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 $\boldsymbol{B}$ be an $m \times n$ matrix. We say that $\boldsymbol{B}$ is in *row echelon form* if it satisfies all of the following conditions:

- If $\boldsymbol{B}$ has rows consisting of *only* 0's, such rows appear consecutively at the bottom of $\boldsymbol{B}$.
- 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 $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 6 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 6 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ are all in row echelon form, but $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 3 & 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 2 & 6 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 6 & 7 \end{bmatrix}$ are not.

We define three *elementary row operations* on $\boldsymbol{B}$:

1. Switch two rows of $\boldsymbol{B}$.
2. Multiply all numbers of a row by the same non-zero value.
3. Let $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ be two row vectors of $\boldsymbol{B}$. Update row $\boldsymbol{r}_i$ to $\boldsymbol{r}_i + c\,\boldsymbol{r}_j$, where c can be any real value.

Any matrix $\boldsymbol{B}$ 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:

$$\begin{bmatrix} 0 & 3 & 0 & 4 \\ 2 & 1 & 6 & 3 \\ 1 & 0 & 5 & 1 \\ 0 & 8 & 3 & 2 \end{bmatrix} \tag{1}$$

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 $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, ..., $\boldsymbol{r_4}$ be the 1st, 2nd, ..., and 4th rows, respectively. Our next goal is to convert the first element of $\boldsymbol{r_2}$, $\boldsymbol{r_3}$, and $\boldsymbol{r_4}$ to 0. Rows $\boldsymbol{r_3}$ and $\boldsymbol{r_4}$ already satisfy the condition. As for $\boldsymbol{r_2}$, we can make it satisfy the condition by replacing it with $-\frac{1}{2}\boldsymbol{r_1} + \boldsymbol{r_2}$, 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 $\boldsymbol{r_3}$ and $\boldsymbol{r_4}$ to 0. Regarding $\boldsymbol{r_3}$, this can be achieved by replacing it with $6\boldsymbol{r_2} + \boldsymbol{r_3}$, 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 $\boldsymbol{r_4}$ with $16\boldsymbol{r_2} + \boldsymbol{r_4}$ 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 $\boldsymbol{r_4}$ to 0, as can be achieved by replacing it with $-\frac{35}{12}\boldsymbol{r_3} + \boldsymbol{r_4}$, giving:

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

The matrix is now in row echelon form. $\square$

2 Matrix Form of Linear Equations

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

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

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

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

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

$$\boldsymbol{Ax} = \boldsymbol{b}.$$

If $\boldsymbol{b} = \boldsymbol{0}$, 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 $\boldsymbol{A}$, denoted as $\tilde{\boldsymbol{A}}$, by including $\boldsymbol{b}$ into A as the last column, namely:

$$\tilde{\boldsymbol{A}} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array}\right]$$

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:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ 2x_1 - x_2 - 2x_3 &= 2 \end{aligned}$$

The corresponding augmented matrix is:

$$\tilde{\boldsymbol{A}} = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & -1 & -2 & 2 \end{array}\right]$$

□

3 Gauss Elimination

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

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

$$\left[\begin{array}{ccc|c} 2 & -1 & -2 & 2 \\ 1 & 2 & 3 & 4 \end{array}\right], \left[\begin{array}{ccc|c} 2 & -1 & -2 & 2 \\ 2 & 4 & 6 & 8 \end{array}\right], \left[\begin{array}{ccc|c} 2 & -1 & -2 & 2 \\ 4 & 3 & 4 & 10 \end{array}\right]$$

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

$$\begin{aligned} 2x_1 - x_2 - 2x_3 &= 2 \\ 4x_1 + 3x_2 + 4x_3 &= 6 \end{aligned}$$

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 $\boldsymbol{Ax} = \boldsymbol{b}$ by converting it to another linear system $\boldsymbol{A'x} = \boldsymbol{b'}$ whose augmented matrix is in row echelon form, as demonstrated in the next few examples.

Example 4. Consider the following linear system:

$$\begin{aligned} 3x_2 &= 4 \\ 2x_1 + x_2 + 6x_3 &= 3 \\ x_1 + 5x_3 &= 1 \\ 8x_2 + 3x_3 &= 2. \end{aligned}$$

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:

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

The system clearly has no solution. □

Example 5. Consider the following linear system:

$$\begin{aligned} 3x_2 &= 4 \\ 2x_1 + x_2 + 6x_3 &= 3 \\ x_1 + 5x_3 &= 1. \end{aligned}$$

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:

$$2x_1 + x_2 + 6x_3 = 3 \quad (3)$$
$$(-0.5)x_2 + 2x_3 = -0.5 \quad (4)$$
$$12x_3 = 1. \quad (5)$$

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 $x_1 = 7/12$. □

Example 6. Consider the following linear system:

$$
\begin{aligned}
3x_2 &= 4 \\
2x_1 + x_2 + 6x_3 &= 3 \\
4x_1 + 5x_2 + 12x_3 &= 10
\end{aligned}
$$

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:

$$
\begin{aligned}
2x_1 + x_2 + 6x_3 &= 3 \\
3x_2 &= 4
\end{aligned}
$$

The system has infinitely many solutions. $\square$

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*.