution** is a list of numbers x, y, z, ... that make *all* of the equations true.
- ► The **solution set** is the collection of all solutions.
- ▶ **Solving** the system means finding the solution set.

Consider the following system of two equations in two unknowns

$$x_1 - 2x_2 = 1$$
$$3x_1 + 2x_2 = 11$$

This system could be expressed in matrix notation as:

$$\left[\begin{array}{cc} 1 & -2 \\ 3 & 2 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} 1 \\ 11 \end{array}\right]$$

# Systems of Equations – 2D – Row vs. Column Picture

$$\left[ egin{array}{cc} 1 & -2 \ 3 & 2 \end{array} 
ight] \left[ egin{array}{c} x_1 \ x_2 \end{array} 
ight] = \left[ egin{array}{c} 1 \ 11 \end{array} 
ight]$$

Row picture: 
$$(1,-2) \cdot (x_1,x_2) = 1 \implies x_1 - 2x_2 = 1$$
  
 $(3,2) \cdot (x_1,x_2) = 11 \implies 3x_1 + 2x_2 = 11$ 

# Systems of Equations – 2D – Row vs. Column Picture

$$\left[ egin{array}{cc} 1 & -2 \ 3 & 2 \end{array} 
ight] \left[ egin{array}{c} x_1 \ x_2 \end{array} 
ight] = \left[ egin{array}{c} 1 \ 11 \end{array} 
ight]$$

Column picture: 
$$x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$$

#### Systems of Equations – 3D – Row vs. Column Picture

If we have three equations with three unknowns, it is still possible to draw a picture of what a solution looks like. Each of the three equations represents a plane in 3D, and their intersection gives the solution of the system. As soon as you go above 3D, visualization becomes impossible.

Consider the following system of three equations in three unknowns

# Systems of Equations – 3D – Row vs. Column Picture

$$\left[egin{array}{cccc} 1 & 2 & 3 \ 2 & 5 & 2 \ 6 & -3 & 1 \end{array}
ight] \left[egin{array}{c} x_1 \ x_2 \ x_3 \end{array}
ight] = \left[egin{array}{c} 6 \ 4 \ 2 \end{array}
ight]$$

**Row picture**: 
$$(1,2,3) \cdot (x_1, x_2, x_3) = 6 \implies x_1 + 2x_2 + 3x_3 = 6$$
  
 $(2,5,2) \cdot (x_1, x_2, x_3) = 4 \implies 2x_1 + 5x_2 + 2x_3 = 4$   
 $(6,-3,1) \cdot (x_1, x_2, x_3) = 2 \implies 6x_1 - 3x_2 + x_3 = 2$ 

# Systems of Equations – 3D – Row vs. Column Picture

$$\left[egin{array}{cccc} 1 & 2 & 3 \ 2 & 5 & 2 \ 6 & -3 & 1 \end{array}
ight] \left[egin{array}{c} x_1 \ x_2 \ x_3 \end{array}
ight] = \left[egin{array}{c} 6 \ 4 \ 2 \end{array}
ight]$$

Column picture: 
$$x_1 \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$$