Lecture 2

(2.0) Recall: last time we learnt how to encode a linear system into a matrix, called <u>the augmented matrix</u> of the system. We listed <u>three elementary operations</u> which do not change the set of solutions of a linear system:

	Linear Systems	$\longleftrightarrow$ Matrices	
Swap	$E_i \leftrightarrow E_j$	$R_i \leftrightarrow R_j$	Elementary Operations
Scale	$E_i \to \alpha E_i$	$R_i \to \alpha R_i$	
Combine	$E_i \to E_i + \lambda E_j \quad (j \neq i)$	$R_i \to R_i + \lambda R_j$	

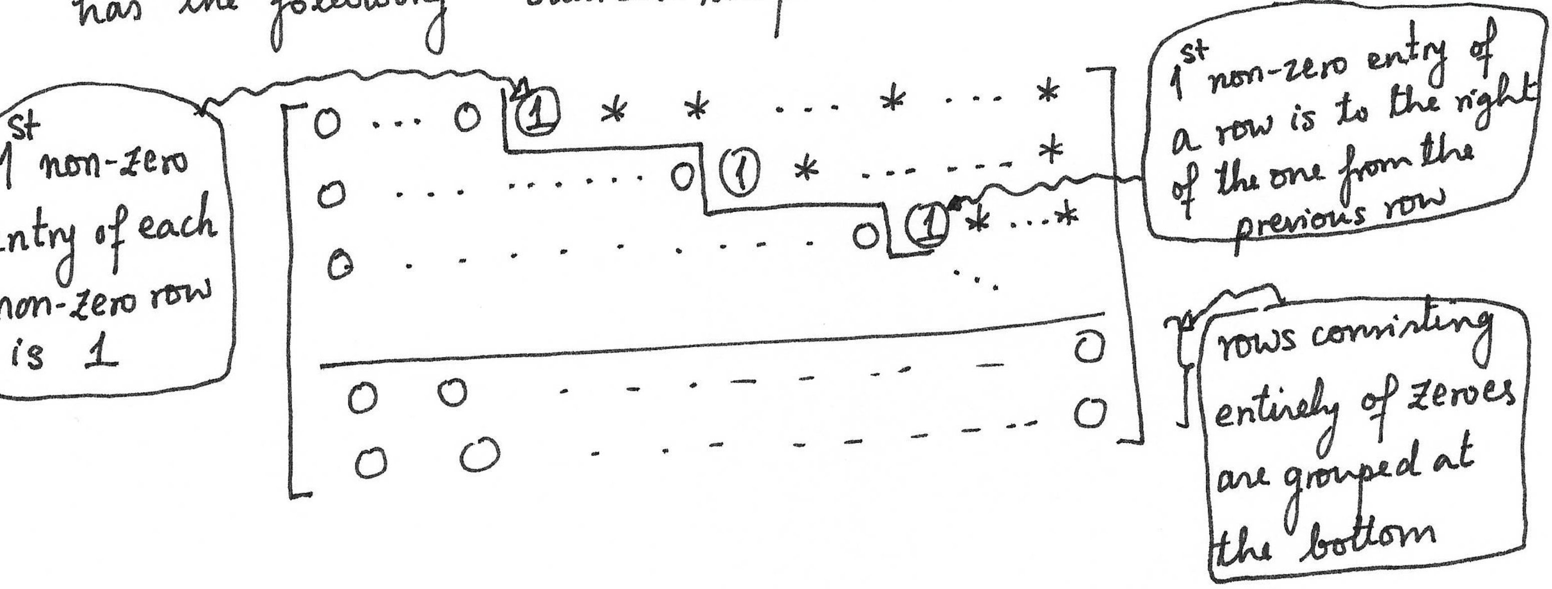
Here:
- $E_1, E_2, \dots, E_m$ are the linear equations our system is composed of.
- $R_1, R_2, \dots, R_m$ are the rows of an $(m \times n)$-matrix.
- $\alpha \neq 0$ and λ arbitrary numbers

Strategy to solve a linear system: simplify the system by performing the elementary operations. Solve the simpler system.

By "simpler system" we mean a linear system whose augmented matrix is in <u>echelon form</u>.

Today we will learn that <u>this strategy always works</u>. That is, <u>any matrix can be put in echelon form</u> using the three elementary operations.

(2.1) Echelon and reduced echelon forms. Recall from the last lecture that a matrix B is in echelon form if it has the following "staircase shape"



We say B is in a reduced echelon form if

(i) B is in echelon form; and

(ii) if $b_{ij} = 1$ is the 1st non-zero entry of row R_i then it is the only non-zero entry of the j^{th} column C_j.

e.g. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is in reduced echelon form.

$\begin{bmatrix} 1 & 5 & 0 & 7 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is also in reduced echelon form.

$B = \begin{bmatrix} 0 & 1 & 4 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is in echelon, but not in reduced echelon form because $b_{15} = 3 \neq 0$

$$\xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 0 & 1 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ reduced ✓

(2.2) <u>Theorem.</u> Given any matrix B, we can put B in a reduced echelon form using the three elementary operations.

This is achieved via two protocals. (or algorithms).

I. Input : an $(m\times n)$-matrix B.
Output : $(m\times n)$-matrix $\mathcal{E}(B)$ in echelon form, obtained from B by a sequence of elementary operations.

Step 0. If B is 0 (i.e., if all entries of B are 0's) then return $\mathcal{E}(B) = 0$.

Step 1. Find smallest i such that $b_{i1} \neq 0$.

- if there is no such b_{i1}, then the 1^{st} column of B consists entirely of 0's. That is, $B = \left[\begin{array}{c|c} 0 & \\ 0 & B' \\ \vdots & m\times(n-1) \\ 0 & \text{matrix} \end{array}\right]$

In this case, run the protocol for the (smaller) matrix B' to find $\mathcal{E}(B')$ and return

$$\mathcal{E}(B) = \left[\begin{array}{c|c} 0 & \\ 0 & \mathcal{E}(B') \\ \vdots & \\ 0 & \end{array}\right]$$

Step 2. - Let $\alpha = b_{i1}$ $(\neq 0)$. Swap $R_1 \leftrightarrow R_i$ and Scale $R_1 \to \frac{1}{\alpha} R_1$.

(after this step, (1^{st} row, 1^{st} column)-entry $= 1$)

Step 3. For each $j = 2, \dots, m$; Combine

$$R_j \longrightarrow R_j - b_{j1} R_1 .$$

Step 4. Now our matrix has the following form:

$$B = \left[\begin{array}{c|c} 1 & * \quad * \quad \cdots \quad * \\ \hline 0 & \\ 0 & \boxed{\begin{matrix} B'' \\ (m-1)\times(n-1) \text{ size} \\ \text{matrix} \end{matrix}} \\ \vdots & \\ 0 & \end{array}\right]$$

Run the protocol on the smaller matrix B'' to get $\mathcal{E}(B'')$.

Return $\quad \mathcal{E}(B) = \left[\begin{array}{c|c} 1 & * \quad * \quad \cdots \quad * \\ \hline 0 & \\ 0 & \boxed{\quad \mathcal{E}(B'') \quad} \\ \vdots & \\ 0 & \end{array}\right].$

II. Input : an $(m\times n)$ matrix C in <u>an echelon form</u>.

Output : $(m\times n)$ matrix $R\mathcal{E}(C)$ in a <u>reduced echelon form</u>, obtained from C by elementary operations.

For this, we scan the matrix from <u>bottom to top</u>.

Step 0. If $C = 0$, Return $R\mathcal{E}(C) = 0$.
(If C is empty, return $R\mathcal{E}(C) = $ empty.)

Step 1. Find largest k such that k^{th} row is non-zero.
Let ℓ be such that $C_{k,\ell} = 1$ is the 1^{st} non-zero entry of $\cancel{0}$ the k^{th} row of C.

Step 2. For each $i = 1, 2, \ldots, k$:
(Combine) $R_i \to R_i - c_{i\ell}\cdot R_k$

Now our matrix is in the form

$$C = \left[\begin{array}{ccccccc} 1 & & & 0 & * & \cdots & * \\ & 1 & & \vdots & & & \\ 0\text{'s} & & 1 & 0 & * & \cdots & * \\ \hline 0 & \cdots & 0 & 1 & * & \cdots & * \\ \hline 0 & & & \cdots & & & 0 \\ 0 & & & \uparrow & & & 0 \end{array}\right] \begin{array}{l} \\ \\ \\ \leftarrow k^{th} \text{ row} \\ \\ \\ \end{array}$$

ℓ^{th} column

Let C' be the top $(k-1)\times n$ matrix

Step 3. Repeat the protocol on C' to get $RE(C')$

return
$$RE(C) = \begin{array}{c} \left[\begin{array}{ccccccc} & & & RE(C') & & & \\ \hline 0 & \cdots & 0 & 1 & * & \cdots & * \\ \hline 0 & & & \cdots & & & 0 \\ 0 & & & \cdots & & & 0 \end{array}\right] \end{array} \begin{array}{l} \\ \leftarrow k^{th} \text{ row of } C \\ \\ \\ \end{array}$$

(2.3) Example. (i) Let $B = \begin{bmatrix} 1 & 3 \\ 4 & 10 \end{bmatrix}$.

$\mathcal{E}(B)$: $\begin{bmatrix} 1 & 3 \\ 4 & 10 \end{bmatrix} \xrightarrow[R_2 \to R_2 - 4R_1]{} \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} \xrightarrow[R_2 \to \frac{-1}{2} R_2]{} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} = \mathcal{E}(B)$ ✓

$R\mathcal{E}(B)$: $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \xrightarrow[R_1 \to R_1 - 3R_2]{} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = R\mathcal{E}(B)$ ✓

(ii) Let $B = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 7 & 8 \\ -1 & 1 & -5 & -5 \end{bmatrix}$

$\mathcal{E}(B)$: $\begin{bmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 7 & 8 \\ -1 & 1 & -5 & -5 \end{bmatrix} \xrightarrow[\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 + R_1}]{} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & -3 & 9 & 6 \\ 0 & 2 & -6 & -4 \end{bmatrix}$

$\xrightarrow[R_2 \to \frac{-1}{3} R_2]{} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -3 & -2 \\ 0 & 2 & -6 & -4 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 2R_1]{} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathcal{E}(B)$ ✓

$R\mathcal{E}(B)$: $\begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[R_1 \to R_1 - R_2]{} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R\mathcal{E}(B)$ ✓

(2.4) Linear system represented by a matrix in reduced echelon form.

- If it contains the following row - $[0\ 0 \cdots 0 \mid 1]$ then the corresponding linear system is inconsistent (i.e., has no solutions).
- The variables which correspond to columns containing pivoted 1's (corners of our staircase) are called dependent variables.
- The variables that are not dependent are called independent (or free parameters).

e.g. Take the matrix $B = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & -1 & 7 & 8 \\ -1 & 1 & -5 & -5 \end{array}\right]$ from the last page.

$$RE(B) = \begin{array}{c} \begin{array}{cccc} 1^{st} & 2^{nd} & & \\ \downarrow & \downarrow & & \end{array} \\ \left[\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right] \end{array}$$

pivoted 1's

. We conclude immediately that

(i) The system $\boxed{\begin{array}{l} x_1 + x_2 - x_3 = 1 \\ 2x_1 - x_2 + 7x_3 = 8 \\ -x_1 + x_2 - 5x_3 = -5 \end{array}}$ is consistent.

(ii) x_3 is a free parameter. Write the system represented by

$RE(B)$: $\boxed{\begin{array}{l} x_1 \qquad + 2x_3 = 3 \\ \qquad x_2 - 3x_3 = -2 \end{array}}$

(iii) For every choice of $x_3 = t$ $\boxed{\begin{array}{l} x_1 = 3 - 2t \\ x_2 = -2 + 3t \end{array}}$ is a solution.

(2.5) Example: Consider the system
$$\begin{aligned} x_1 + x_2 - x_5 &= 1 \\ x_2 + 2x_3 + x_4 + 3x_5 &= 1 \\ x_1 - x_3 + x_4 + x_5 &= 0. \end{aligned}$$

Determine whether this system is consistent or not. If yes, describe its set of solutions.

Sol.:
$$B = \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 1 & 0 & -1 & 1 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 - R_1]{} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 0 & -1 & -1 & 1 & 2 & -1 \end{array}\right] \xrightarrow[R_3 \to R_3 + R_2]{} \underbrace{\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 5 & 0 \end{array}\right]}_{\mathcal{E}(B)}$$

$$\xrightarrow[R_2 \to R_2 - 2R_3]{} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & -2 & 1 \\ 0 & 0 & 1 & 2 & 5 & 0 \end{array}\right] \xrightarrow[R_1 \to R_1 - R_2]{} \underbrace{\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & -2 & 1 \\ 0 & 0 & 1 & 2 & 5 & 0 \end{array}\right]}_{R\mathcal{E}(B)}$$

The system is consistent ✓.

x_4, x_5 are free parameters.

$$\boxed{\begin{aligned} x_1 + x_4 + x_5 &= 0 \\ x_2 - x_4 - 2x_5 &= 1 \\ x_3 + 2x_4 + 5x_5 &= 0 \end{aligned}}$$
system corresponding to $R\mathcal{E}(B)$

for every choice $\boxed{\begin{aligned} x_4 &= s \\ x_5 &= t \end{aligned}}$ of 2 numbers $\leadsto$ $\boxed{\begin{aligned} x_1 &= -s-t \\ x_2 &= 1+s+2t \\ x_3 &= -2s-5t \end{aligned}}$ a solution.

(2.6) Some terminology.–

We say two linear systems are <u>equivalent</u> if they have the <u>same set of solutions</u>.

Two matrices B_1 and B_2 are said to be <u>row equivalent</u> if we can go from B_1 to B_2 by performing elementary row operations. Recall (from Lecture 1) that the <u>row operations are reversible</u>. Thus, for example, if $\underline{RE(B_1) = RE(B_2)}$ then B_1 and B_2 are <u>row equivalent</u>.

$$B_1 \xrightarrow[\text{row operations}]{} RE(B_1) = RE(B_2) \xleftarrow[\text{row operations}]{} B_2$$

e.g. let $B_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B_2 = \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix}$

$$B_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow[R_2 \to R_2 - 3R_1]{} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \xrightarrow[R_2 \to -\frac{1}{2}R_2]{} \underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}}_{E(B_1)} \xrightarrow[R_1 \to R_1 - 2R_2]{} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = RE(B_1)$$

$$\|$$

$$B_2 = \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix} \xrightarrow[R_2 \to R_2 + R_1]{} \begin{bmatrix} 1 & 3 \\ 0 & 8 \end{bmatrix} \xrightarrow[R_2 \to \frac{1}{8}R_2]{} \underbrace{\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}}_{E(B_2)} \xrightarrow[R_1 \to R_1 - 3R_2]{} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = RE(B_2)$$

So B_1 and B_2 are <u>row equivalent</u>.