Linear Algebra I

Elementary row operations

1 Augmented matrix

A *solution* to the system of linear equations over the real numbers[1]

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\ \cdots \quad & \quad \cdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

in terms of the matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & & \cdots & \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

and vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$

is a vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

such that

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

In order to find the solution set of the given system of equations by Gaussian elimination we form the *augmented matrix*

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \cdots & & \cdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right] \in \mathbb{R}^{m\times(n+1)}.$$

(The vertical line is there to remind us that it represents a system of linear equations, but can be dropped. We can simply write $[A\ \ \mathbf{b}]$, as in Strang's textbook.)

[1] Sometimes a different field to $\mathbb{R}$ is considered: for example, the set of rational numbers $\mathbb{Q}$, or the set of complex numbers $\mathbb{C}$. Most results for linear equations over $\mathbb{R}$ carry over to these other fields, as the relevant properties of $\mathbb{R}$ that are used, such as existence of multiplicative inverses (the element t^{-1} such that $t^{-1}t = 1 = tt^{-1}$ for $t \neq 0$) or distributivity of multiplication over addition ($a(b+c) = ab + ac$), are properties shared by any field. Later in the course we shall look at what makes a field a field (in the mathematical sense!).

2 Elementary row operations

We describe elementary row operations on a general $m \times n$ matrix

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \cdots & & \cdots & \\ c_{m,1} & c_{m,2} & \cdots & c_{m,n} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

In practice, C will be the augmented matrix $[A \mid \mathbf{b}]$ for a system of linear equations.

Proposition 2.1. *The two operations of (1) scalar multiplication of row i by $t \neq 0$ and (2) replacing row i by its sum with row j,*

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \end{bmatrix} \longmapsto \begin{bmatrix} \cdots & & \cdots & \\ tc_{i,1} & tc_{i,2} & \cdots & tc_{i,n} \\ \cdots & & \cdots & \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \longmapsto \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1}+c_{j,1} & c_{i,2}+c_{j,2} & \cdots & c_{i,n}+c_{j,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \tag{2}$$

can be composed to produce the operations of (3) adding a scalar multiple of row j to row i and (4) swapping rows i and j:[2]

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \longmapsto \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1}+tc_{j,1} & c_{i,2}+tc_{j,2} & \cdots & c_{i,n}+tc_{j,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \longmapsto \begin{bmatrix} \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \end{bmatrix} \tag{4}$$

Proof. We write $C \underset{(1)}{\sim} C'$ if the matrix C' is obtained from C by an application of operation (1) multiplying a row by a non-zero scalar, and likewise $C \underset{(2)}{\sim} C'$ of C' is obtained from C by an application of operation (2) swapping two rows.

[2] In Sage operation (3) is given by `add_multiple_of_row(`i`, `j`, `t`)` and operation (4) by `swap_rows(`i`, `j`)`. In Sage rows and columns are counted starting at zero: for the first row $i = 0$ and the first column $j = 0$. E.g. for the 2nd row and 3rd column you should enter $i = 1$ and $j = 2$ in these Sage commands. For an example worksheet see `http://arcikam.kam.mff.cuni.cz:12080/home/pub/1/`

The operation (3) is for $t \neq 0$ realized by the sequence:

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ tc_{j,1} & tc_{j,2} & \cdots & tc_{j,n} \\ \cdots & & \cdots & \end{bmatrix}$$

$$\underset{(2)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} + tc_{j,1} & c_{i,2} + tc_{j,2} & \cdots & c_{i,n} + tc_{j,n} \\ \cdots & & \cdots & \\ tc_{j,1} & tc_{j,2} & \cdots & tc_{j,n} \\ \cdots & & \cdots & \end{bmatrix}$$

$$\underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} + tc_{j,1} & c_{i,2} + tc_{j,2} & \cdots & c_{i,n} + tc_{j,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix}$$

where in the last line we have used the fact that $t \neq 0$ has a multiplicative inverse (so that we can return row j to its original value by dividing through by t).

(When $t = 0$ operation (3) does nothing to the matrix, so this case does not need to be considered.)

There are other sequence of operations (1) and (2) that can produce the same effect as operation (3), but the one above involving three steps is the shortest.

The operation (4), swapping rows, is, for example, realized by the sequence:

$$
\begin{aligned}
\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix}
&\underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ -c_{i,1} & -c_{i,2} & \cdots & -c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \\
&\underset{(2)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ -c_{i,1} & -c_{i,2} & \cdots & -c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1}-c_{i,1} & c_{j,2}-c_{i,2} & \cdots & c_{j,n}-c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \\
&\underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1}-c_{i,1} & c_{j,2}-c_{i,2} & \cdots & c_{j,n}-c_{i,n} \\ \cdots & & \cdots & \end{bmatrix} \\
&\underset{(2)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \\ c_{j,1}-c_{i,1} & c_{j,2}-c_{i,2} & \cdots & c_{j,n}-c_{i,n} \\ \cdots & & \cdots & \end{bmatrix} \\
&\underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1}-c_{j,1} & c_{i,2}-c_{j,2} & \cdots & c_{i,n}-c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \\
&\underset{(2)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n}- \\ \cdots & & \cdots & \end{bmatrix}
\end{aligned}
$$

There are other sequence of operations (1) and (2) that can produce the same effect as operation (4): have we found the shortest such sequence here?

Note: having already proved that operation (3) is given by a sequence of operations of types (1) and (2), we could give a shorter proof for operation (4) as follows (in which $C \underset{(3)}{\sim} C'$ means

C' is obtained by applying operation (3) to C):

$$\begin{bmatrix} \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix} \underset{(3)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1}-c_{j,1} & c_{i,2}-c_{j,2} & \cdots & c_{i,n}-c_{j,n} \\ \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \end{bmatrix}$$

$$\underset{(2)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{i,1}-c_{j,1} & c_{i,2}-c_{j,2} & \cdots & c_{i,n}-c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \end{bmatrix}$$

$$\underset{(3)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ -c_{j,1} & -c_{j,2} & \cdots & -c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \end{bmatrix}$$

$$\underset{(1)}{\sim} \begin{bmatrix} \cdots & & \cdots & \\ c_{j,1} & c_{j,2} & \cdots & c_{j,n} \\ \cdots & & \cdots & \\ c_{i,1} & c_{i,2} & \cdots & c_{i,n} \\ \cdots & & \cdots & \end{bmatrix}$$

□

3 Elementary operations on the augmented matrix for a system of linear equations

Lemma 3.1. *Each of the elementary row operations (1) and (2) preserve the solution sets of* $A\mathbf{x} = \mathbf{b}$. *That is, if*

(a) $[A \mid \mathbf{b}] \underset{(1)}{\sim} [A' \mid \mathbf{b}']$ *, or*

(b) $[A \mid \mathbf{b}] \underset{(2)}{\sim} [A' \mid \mathbf{b}']$

then the solution set of $A\mathbf{x} = \mathbf{b}$ *is equal to the solution set of* $A'\mathbf{x} = \mathbf{b}'$.

Proof. Both operations (1) and (2) just affect row i, the remaining rows staying the same. Hence the only difference between the system of equations $A\mathbf{x} = \mathbf{b}$ and the system of equations $A'\mathbf{x} = \mathbf{b}'$ is the ith equation. Let the (k,l)-entry of A be $a_{k,l}$ and the (k,l)-entry of A' be $a'_{k,l}$. Then $a'_{k,l} = a_{k,l}$ and $b'_k = b_k$ when $k \neq i$.

Let S be the solution set of $A\mathbf{x} = \mathbf{b}$ and let S' be the solution set of $A'\mathbf{x} = \mathbf{b}'$. In order to prove that $S = S'$ we show that

$(S \subseteq S')$ if $A\mathbf{x} = \mathbf{b}$ then $A'\mathbf{x} = \mathbf{b}'$, and

$(S' \subseteq S)$ if $A'\mathbf{x} = \mathbf{b}'$ then $A\mathbf{x} = \mathbf{b}$.

Since only the ith equations differ, all we need to prove is that

$(S \subseteq S')$ if $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$ then $a'_{i,1}x_1 + a'_{i,2}x_2 + \cdots + a'_{i,n}x_n = b'_i$, and

$(S' \subseteq S)$ if $a'_{i,1}x_1 + a'_{i,2}x_2 + \cdots + a'_{i,n}x_n = b'_i$ then $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$.

(a) Under operation (1) we have $a'_{i,l} = ta_{i,l}$ and $b'_i = tb_i$ and the ith equation in $A'\mathbf{x} = \mathbf{b}'$ is

$$ta_{i,1}x_1 + ta_{i,2}x_2 + \cdots + ta_{i,n}x_n = tb_i$$

while in $A\mathbf{x} = \mathbf{b}$ it is

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i.$$

$S \subseteq S'$: if $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$ then by multiplying by $t \in \mathbb{R}$ we obtain $t(a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n) = tb_i$, and this is the same as $ta_{i,1}x_1 + ta_{i,2}x_2 + \cdots + ta_{i,n}x_n = tb_i$, which is the ith equation of $A'\mathbf{x} = \mathbf{b}'$.

$S' \subseteq S$: if $ta_{i,1}x_1 + ta_{i,2}x_2 + \cdots + ta_{i,n}x_n = tb_i$ then, since $t \neq 0$, we can multiply by t^{-1} to obtain the equation $t^{-1}(ta_{i,1}x_1 + ta_{i,2}x_2 + \cdots + ta_{i,n}x_n) = t^{-1}(tb_i)$, which is the same as $t^{-1}ta_{i,1}x_1 + t^{-1}ta_{i,2}x_2 + \cdots + t^{-1}ta_{i,n}x_n = t^{-1}tb_i$, and since $t^{-1}t = 1$ this is the same as $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$, which is the ith equation of $A\mathbf{x} = \mathbf{b}$.

(b) Under operation (2) we have $a'_{i,l} = a_{i,l} + a_{j,l}$ and $b'_i = b_i + b_j$ and the ith equation in $A'\mathbf{x} = \mathbf{b}'$ is

$$(a_{i,1} + a_{j,l})x_1 + (a_{i,2} + a_{j,2})x_2 + \cdots + (a_{i,n} + a_{j,n})x_n = b_i + b_j$$

while in $A\mathbf{x} = \mathbf{b}$ it is

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i.$$

$S \subseteq S'$: if $a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$ and $a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n = b_j$ then by adding these two equalities we obtain

$$(a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n) + (a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n) = b_i + b_j,$$

and, collecting together coefficients of the x_l, this is the same as

$$(a_{i,1} + a_{j,1})x_1 + (a_{i,2} + a_{j,2})x_2 + \cdots + (a_{i,n} + a_{j,n})x_n = b_i + b_j,$$

which is the ith equation of $A'\mathbf{x} = \mathbf{b}'$.

$S' \subseteq S$: the ith and jth equations of $A'\mathbf{x} = \mathbf{b}'$ are

$$(a_{i,1} + a_{j,1})x_1 + (a_{i,2} + a_{j,2})x_2 + \cdots + (a_{i,n} + a_{j,n})x_n = b_i + b_j,$$

and

$$a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n = b_j$$

from which

$$(a_{i,1}+a_{j,1})x_1+(a_{i,2}+a_{j,2})x_2+\cdots+(a_{i,n}+a_{j,n})x_n-(a_{j,1}x_1+a_{j,2}x_2+\cdots+a_{j,n}x_n) = b_i+b_j-b_j,$$

Multiplying out the brackets, rearranging, and collecting together the coefficients of the x_l, this is to say

$$(a_{i,1} + a_{j,1} - a_{j,1})x_1 + (a_{i,2} + a_{j,2} - a_{j,2})x_2 + \cdots + (a_{i,n} + a_{j,n} - a_{j,n})x_n = b_i,$$

which is to say

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i,$$

which is the ith equation of $A\mathbf{x} = \mathbf{b}$.

$\square$

If there is a finite sequence of elementary row operations of types (1) and (2) that change matrix C to matrix C' then we write $C \sim C'$. (By Proposition 2.1 we can also use operations (3) and (4) in this definition.)

Proposition 3.2. *If* $[A \mid \mathbf{b}] \sim [A' \mid \mathbf{b}']$ *then the solution set of* $A\mathbf{x} = \mathbf{b}$ *is equal to the solution set of* $A'\mathbf{x} = \mathbf{b}'$.

Proof. By definition, if $[A \mid \mathbf{b}] \sim [A' \mid \mathbf{b}']$ then there is a finite sequence of elementary row operations of types (1) and (2) which transforms $[A \mid \mathbf{b}]$ into $[A' \mid \mathbf{b}']$.

We proceed by induction on the length of this sequence of elementary row operations. For length zero there is no operation at all, $[A \mid \mathbf{b}] = [A' \mid \mathbf{b}']$, and the assertion is trivial. (For length one, the statement is equivalent to Lemma 3.1 and hence also holds as a base case.)

Assume as induction hypothesis that when there is a sequence of at most r elementary row operations that transforms $[A \mid \mathbf{b}]$ into $[A' \mid \mathbf{b}']$ it is true that the solution set of $A\mathbf{x} = \mathbf{b}$ is equal to the solution set of $A'\mathbf{x} = \mathbf{b}'$. (We have verified the cases $r = 0$ and $r = 1$.)

Suppose $[A \mid \mathbf{b}] \sim [A' \mid \mathbf{b}']$ by a sequence of $r + 1$ elementary row operations. Let $[A'' \mid \mathbf{b}'']$ be the matrix obtained after applying the first r operations. Then $[A \mid \mathbf{b}] \sim [A'' \mid \mathbf{b}'']$ and by induction hypothesis the solution set S of $A\mathbf{x} = \mathbf{b}$ is equal to the solution set S'' of $A''\mathbf{x} = \mathbf{b}''$. Since we apply a single elementary row operation of type (1) or type (2) to transform $[A'' \mid \mathbf{b}'']$ into $[A' \mid \mathbf{b}']$, by Lemma 3.1 the solution set S'' of $A''\mathbf{x} = \mathbf{b}''$ is equal to the solution set S' of $A'\mathbf{x} = \mathbf{b}'$.

Since $S = S''$ and $S'' = S'$ it follows that $S = S'$, i.e., the solution set of $A\mathbf{x} = \mathbf{b}$ is equal to the solution set of $A'\mathbf{x} = \mathbf{b}'$. This proves the induction step and by mathematical induction the proof of the proposition is complete. □

A system of linear equations $A\mathbf{x} = \mathbf{b}$ represented by augmented matrix $[A \mid \mathbf{b}]$ can be reduced by elementary row operations to *row echelon form* (see lecture notes). In other words, there is an echelon form matrix $[A' \mid \mathbf{b}']$ which by Proposition 3.2 gives a system of linear equations $A'\mathbf{x} = \mathbf{b}'$ with the same solution set as the original system $A\mathbf{x} = \mathbf{b}$.

The Gaussian elimination algorithm converts a given matrix $[A \mid \mathbf{b}]$ by a sequence of operations of types (1) and (2) (in practice, a sequence of operations of types (3) and (4)) into an echelon form matrix $[A' \mid \mathbf{b}']$. Since *back-substitution* can be used to solve $A'\mathbf{x} = \mathbf{b}'$, using the Gauss elimination algorithm we can determine the solution set of any system of linear equations.

Furthermore, by Gauss-Jordan reduction there is a unique matrix $[A' \mid \mathbf{b}']$ in *reduced row echelon form* which is equivalent to $[A \mid \mathbf{b}]$ by elementary row operations. In this way, using Gauss-Jordan reduction you can determine whether a system of linear equations $A_1\mathbf{x} = \mathbf{b}_1$ has the same solution set as another system of linear equations $A_2\mathbf{x} = \mathbf{b}_2$, without having to determine their solution sets.