## Chapter 9

## The General Linear Systems

### 9.1 Row Echelon Form

Recall the nonsingular linear-systems problem

$$
A \cdot \mathbf{x}=\mathbf{b}
$$

where $A \in \mathbb{R}^{n \times n}$ is a square, regular matrix and $\mathbf{b} \in \mathbb{R}^{n}$ is known. To solve this problem, we transformed this system into an equivalent problem

$$
U \cdot \mathbf{x}=\mathbf{y}
$$

where $U \in \mathbb{R}^{n \times n}$ is an upper-triangular matrix with nonzero diagonal elements. To transform our original system into our desired form, we multiply both sides of our equation on the left by a sequence of $t \in \mathbb{N}$ elementary matrices

$$
U=E_{t} \cdots E_{2} \cdot E_{1} \cdot A, \quad \mathbf{y}=E_{t} \cdots E_{2} \cdot E_{1} \cdot \mathbf{b}
$$

To find our solution, we apply Backward Substitution to solve the resulting system.
While this algorithm works well to solve square linear-systems problems with regular coefficient matrices, not all linear systems problems can be solve using this technique. In this section, we generalize our strategy to develop a general Gaussian Elimination algorithm that can be used to solve all types of linear systems problem.

## Definition 9.1: The General Linear-Systems Problem

Let $m, n \in \mathbb{N}$. Let $A \in \mathbb{R}^{m \times n}$ be a given rectangular matrix and $\mathbf{b} \in \mathbb{R}^{m}$ be a given vector. Then the general linear-systems problem is to find an unknown vector $\mathrm{x} \in \mathbb{R}^{n}$ such that

$$
A \cdot \mathbf{x}=\mathbf{b}
$$

Let's begin our discussion by considering general linear-systems problems with rectangular coefficient matrices. Suppose we are given matrix $A \in \mathbb{R}^{m \times n}$ and vector $\mathbf{b} \in \mathbb{R}^{m}$ that define the linear-systems problem

$$
A \cdot \mathrm{x}=\mathbf{b}
$$

The strategy we use for solving our general linear-systems problem is to replace our original "hard" problem with an equivalent "easier" problem

$$
U \cdot \mathbf{x}=\mathbf{y}
$$

We do so in such a way that our transformed system has the exact same solution set as our original system. We create matrix $U \in \mathbb{R}^{m \times n}$ and vector $\mathbf{y} \in \mathbb{R}^{m}$ using a series of elementary row operations on matrix $A$ and vector $\mathbf{b}$. In particular, we transform $A$ and $\mathbf{b}$ by multiplying on the left-hand side by a series of elementary matrices.

In our previous section, we worked to transform our regular matrix $A$ into uppertriangular form. In our more general situation, we will instead work to transform $A$ into row echelon or reduced row echelon form.

## Definition 9.2: Row Echelon Form

Let $U \in \mathbb{R}^{m \times n}$ be a given matrix. We say that $U$ is in row echelon form if and only if $U$ satisfies the following two conditions
i. All zeros rows are below all nonzero rows
ii. The column index of the first nonzero entry in a row is larger than the column index of the first nonzero entry in any previous row.
iii. All entries in a column below an leading entry are zeros.

An intuitive, though imprecise, way to think about row echelon form is a matrix $U$ with all zero rows of $U$ are found at the "bottom" of the matrix and the first non-zero entry in any row is to the right of the first nonzero entries in the rows above it. While the row echelon form of a matrix $U$ is extremely helpful to evince information about the solution to a given linear-systems problem $U \mathbf{x}=\mathbf{y}$, we have a more powerful tool in our toolset.

## Definition 9.3: Reduced Row Echelon Form

Let $A \in \mathbb{R}^{m \times n}$ be a given matrix. We say that $A$ is in reduce row echelon form if and only if $A$ satisfies the following three conditions
i. $A$ is in row echelon form
ii. All leading entries of the rows of $A$ are equal to 1
iii. For any column that includes a leading entry, all other coefficients in that column are zero

## EXAMPLE 9.1.1

Let's look at some examples of matrices in row echelon form. To do so, we will highlight the sparsity structure of these matrices, focusing on the position of zero and nonzero entries to study the general patterns of matrices in row echelon form. In each of the matrices below, we denote leading entries with the bullet symbol •, representing any nonzero real number. On the other hand, the starred entries $\times$ represent any real number including zero. Zero entires are marked as such. Let's begin with three matrix structures that satisfy the conditions of row echelon form.

$$
\left[\begin{array}{lll}
\star & \times & \times \\
0 & \star & \times \\
0 & 0 & \star
\end{array}\right],\left[\begin{array}{llllll}
\star & \times & \times & \times & \times & \times \\
\times \\
0 & 0 & \star & \times & \times & \times \\
0 & 0 & 0 & 0 & 0 & \star \\
0 & 0 & 0 & 0 & 0 & 0
\end{array} \star\left\{\begin{array}{lllllllllll}
0 & \star & \times & \times & \times & \times & \times & \times & \times & \times & \times \\
0 & 0 & \star & \times & \times & \times & \times & \times & \times & \times & \times \\
0 & 0 & 0 & 0 & \star & \times & \times & \times & \times & \times & \times \\
0 & 0 & 0 & 0 & 0 & \star & \times & \times & \times & \times & \times \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \star & \times & \times \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]\right.
$$

In each of these matrices, we highlight the region in blue the entries of the matrix that may be nonzero including all leading entries. As is evident from this highlighting, matrices in row echelon form have a neat stair-step structure. Taking the general row echelon structures from above, we can refine these matrices into reduce row echelon form. In each case, we force the leading entries in each row to be one. Further, we ensure that any column with a leading entry has one, and only one, nonzero entry with all other coefficients in that column being set to zero. Below are the same matrix structures transformed into their corresponding reduced row echelon form:

$$
\left[\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{array}\right],\left[\begin{array}{ccccccc}
1 & \times & 0 & \times & \times & 0 & 0 \\
0 & 0 & 1 & \times & \times & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right],\left[\begin{array}{ccccccccccc}
0 & 1 & 0 & \times & 0 & 0 & \times & \times & 0 & \times & \times \\
0 & 0 & 1 & \times & 0 & 0 & \times & \times & 0 & \times & \times \\
0 & 0 & 0 & 0 & 1 & 0 & \times & \times & 0 & \times & \times \\
0 & 0 & 0 & 0 & 0 & 1 & \times & \times & 0 & \times & \times \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \times & \times \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

We work to transform $A$ into either row echelon form or reduced row echelon form. As we will see, if we are lucky enough to have a square coefficient matrix $A \in \mathbb{R}^{n \times n}$ in our original linear system, such transformations result in an uppertriangular matrix or diagonal matrix $U$. Moreover, solutions to linear-systems problems with upper-triangular coefficient matrices can be quickly computed using a special technique known as backward substitution. Let's begin with a neat example.

## EXAMPLE 9.1.2

Suppose we are hired by Road and Track Magazine to analyze the acceleration performance of a Tesla Model S. As part of our work, we conduct an acceleration experiment and collect two data points:

| Index $i$ | Time $t_{i}$ in seconds (s) | Velocity $v_{i}$ in mph |
| :---: | :---: | :---: |
| 1 | $t_{1}=0.5$ | $v_{1}=10$ |
| 2 | $t_{2}=2.5$ | $v_{2}=50$ |

After careful analysis, we decide that for the domain $0 \leq t \leq 14$, we want to fit our two data points with a quadratic function as illustrated below:


In this case, the coefficients $a_{0}, a_{1}, a_{2}$ of our quadratic model our unknown an desired. However, we can impose the conditions that our desired model fit the observed behavior exactly. To this end, we set up a system of 2 equations with 3 unknowns given by

$$
\begin{aligned}
& v\left(t_{1}\right)=v(0.5)=a_{0}+0.5 a_{1}+0.25 a_{2}=10=v_{1} \\
& v\left(t_{2}\right)=v(2.5)=a_{0}+2.5 a_{1}+6.25 a_{2}=40=v_{2}
\end{aligned}
$$

We can write this as a general linear-systems problem $A \mathbf{x}=\mathbf{b}$ in the following way

$$
\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
1 & 2.5 & 6.25
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{l}
10 \\
40
\end{array}\right]
$$

To solve this general system, we transform the coefficient matrix $A$ into matrix $U \in \mathbb{R}^{2 \times 3}$ in RREF. We begin by identifying our first pivot.

## STEP 1: Identify the first pivot

Identify the entry with row index 1 and column index 1 . Make sure this entry is nonzero and refer to this nonzero entry as the first pivot of our matrix. Refer to column 1 as the first pivot column. If the first entry is zero, multiply this matrix on the left by a permutation matrix $P_{i k}$ to swap rows and place a nonzero entry in this position.

In this case we see $a_{11}=1$ is nonzero and we call this entry our first pivot.

$$
\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
1 & 2.5 & 6.25
\end{array}\right]
$$

Because our first pivot is in column 1, we say that $A(:, 1)$ is the first pivot column of our coefficient matrix $A$. The next step of our elimination algorithm is to zero out all nonpivot entries of our first pivot column.

STEP 2: Create zeros in all entries below the first pivot
Multiply the original system of equation by a sequence of shear matrices to introduce zeros in all entries below our first pivot.

In this application, we introduce a zero value into entry $a_{21}$ in the first column

$$
\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
1 & 2.5 & 6.25
\end{array}\right]
$$

We transform all entries below this pivot into zero via left multiplication by the appropriate shear matrix. We see that the linear combination $-1 \cdot A(1,:)+A(2,:)$ is given by:

$$
-1 \cdot\left[\begin{array}{lll}
1 & 0.5 & 0.25
\end{array}\right]+\left[\begin{array}{lll}
1 & 2.5 & 6.25
\end{array}\right]=\left[\begin{array}{lll}
0 & 2.0 & 6.00
\end{array}\right]
$$

This yields a zero in entry $(2,1)$, as desired. To accomplish this transformation, we left multiply both sides of our general linear-systems problem by shear matrix $S_{21}(-1)$ as follows

$$
\begin{array}{cc}
{\left[\begin{array}{rr}
1 & 0 \\
-1 & 1
\end{array}\right]\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
1 & 2.5 & 6.25
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{rr}
1 & 0 \\
-1 & 1
\end{array}\right]\left[\begin{array}{l}
10 \\
40
\end{array}\right]} \\
\Longrightarrow & {\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
0 & 2.0 & 6.00
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{l}
10 \\
30
\end{array}\right]}
\end{array}
$$

STEP 3: Identify next pivot and introduce zeros under pivot

Move to the next row down and next column to the right. Call the first nonzero entry in this row the second pivot. The column that this pivot is in is called the second pivot column.

In our new, equivalent system we identify entry $(2,2)$ as our second pivot

$$
\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
0 & 2.0 & 6.00
\end{array}\right]
$$

We now want introduce zeros below pivot 2 in the second pivot column. Since our original coefficient matrix $A$ only has two rows, we are done with this step.

In the case of a square linear-system problem, our reduction algorithm would end here and we would solve the equivalent upper-triangular system using backward substitution. However, for this general linear-systems problem, we see that our equivalent system is not as easy to solve as our upper-triangular case due to the extra column. With this in mind, we continue our elimination procedure to transform our coefficient matrix into RREF.

STEP 4: Turn all pivots to one
Transform each pivot to one using the appropriate dilation matrix.

Next, we will transform the second pivot in entry $(2,2)$ into one by multiplying the second row by the scalar $c=0.5$. In other words, we will multiply our system by $D_{2}(0.5)$ resulting in the transformed system

$$
\begin{gathered}
{\left[\begin{array}{cc}
1 & 0 \\
0 & 0.5
\end{array}\right]\left[\begin{array}{lll}
1 & 0.5 & 0.25 \\
0 & 2.0 & 6.00
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{cc}
1 & 0 \\
0 & 0.5
\end{array}\right]\left[\begin{array}{l}
10 \\
30
\end{array}\right]} \\
{\left[\begin{array}{rrr}
1 & 0.5 & 0.25 \\
0 & 1 & 3.00
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{l}
10 \\
15
\end{array}\right]}
\end{gathered}
$$

This guarantees that all pivots in our system are equal to one, which is one of the conditions for a matrix in RREF. The next condition is that each pivot column has a unique nonzero entry.

STEP 5: Cancel out all nonpivot entries in each pivot column
Use left multiplication by shear matrices to cancel all nonpivot elements in each pivot column.

In this case, we see that in the second pivot column we have a single nonzero entry above pivot 2

$$
\left[\begin{array}{ccc}
1 & 0.5 & 0.25 \\
0 & 1 & 3.00
\end{array}\right]
$$

We can cancel out the entry in position $(1,2)$ using a shear matrix mulitplication

$$
\begin{array}{cc}
{\left[\begin{array}{cc}
1 & -0.5 \\
0 & 1
\end{array}\right]\left[\begin{array}{rrr}
1 & 0.5 & 0.25 \\
0 & 1 & 3.00
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{cc}
1 & -0.5 \\
0 & 1
\end{array}\right]\left[\begin{array}{l}
10 \\
15
\end{array}\right]} \\
\Longrightarrow & {\left[\begin{array}{rrr}
1 & 0 & -1.25 \\
0 & 1 & 3.00
\end{array}\right]\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2}
\end{array}\right]=\left[\begin{array}{l}
2.5 \\
15
\end{array}\right]}
\end{array}
$$

Now, we have an equivalent general linear-system problem $U \mathbf{x}=\mathbf{y}$, where the matrix $U=\operatorname{RREF}(A)$. In the next section, we discuss how to create the solution set to this system.

The algorithm above works in general. The general form of the matrix at the end of this discussion is known as reduced row echelon form of a matrix.

## Definition 9.4: Elementary Row Operations on Linear Systems

1. Replace an equation by the sum of itself and a scalar multiple times another equation
2. Interchange two equations
3. Multiply an equation by a nonzero constant


## Definition 9.5: Elementary Matrices for Row Operations

1. Shear Matrix $S_{i k}(c)$ : Replace a row $k$ by the sum of itself and a scalar multiple $c$ times row $k$
2. Transposition Matrix $P_{i k}$ : Interchange rows $i$ and $k$
3. Dilation Matrix $D_{i}(c)$ : Multiply row $i$ by constant $c$

When transforming a matrix $A$ into matrix $U$ in either row echelon form or reduced row echelon form, we multiply $A$ on the left by a sequence $t$ of elementary matrices

$$
\underbrace{E_{t} \cdot E_{t-1} \cdots E_{2} \cdot E_{1}}_{E} \cdot A=U
$$

In this case, each elementary matrix $E_{j}$ is either a shear, transposition or dilation matrix. As we will see, these elementary matrices have very beautiful properties and ensure that our resulting linear systems problem has an identical solution set.

## Lesson 16: Row Echelon Form- Suggested Problems

1. Determine if the given matrix is in RREF, in REF or neither. For each matrix in listed below, please
i. Write the matrix out fully (don't be lazy)
ii. If the matrix is in RREF or REF, annotate each matrix to highlight important properties of the definition of REF or RREF
iii. If the matrix is not in RREF nor REF, describe in detail which properties of the definition of REF or RREF that the matrix does NOT satisfies.

Think of your solutions as a study guide for the next exam. Be sure to write as much detail as possible. Also, make your solutions neat, organized and easy to read.
A. $\left[\begin{array}{rrrr}1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0\end{array}\right]$
B. $\left[\begin{array}{rrrr}1 & -4 & 3 & 9 \\ 0 & 2 & -2 & -7 \\ 0 & 0 & 2 & 1\end{array}\right]$ C. $\left[\begin{array}{rrrrrr}1 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & -3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0\end{array}\right]$
D. $\left[\begin{array}{rrrr}1 & 4 & 0 & 2 \\ 0 & 0 & 1 & -1\end{array}\right]$
E. $\left[\begin{array}{ll}3 & 2 \\ 0 & 1 \\ 0 & 0\end{array}\right]$
F. $\left[\begin{array}{llllll}1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1\end{array}\right]$
2. Transform each matrix below into RREF using two different methods
i. Multiply the given matrix on the left by a sequence of elementary matrices $E_{1}, E_{2}, \ldots, E_{t}$ (where each $E_{i}$ is either a shear, dilation or transposition matrix). For examples on how to carry out these steps, see the blue boxes titled "STEP 1" through "STEP 5" on p. 15-17 in your textbook.
ii. Transform using a TI Calculator (see video: https://youtu.be/JKJ461c0k7c)
A. $\left[\begin{array}{rrrr}1 & -3 & 2 & -1 \\ 2 & -6 & 4 & -2 \\ 3 & 9 & 6 & -4\end{array}\right]$
B. $\left[\begin{array}{rrrrrr}0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15\end{array}\right]$
3. Model for potato gun.

