• Instructor: He Wang Email: he.wang@northeastern.edu

§1.2 Matrices, Vectors, and Gauss–Jordan Elimination

▶ Matrices

Recall the linear system $\begin{cases} x_1 - 3x_2 - 5x_3 = 1 \\ x_1 - x_2 - 2x_3 = 0 \\ 3x_1 - x_2 + x_3 = 3 \end{cases}$ in §1.1. Once we keep the order of the variables, all information of the linear system is captured by the following ***matrix:***

$$M = \begin{bmatrix} 1 & -3 & -5 & 1 \\ 1 & -1 & -2 & 0 \\ 3 & -1 & 1 & 3 \end{bmatrix}$$

It is a 3×4 matrix (called 3 by 4 matrix). It has 3 **rows** and 4 **columns.**

The above matrix M is called the **augmented matrix** of the linear system. Sometimes, we separate the last column as $M = \left[\begin{array}{ccc|c} 1 & -3 & -5 & 1 \\ 1 & -1 & -2 & 0 \\ 3 & -1 & 1 & 3 \end{array}\right]$

The 3×3 matrix $A = \begin{bmatrix} 1 & -3 & -5 \\ 1 & -1 & -2 \\ 3 & -1 & 1 \end{bmatrix}$ is called the **coefficient matrix** of the linear system.

In general, we denote a **size** 3×4 matrix A by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

using double subscripts to refer the position of row and column. The real number a_{ij} is called the **(i, j)-th entry** of A.

Warning on the notation: We can write a matrix as $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, but we **NEVER** write a matrix as $A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$. This means the determinant of a matrix.

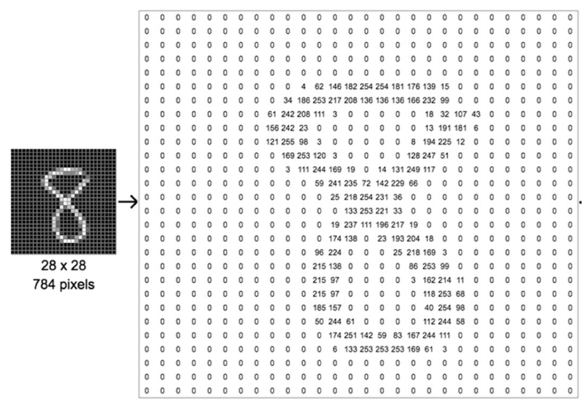
Digital world examples: An HD 1080P picture has 1920×1080 pixels. A 4K picture has 3840×2160 pixels. They are stored by matrices.

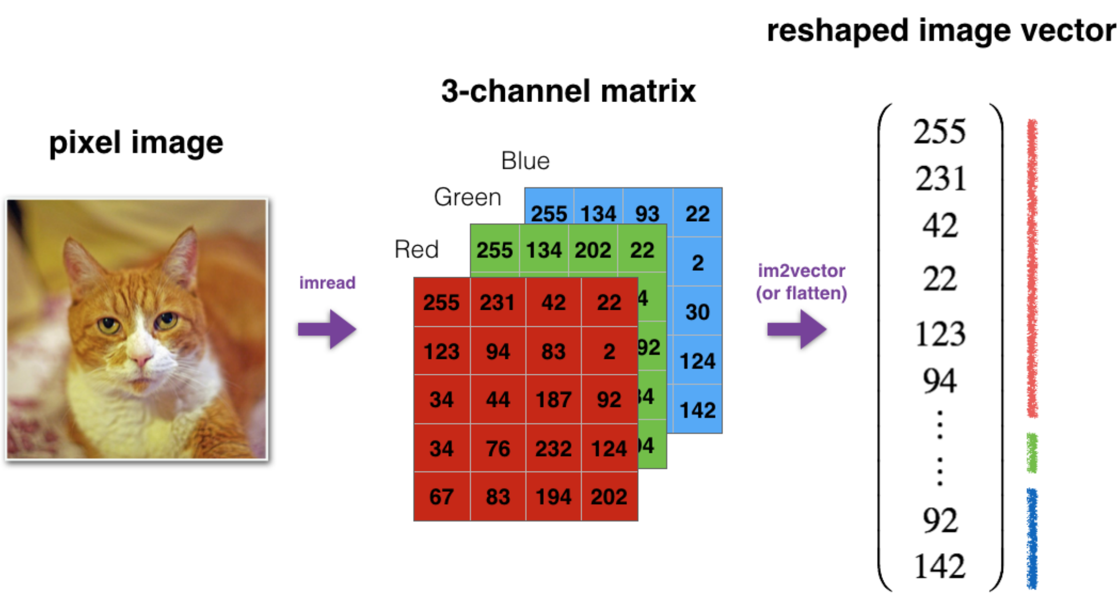
For each pixel, Black-White (1 bpp) has $2^1 = 2$ colors, gray picture has $2^8 = 256$ colors, color picture (rbg) has three layers in the matrix, while the "Truecolor" (24 bpp) has $2^{24} = 16,777,216$ colors. 

Digital image processing is a popular research field in computer science and in applied mathematics.

A music CD is also stored as a $n \times 1$ matrix. For example, $n = 440,000$ for a 10 second music.

Example: The MNIST database of handwritten digits, has a training set of 60,000 examples, and a test set of 10,000 examples. Many methods have been tested with this database.(e.g., `https://www.cs.ryerson.ca/~aharley/vis/conv/`)





Matrix comes from many other real world problems like, dynamical system (Google's page-rank algorithm, by Page and Brin in 1998).

Definition.

Two matrices A and B are said to be **equal** (denoted as $A = B$) if they have the same size, and if corresponding entries are equal: $a_{ij} = b_{ij}$.

The **zero matrix**, written simply as 0, is any $m \times n$ matrix all of whose entries are 0. For example,

$$0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad 0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Definition. Some square matrix terminologies:

1. Let A be an $n \times n$ matrix which is called an **square matrix**.
2. The **diagonal entries** of A are the entries $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ forming the **diagonal** of A.
3. A is called a **diagonal matrix** if all off diagonal entries of A are zeros. (i.e., $a_{ij} = 0$ if $i \neq j$.)

   In particular, the $n \times n$ **identity matrix** I_n whose diagonal entries are all 1 (and all off diagonal entries are 0).
4. A is called **lower triangular** if $a_{ij} = 0$ whenever $j > i$, (that is if all entries a_{ij} above the main diagonal are zero.) Similarly, we can define **upper triangular matrix** if if $a_{ij} = 0$ whenever $i > j$.

Examples:

$$A = \begin{bmatrix} 4 & 2 & 1.2 \\ 0 & 0 & 6 \\ 1/3 & 4 & 2 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 5 \end{bmatrix}, D = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 5 \end{bmatrix}, I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$U = \begin{bmatrix} 3 & \sqrt{2} & 2.7 & 2 \\ 0 & 0 & \pi & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \qquad L = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

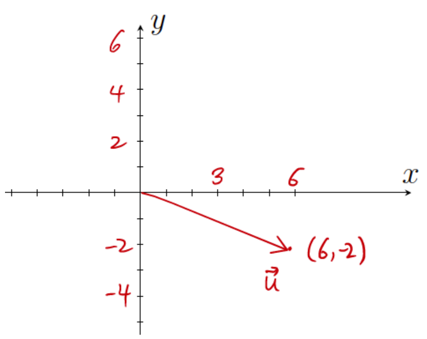
Definition. ▶ Vectors

An $m \times 1$ matrix is called a (column) ***vector*** in $\mathbb{R}^m$.

$$\vec{v} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix},$$

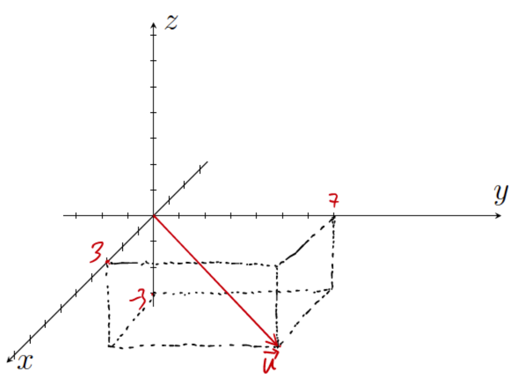
where $c_1, c_2, \ldots, c_m$ are real numbers. $\mathbb{R}^n$ is called a **vector space**.

Example 1. Vectors in $\mathbb{R}^2$



$$\vec{u} = \begin{bmatrix} 6 \\ -2 \end{bmatrix} \in \mathbb{R}^2$$

Example 2. Vectors in $\mathbb{R}^3$



$$\vec{u} = \begin{bmatrix} 3 \\ 7 \\ -3 \end{bmatrix} \in \mathbb{R}^3$$

Example 3. Vectors in $\mathbb{R}^4$ and $\mathbb{R}^5$

$$\vec{u} = \begin{bmatrix} 3 \\ -1 \\ \sin(2) \\ \ln(3) \end{bmatrix}, \qquad \vec{v} = \begin{bmatrix} \sqrt{3} \\ 2.6 \\ \pi \\ 3.8 \\ e \end{bmatrix} \in \mathbb{R}^5.$$

We can always illustrate a vector $\vec{v} \in \mathbb{R}^n$ as an arrow starting from origin zero.

▶ **Gauss–Jordan Elimination** Now, we introduce an algorithm for solving a linear system of any size. Recall the linear system $\begin{cases} x_1 - 3x_2 - 5x_3 = 1 \\ x_1 - x_2 - 2x_3 = 0 \\ 3x_1 - x_2 + x_3 = 3 \end{cases}$. The linear system is fully captured by the **augmented matrix**: $\left[\begin{array}{ccc|c} 1 & -3 & -5 & 1 \\ 1 & -1 & -2 & 0 \\ 3 & -1 & 1 & 3 \end{array}\right]$

Let's compare the procedures solving this linear system:

$$\begin{array}{r} \\ L_2 - L_1 \\ L_3 - 3L_1 \end{array} \begin{cases} x_1 - 3x_2 - 5x_3 = 1 \\ 2x_2 + 3x_3 = -1 \\ 8x_2 + 16x_3 = 0 \end{cases} \qquad \xrightarrow[R_3 - 3R_1]{R_2 - R_1} \left[\begin{array}{ccc|c} 1 & -3 & -5 & 1 \\ 0 & 2 & 3 & -1 \\ 0 & 8 & 16 & 0 \end{array}\right]$$

$$\begin{array}{r} \\ \\ L_3 - 4L_2 \end{array} \begin{cases} x_1 - 3x_2 - 5x_3 = 1 \\ 2x_2 + 3x_3 = -1 \\ 4x_3 = 4 \end{cases} \qquad \xrightarrow{R_3 - 4R_2} \left[\begin{array}{ccc|c} 1 & -3 & -5 & 1 \\ 0 & 2 & 3 & -1 \\ 0 & 0 & 4 & 4 \end{array}\right]$$

$$\begin{array}{r} \\ \\ \frac{1}{4}L_3 \end{array} \begin{cases} x_1 - 3x_2 - 5x_3 = 1 \\ 2x_2 + 3x_3 = -1 \\ x_3 = 1 \end{cases} \qquad \xrightarrow{\frac{1}{4}R_4} \left[\begin{array}{ccc|c} 1 & -3 & -5 & 1 \\ 0 & 2 & 3 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{array}{r} L_1 + 5L_3 \\ L_2 - 3L_3 \\ \\ \end{array} \begin{cases} x_1 - 3x_2 = 6 \\ 2x_2 = -4 \\ x_3 = 1 \end{cases} \qquad \xrightarrow[R_2 - 3R_3]{R_1 + 5R_3} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 6 \\ 0 & 2 & 0 & -4 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{array}{r} \\ \frac{1}{2}L_2 \\ \\ \end{array} \begin{cases} x_1 - 3x_2 = 6 \\ x_2 = -2 \\ x_3 = 1 \end{cases} \qquad \xrightarrow{\frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{array}{r} L_1 + 3L_2 \\ \\ \\ \end{array} \begin{cases} x_1 = 0 \\ x_2 = -2 \\ x_3 = 1 \end{cases} \qquad \xrightarrow{R_1 + 3R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

So, $x_1 = 0, x_2 = -1$ and $x_3 = 1$.

The first nonzero entry of a row is called **leading entry**(or **pivot**) in this row.

Definition.

A matrix is in ***row-echelon form* (ref)** if it has the following conditions:
1. If a column contains a leading entry, then all entries in a column below a leading entry are zeros. (Zeros Below)
2. If a column contains a leading entry, then each row above it contains a leading entry further to the left. (Above-Left or Down-Below.)

Condition 2 implies that all zero rows are at the bottom of the matrix.

Definition.

A matrix is in ***reduced row-echelon form* (rref)**, if it satisfies the above two conditions together with the following two conditions:
3. The leading entry in each nonzero row is 1. (Leading 1.)
4. If a column contains a leading entry, then all entries in a column above a leading entry are zeros. (Zeros Above)

$$\begin{bmatrix} \blacksquare & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & * & 0 & 0 & 0 & * \\ 0 & 0 & 1 & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\blacksquare$: Non zero number

$*$: any number

Example 4. The following matrices are **not** in **ref**.

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

Example 5. The following matrices are in **ref**, but **not** in **rref**.

$$\begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Definition.

Types of ***elementary row operations*** :

1. Scaling: Multiply(or divide) a row by a nonzero scalar.
2. Replacement: Add (or substract) to a multiple of a row from another row.
3. Interchange: Interchange two rows.

Using the elementary row operations, one can change any matrix to a (**unique**) reduced row-echelon form (rref). This method of solving a linear system is called ***Gauss–Jordan Elimination.***

Example 6. Using Row Reduction Algorithm (Gauss-Elimination) to solve a linear system with the following augmented matrix. (Write down all solutions using free variables)

$$\left[\begin{array}{ccccc|c} 0 & 0 & 2 & -8 & -1 & 3 \\ 1 & 6 & 2 & -5 & -2 & -4 \\ 2 & 12 & 2 & -2 & -3 & -11 \\ 1 & 6 & 0 & 3 & -2 & -14 \end{array}\right]$$

$$\begin{bmatrix} 0 & 0 & 2 & -8 & -1 & 3 \\ 1 & 6 & 2 & -5 & -2 & -4 \\ 2 & 12 & 2 & -2 & -3 & -11 \\ 1 & 6 & 0 & 3 & -2 & -14 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 2 & 12 & 2 & -2 & -3 & -11 \\ 1 & 6 & 0 & 3 & -2 & -14 \end{bmatrix} \xrightarrow[R_4-R_1]{R_3-2R_1}$$

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & -2 & 8 & 1 & -3 \\ 0 & 0 & -2 & 8 & 0 & -10 \end{bmatrix} \xrightarrow[R_4+R_2]{R_3+R_2} \begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -7 \end{bmatrix} \xrightarrow[R_3 \leftrightarrow R_4]{-R_4} \begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ ref}$$

$$\xrightarrow[R_1+2R_3]{R_2+R_3} \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1-2R_2} \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ rref}$$

$$\begin{cases} x_1 + 6x_2 + 3x_4 = 0 \\ x_3 - 4x_4 = 5 \\ x_5 = 7 \end{cases} \Longrightarrow \begin{cases} x_1 = -6x_2 - 3x_4 \\ x_3 = 5 + 4x_4 \\ x_5 = 7 \\ x_2, x_4 \text{ are free!} \end{cases}$$

Definition.

- **basic variables**: All x_i corresponding to the pivot cloumns.
- **free variables**: the rest variables. (Act as parameters.)

Parametric vector form of all solutions of a linear system:

For the above example,

Using free variables, we can write all solutions of $A\vec{x}$
The parametric vector form of all solutions is

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -6x_2 - 3x_4 \\ x_2 \\ 5 + 4x_4 \\ x_4 \\ 7 \end{bmatrix} = x_2 \begin{bmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 4 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \\ 0 \\ 0 \\ 7 \end{bmatrix}$$

◊ The Row Reduction Algorithm (Gauss-Jordan elimination):

1. Begin with the **leftmost *nonzero*** column. This is a pivot column.

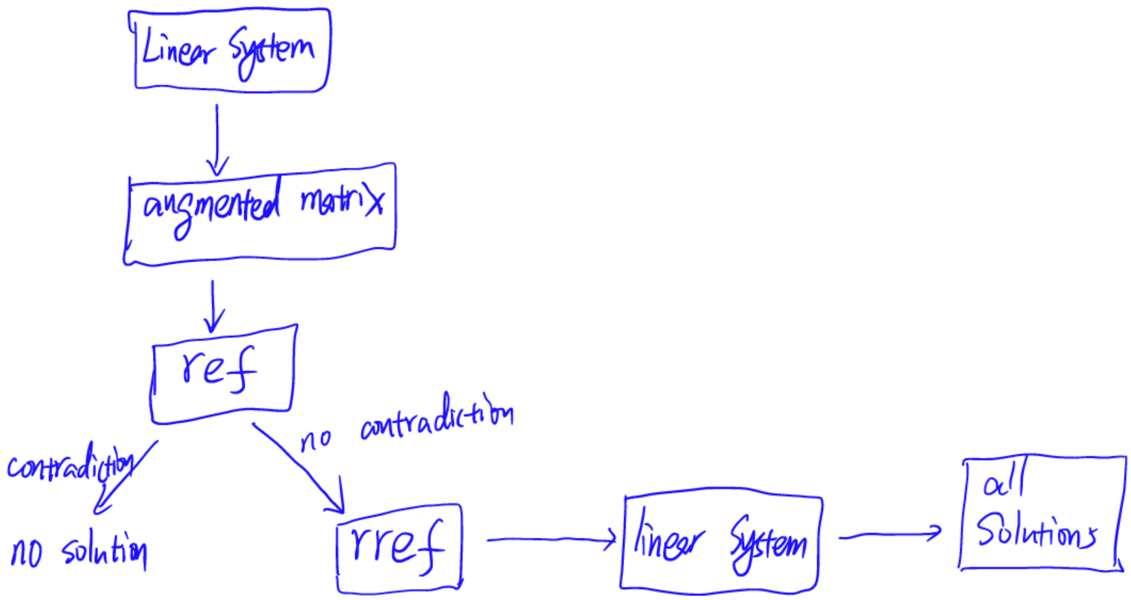
2. Select a *nonzero* entry in the pivot column as a pivot. (If necessary, interchange a pair of rows to move the pivot into the pivot position of that pivot column.)

3. Use row replacement operations to create zeros in all positions below the pivot.

4. Apply steps 1–3 to the matrix obtained by deleting all rows above the row containing the pivot from Step 3. Repeat this process until there are no more nonzero rows remaining.

5. Beginning with the **rightmost** pivot and working upward and to the left, create zeros above each pivot (using row replacement). If a pivot is not equal to 1, use scaling to change it to 1.

Procedure of solving a linear system.



Now, we can find all solutions for the consistent linear system using the Gauss-Jordan elimination method.

We will deal with the inconsistent case using least squares method later in chapter 5.

Example 2 : Using Row Reduction Algorithm to solve a linear system.

$$\begin{cases} 3x_2 - 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases}$$

sketch solutions

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & 7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$ augmented matrix

$$\longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$ row-echelon form (ref) not unique

$$\longrightarrow \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$ reduced row-echelon form (rref) unique

$$\begin{cases} x_1 - 2x_3 + 3x_4 = -24 \\ x_2 - 2x_3 + 2x_4 = -7 \\ x_5 = 4 \end{cases}$$

$$\begin{cases} x_1 = 2x_3 - 3x_4 - 24 \\ x_2 = -2x_3 + 2x_4 - 7 \\ x_5 = 4 \end{cases}$$

x_3, x_4 are free variables

The parametric vector form of all solutions is

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_3 - 3x_4 - 24 \\ -2x_3 + 2x_4 - 7 \\ x_3 \\ x_4 \\ 4 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -24 \\ 0 \\ 0 \\ 0 \\ 4 \end{bmatrix}$$