Lecture Notes: Solving Linear Systems with Gauss Elimination

Yufei Tao
Department of Computer Science and Engineering
Chinese University of Hong Kong
taoyf@cse.cuhk.edu.hk

1 Echelon Form and Elementary Row Operations

Let \( B \) be an \( m \times n \) matrix. We say that \( B \) is in row echelon form if it satisfies all of the following conditions:

- If \( B \) has rows consisting of only 0’s, such rows appear consecutively at the bottom of \( B \).
- For \( i \in [1, m - 1] \), the leftmost non-zero element of the \( i \)-th row is at a column that is strictly to the left of the column containing the leftmost non-zero element of the \((i + 1)\)-th row.

For example, matrices

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 5 & 6 & 7 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad \begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 0 & 6 & 7 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad \begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 0 & 6 & 7 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

are all in row echelon form, but

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 2 & 3 & 4 \\
0 & 0 & 0 & 0
\end{bmatrix}, \quad \begin{bmatrix}
0 & 2 & 6 & 7 \\
0 & 0 & 1
\end{bmatrix}, \quad \begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 0 & 6 & 7
\end{bmatrix}
\]

are not.

We define three elementary row operations on \( B \):

1. Switch two rows of \( B \).
2. Multiply all numbers of a row by the same non-zero value.
3. Let \( r_i \) and \( r_j \) be two row vectors of \( B \). Update row \( r_i \) to \( r_i + cr_j \), where \( c \) can be any real value.

Any matrix \( B \) can be converted into a matrix in row echelon form by performing only elementary row operations. We demonstrate the steps using an example.

**Example 1.** We will convert the matrix below into row echelon form:

\[
\begin{bmatrix}
0 & 3 & 0 & 4 \\
2 & 1 & 6 & 3 \\
1 & 0 & 5 & 1 \\
0 & 8 & 3 & 2
\end{bmatrix}
\]

First, switch the rows so that the leftmost non-zero element of any row starts at a column that is the same or to the left of the column containing the leftmost non-zero element of the next row. The following is a matrix satisfying the condition:

\[
\begin{bmatrix}
2 & 1 & 6 & 3 \\
1 & 0 & 5 & 1 \\
0 & 3 & 0 & 4 \\
0 & 8 & 3 & 2
\end{bmatrix}
\]
Let $r_1$, $r_2$, ..., $r_4$ be the 1st, 2nd, ..., and 4th rows, respectively. Our next goal is to convert the first element of $r_2$, $r_3$, and $r_4$ to 0. Rows $r_3$ and $r_4$ already satisfy the condition. As for $r_2$, we can make it satisfy the condition by replacing it with $\frac{-1}{2}r_1 + r_2$, which gives the following matrix:

$$
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & -0.5 & 2 & -0.5 \\
0 & 3 & 0 & 4 \\
0 & 8 & 3 & 2 \\
\end{bmatrix}
$$

Henceforth, we will not touch the first row any more. Our next goal is to convert the second element of $r_3$ and $r_4$ to 0. Regarding $r_3$, this can be achieved by replacing it with $6r_2 + r_3$, leading to:

$$
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & -0.5 & 2 & -0.5 \\
0 & 0 & 12 & 1 \\
0 & 8 & 3 & 2 \\
\end{bmatrix}
$$

Similarly, replacing $r_4$ with $16r_2 + r_4$ gives:

$$
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & -0.5 & 2 & -0.5 \\
0 & 0 & 12 & 1 \\
0 & 0 & 35 & -6 \\
\end{bmatrix}
$$

Henceforth, we will not touch the first two rows any more. Our next goal is to convert the third element of $r_4$ to 0, as can be achieved by replacing it with $-\frac{35}{12}r_3 + r_4$, giving:

$$
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & -0.5 & 2 & -0.5 \\
0 & 0 & 12 & 1 \\
0 & 0 & 0 & -107/12 \\
\end{bmatrix}
$$

The matrix is now in row echelon form.

\section{Matrix Form of Linear Equations}

Consider that we have a system of line equations (such as a system is called a linear system):

\begin{align*}
    a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\
    &... \\
    a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m
\end{align*}

Note that the system has $m$ equations about $n$ variables $x_1, ..., x_n$. If we introduce:

$$
A = \begin{bmatrix}
a_{11} & a_{12} & ... & a_{1n} \\
a_{21} & a_{22} & ... & a_{2n} \\
... \\
a_{n1} & a_{n2} & ... & a_{mn} \\
\end{bmatrix}, \quad x = \begin{bmatrix}
x_1 \\
x_2 \\
... \\
x_n \\
\end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix}
b_1 \\
b_2 \\
... \\
b_m \\
\end{bmatrix}
$$

2
then we can concisely represent the linear system with matrix multiplication:

\[ Ax = b. \]

If \( b = 0 \), we say that the system is *homogeneous system*; otherwise, it is *nonhomogeneous system*. If the system has at least one solution, we say that the system is *consistent*; otherwise, it is *inconsistent*.

We define the *augmented matrix* of \( A \), denoted as \( \tilde{A} \), by including \( b \) into \( A \) as the last column, namely:

\[
\tilde{A} = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
\vdots & & & & \\
a_{m1} & a_{m2} & \cdots & a_{mn} & b_n
\end{bmatrix}
\]

Note that the vertical bar between the last two columns is just a reminder that this is an augmented matrix; the bar can be omitted if as desired. It is obvious that a linear system uniquely corresponds to an augmented matrix, and vice versa.

**Example 2.** Consider the following linear system:

\[
x_1 + 2x_2 + 3x_3 = 4 \\
2x_1 - x_2 - 2x_3 = 2
\]

The corresponding augmented matrix is:

\[
\tilde{A} = \begin{bmatrix}1 & 2 & 3 & 4 \\2 & -1 & -2 & 2\end{bmatrix}
\]

**3 Gauss Elimination**

Suppose that we are given a linear system \( Ax = b \). Let \( \tilde{A} \) be the augmented matrix of \( A \). Consider that we perform elementary row operations to convert \( \tilde{A} \) into another matrix \( \tilde{A}' \). The linear system corresponding to \( \tilde{A}' \) has *exactly the same solutions* as the linear system corresponding to \( \tilde{A} \). In other words, elementary row operations do not change the solutions of a linear system. We say that \( \tilde{A} \) and \( \tilde{A}' \) are *row equivalent*.

**Example 3.** Consider the augmented matrix \( \tilde{A} \) shown in Example 2. All the following matrices are row equivalent to \( \tilde{A} \) (think: which elementary row operations were used to derive them?):

\[
\begin{bmatrix}2 & -1 & -2 & 2 \\1 & 2 & 3 & 4\end{bmatrix}, \begin{bmatrix}2 & -1 & -2 & 2 \\2 & 4 & 6 & 8\end{bmatrix}, \begin{bmatrix}2 & -1 & -2 & 2 \\4 & 3 & 4 & 10\end{bmatrix}
\]

Note that the last matrix corresponds to the following linear system:

\[
2x_1 - x_2 - 2x_3 = 2 \\
4x_1 + 3x_2 + 4x_3 = 6
\]

Verify that this system has the same solutions as the system in Example 2.
Motivated by the above observation, we can solve the linear system $Ax = b$ by converting it to another linear system $A'x = b'$ whose augmented matrix is in row echelon form, as demonstrated in the next few examples.

**Example 4.** Consider the following linear system:

\[
\begin{align*}
3x_2 & = 4 \\
2x_1 + x_2 + 6x_3 & = 3 \\
x_1 + 5x_3 & = 1 \\
8x_2 + 3x_3 & = 2.
\end{align*}
\]

*Solution.* The augmented matrix of the linear system is matrix (1), which can be converted to the (row-equivalent) matrix in (2) of row echelon form, as shown in Example 1. (2) is the augmented matrix of the following linear system:

\[
\begin{align*}
2x_1 + x_2 + 6x_3 & = 3 \\
(-0.5)x_2 + 2x_3 & = -0.5 \\
12x_3 & = 1 \\
0 & = -107/12.
\end{align*}
\]

The system clearly has no solution. \(\blacksquare\)

**Example 5.** Consider the following linear system:

\[
\begin{align*}
3x_2 & = 4 \\
2x_1 + x_2 + 6x_3 & = 3 \\
x_1 + 5x_3 & = 1.
\end{align*}
\]

*Solution.* The augmented matrix of the linear system is

\[
\begin{bmatrix}
0 & 3 & 0 & 4 \\
2 & 1 & 6 & 3 \\
1 & 0 & 5 & 1
\end{bmatrix}
\]

which can be converted to the following matrix of row echelon form

\[
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & -0.5 & 2 & -0.5 \\
0 & 0 & 12 & 1
\end{bmatrix}
\]

This matrix is the augmented matrix of the following linear system:

\[
\begin{align*}
2x_1 + x_2 + 6x_3 & = 3 \\
(-0.5)x_2 + 2x_3 & = -0.5 \\
12x_3 & = 1.
\end{align*}
\]

Now we can do back substitution to obtain a unique solution. First, (5) gives $x_3 = 1/12$. Then, substituting this into (4), we get $x_2 = 4/3$. Finally, substituting the values of $x_2$ and $x_3$ into (3), we get $x_1 = 7/12$. \(\blacksquare\)
Example 6. Consider the following linear system:

\[
\begin{align*}
3x_2 &= 4 \\
2x_1 + x_2 + 6x_3 &= 3 \\
4x_1 + 5x_2 + 12x_3 &= 10
\end{align*}
\]

Solution. The augmented matrix of the linear system is

\[
\begin{bmatrix}
0 & 3 & 0 & 4 \\
2 & 1 & 6 & 3 \\
4 & 5 & 12 & 10
\end{bmatrix}
\]

which can be converted to the following matrix of row echelon form

\[
\begin{bmatrix}
2 & 1 & 6 & 3 \\
0 & 3 & 0 & 4 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

This matrix is the augmented matrix of the following linear system:

\[
\begin{align*}
2x_1 + x_2 + 6x_3 &= 3 \\
3x_2 &= 4
\end{align*}
\]

The system has infinitely many solutions.

The above method is called Gauss elimination. From the earlier examples, we can see that a linear system may have

- no solution—in this case, we say that the system is over-determined;
- a unique solution—in this case, we say that the system is determined;
- infinitely many solutions—in this case, we say that the system is under-determined.