Gaussian Elimination

Systems of linear equations

A single equation of the form:

$$ax = b,$$

has a solution for all b if $a \neq 0$. The situation is not as obvious if we have many such equations in many real variables. Consider as a first example the following pair of equations:

$$\begin{array}{llll} 1x & +2y & +3z & = 4, \\ 2x & +4y & +6z & = 6. \end{array}$$

The equality of the left-hand side and the right-hand side of an equation is maintained if both sides are multiplied by the same number. Multiplying the first equation by 2, we get:

$$2x + 4y + 6z = 8,$$

and this is inconsistent with the second equation. Hence we see that this pair of equation has no solutions. If the constant on the right hand side of the first equation is 3, the first equation holds if and only if the second equation hold. As a result, and triple $(x, y, z) = (3 - 2y - 3z, y, z)$ gives a solution to the system.

Gaussian elimination provides a systematic approach to the number of solutions to linear systems of equations. A system of m linear equations in n real variables $(x_1, ..., x_n)$ is written as:

$$\begin{array}{cccccc} a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n & = & b_1, \\ a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n & = & b_2, \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & \cdots & +a_{mn}x_n & = & b_m. \end{array} \tag{1}$$

In matrix form this is:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

where

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Write the equations in such an order that $a_{11} \neq 0$. The first equation of (**??**) gives:

$$x_1 = \frac{b_1}{a_{11}} - \frac{a_{12}x_2}{a_{11}} - ... - \frac{a_{1n}x_n}{a_{11}}.$$

By substituting into the other $(m-1)$ equations, we get:

$$\begin{array}{ccccc} (a_{22} - a_{21}\frac{a_{12}}{a_{11}})x_2 & \cdots & + \left(a_{2n} - a_{21}\frac{a_{1n}}{a_{11}}\right) x_n & = & b_2 - a_{21}\frac{b_1}{a_{11}}, \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ (a_{m2} - a_{m1}\frac{a_{12}}{a_{11}})x_2 & \cdots & + \left(a_{mn} - a_{m1}\frac{a_{1n}}{a_{11}}\right) x_n & = & b_m - a_{m1}\frac{b_1}{a_{11}}. \end{array} \tag{2}$$

An equivalent way of writing this set of equations is to add to the k^{th} row of $\boldsymbol{A}$ the first row of $\boldsymbol{A}$ multiplied by $\frac{a_{k1}}{a_{11}}$ and add to b_k $a_{21}\frac{b_1}{a_{11}}$ to the k^{th} row of the column vector $\boldsymbol{b}$. My construction, the first element on each row $k > 1$ of the new matrix is zero and we can solve $(x_2, ..., x_n)$ from the new system with at most $m-1$ equations in at most $n-1$ variables.

If $(a_{k2} - a_{k1}\frac{a_{12}}{a_{11}}) \neq 0$ for some $k > 1$, we can repeat the previous step for system (**??**), i.e. solve x_2 and substitute into the other equations. If $(a_{k2} - a_{k1}\frac{a_{12}}{a_{11}}) = 0$ for all $k > 1$, eliminate the variable x_l where l is the smallest index for which $(a_{kl} - a_{k1}\frac{a_{1l}}{a_{11}}) \neq 0$ for some $k > 1$. Since after each such step we are left with a system that has at least one fewer equation than in the previous step, the process comes to an end in finitely many steps.

The key to this elimination process relies on two basic facts of arithmetics.

1. The solution to an equation is unchanged if both sides of the equation are multiplied by the same non-zero number.

2. If $(x_1, ..., x_n)$ satisfies

$$a_{11}x_1 \quad +a_{12}x_2 \quad \cdots \quad +a_{1n}x_n \quad = \quad b_1,$$

and

$$a_{21}x_1 \quad +a_{22}x_2 \quad \cdots \quad +a_{2n}x_n \quad = \quad b_2,$$

then $(x_1, ..., x_n)$ satisfies:

$$(a_{11} + a_{21})\, x_1 + ... + (a_{1n} + a_{2n})\, x_n = b_1 + b_2.$$

We call the two fundamental steps in this elimination process *elementary row operations*. They are:

i) Swapping equations in the system (i.e. swapping rows in the associated matrix).

ii) Summing multiples of an equation to another equation (ading a multiple of a row in the associated matrix to another row).

Solving systems of equations via elementary row operations

Homogenous systems

Consider the system of equations in matrix form

$$\boldsymbol{A}\boldsymbol{x} = 0,$$

where

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \; \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Since the right hand side of the equation is zero, this is called a homogenous system. It has always a trivial solution $x = (0, ..., 0)$, but we want to know if it has other solutions.

Start by considering matrix $\boldsymbol{A}$. If $a_{11} = 0$, swap row 1 with row k, where $a_{k1} \neq 0$. if no such row exists, the vector $(z, 0, ..., 0)$ satisfies the system of equations for all z and therefore the solution is not unique. In fact, x_1 is not really a variable in this system.

Assume next that for some k, $a_{k1} \neq 0$ and swap rows 1 and $k = 1$ if $a_{11} = 0$. Multiply the first row by $\frac{1}{a_{11}}$ and add this multiplied first row to each row $k > 1$ multiplied by $-\frac{a_{k1}}{a_{11}}$. We get the following new matrix

$$\boldsymbol{A}^{(1)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - a_{21}\frac{a_{12}}{a_{11}} & \cdots & a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} - a_{m1}\frac{a_{12}}{a_{11}} & \cdots & a_{mn} - a_{m1}\frac{a_{1n}}{a_{11}} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2}^{(1)} & \cdots & a_{mn}^{(1)} \end{pmatrix}.$$

If

$$a_{22} - a_{21}\frac{a_{12}}{a_{11}} \neq 0,$$

multiply the second row by

$$\frac{1}{a_{22} - a_{21}\frac{a_{12}}{a_{11}}},$$

and add the resulting second row multiplied by

$$-\frac{a_{k2} - a_{k1}\frac{a_{12}}{a_{11}}}{a_{22} - a_{21}\frac{a_{12}}{a_{11}}}$$

to each row $k > 2$.

If

$$a_{22} - a_{21}\frac{a_{12}}{a_{11}} = 0,$$

swap row 2 and k'' such that

$$a_{k''2} - a_{k''1}\frac{a_{12}}{a_{11}} \neq 0$$

and proceed as before. If $a_{k2}^{(1)} = 0$ for all $k \geq 2$, multiply the second row of $A^{(1)}$ by

$$\frac{1}{a_{23}^{(1)}}.$$

(or swap the rows if $a_{23}^{(1)} = 0$) and proceed as before.

This results in a new matrix

$$\boldsymbol{A}^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(2)} \end{pmatrix}.$$

By repeating the above steps, we get matrices $\boldsymbol{A}^{(3)}, \boldsymbol{A}^{(4)}$ etc. until after k eliminations, we get e.g. for $m = n = 5$,

$$\begin{pmatrix} a_{11}^{(5)} & a_{12}^{(5)} & \cdot & \cdot & a_{15}^{(5)} \\ 0 & a_{22}^{(5)} & a_{23}^{(5)} & \cdot & a_{25}^{(5)} \\ 0 & 0 & a_{33}^{(5)} & a_{34}^{(5)} & a_{35}^{(5)} \\ 0 & 0 & 0 & a_{44}^{(5)} & a_{45}^{(5)} \\ 0 & 0 & 0 & 0 & a_{55}^{(5)} \end{pmatrix},$$

or

$$\begin{pmatrix} a_{11}^{(4)} & a_{12}^{(4)} & \cdot & \cdot & a_{15}^{(4)} \\ 0 & a_{22}^{(4)} & a_{23}^{(4)} & \cdot & a_{25}^{(4)} \\ 0 & 0 & 0 & a_{34}^{(4)} & a_{35}^{(4)} \\ 0 & 0 & 0 & 0 & a_{45}^{(4)} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We say that a matrix $\boldsymbol{A}$ is in *row echelon form* if each row k has a larger number of initial zero elements than row $k-1$. Both of the matrices above are in row echelon form. All matrices can be transformed into row echelon form by elementary row operations.

The number of non-zero rows is the called the *row rank* of a matrix in row echelon form. The top matrix above has row rank 5 and the one below it has row rank 4.

Since each row in the row echelon form starts with more zeros than the previous row, the row rank is always less than or equal to the number of columns. If the row rank is equal to the number of columns, the only solution is the trivial solution $\boldsymbol{x} = 0$. If row rank is less than the number of columns, the system has infinitely many solutions.

In the first case above, the trivial solution is the only solution to the system. This can be seen as follows. The last row in the row echelon form implies that $x_5 = 0$ in any solution. Using this, the second to last row implies that $x_4 = 0$ etc.

In the second case above, the variable x_3 can be chosen freely. For each choice of x_3, the other variables are uniquely determined.

Non-homogenous systems

Consider next the system of n equations in n variables.

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

We will perform elementary row operations to transform $\boldsymbol{A}$ to its row echelon form. It is now useful to consider the *augmented matrix*:

$$\left(\boldsymbol{A}\vdots\boldsymbol{b}\right) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right).$$

We perform the elementary row operations on the entire matrix $\left(\boldsymbol{A}\vdots\boldsymbol{b}\right)$ to keep track of the right hand side. Obviously this was not necessary in the homogenous case where the right hand side is zero.

$$\left(\begin{array}{cccc|c} a_{11}^{(2)} & a_{12}^{(k)} & \cdots & a_{1n}^{(k)} & b_1^{(k)} \\ 0 & a_{22}^{(k)} & \cdots & a_{2n}^{(k)} & b_2^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array}\right).$$

If

$$\text{rank}\left(A\vdots b\right) = \text{rank}\,(A)\,,$$

then the system has a solution.

If

$$\text{rank}\left(A\vdots b\right) > \text{rank}\,(A)\,,$$

it has no solutions.

If

$$\text{rank}\left(A\vdots b\right) = \text{rank}\,(A) = n,$$

the solution is unique.

If

$$\text{rank}\left(A\vdots b\right) = \text{rank}\,(A) < n,$$

then the system has infinitely many solutions.

Examples of elementary row operations

- **Finding the row echelon form**

  Let

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

  Multiply first row by $-\frac{1}{2}$ and add to second and third row:

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & 2-\frac{1}{2} & 2+\frac{1}{2} \\ 0 & 0-\frac{1}{2} & 1+\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

  Multiply second row by $\frac{1}{3}$ and add to third row:

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 0 & \frac{3}{2}+\frac{5}{6} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 0 & \frac{7}{3} \end{pmatrix}.$$

  Since the row echelon form has rank 3, we know that the system

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

  has a unique solution for all b.

Solving a system of equations

Consider a numerical example for the previous system:

$$\begin{array}{cccccc} 2x_1 & +x_2 & -x_3 & & 2 \\ x_1 & +2x_2 & +2x_3 & = & 1 \\ x_1 & & +x_3 & & 0 \end{array}$$

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

The augmented matrix is now:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right).$$

Repeat the elementary row operations:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & \frac{3}{2} & \frac{5}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{3}{2} & -1 \end{array}\right)$$

and

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & \frac{3}{2} & \frac{5}{2} & 0 \\ 0 & 0 & \frac{7}{3} & -1 \end{array}\right).$$

We get:

$$x_3 = \frac{-3}{7}.$$

Substituting into the second row:

$$\frac{3}{2}x_2 + \frac{5}{2}\left(\frac{-3}{7}\right) = 0.$$

Hence:

$$x_2 = \frac{5}{7}.$$

The first row gives:

$$2x_1 + \frac{5}{7} - \frac{-3}{7} = 2$$

or

$$x_1 = \frac{3}{7}.$$

A matrix without full rank

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 4 \\ 1 & 1 & 3 \\ 2 & 1 & 5 \end{pmatrix}$$

Eliminate the first entry in the second and the third row by using the first row:

$$\begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

When eliminating the second entry on the third row by using the second, we get row echelon form:

$$\begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

We see that the matrix $\boldsymbol{A}$ has $rank(\boldsymbol{A}) = 2$ and the system of equations has either zero or infinitely many solutions.

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

I leave it as an exercise using the augmented matrix to show that the system has a solution only if

$$b_3 = b_2 + \frac{1}{2}b_1.$$

Since $\boldsymbol{A}$ does not have full rank, we know that the homogenous equation has non-trivial solutions (in this case any multiple of the vector $(-2, -1, 1)$ solves the system). If $\boldsymbol{x}^0$ is a solution to the homogenous eqyation and $\boldsymbol{x}^P$ is some solution to the non-homogenous system, then $\boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{x}^P$ is also a solution to the non-homogenous equation. You can see this by summing the homogenous and non-homogenous equations.

Invering a matrix via Gaussian elimination We compute $\boldsymbol{A}^{-1}$, for

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

To find the inverse matrix, we need to find three vectors $\boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3$ such that

$$\boldsymbol{A}\boldsymbol{x}^i = \boldsymbol{e}^i \text{ for } i \in \{1, 2, 3\},$$

where $\boldsymbol{e}^i$ is the i^{th} unit vector. To do this in one go, form the augmented matrix:

$$\left(\boldsymbol{A}\vdots\boldsymbol{I}\right) = \left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

After using Gaussian elimination to transform $\boldsymbol{A}$ into the identity matrix, we can read the $\boldsymbol{x}^i$ as the columns on the right hand side of the augmented matrix.

Multiply first row by $\frac{1}{2}$. Add the first row multiplied by $\frac{-1}{2}$ to the third row. At this point, we have:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{array}\right).$$

Multiply the second row by $\frac{1}{2}$ and add $\frac{3}{4}$ of the second row to the third row:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{5}{4} & -\frac{1}{2} & \frac{3}{4} & 1 \end{array}\right).$$

Multiply third row by $\frac{4}{5}$ and add the third row multiplied by $-\frac{2}{5}$ to the second and the first row:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{3}{2} & 0 & \frac{7}{10} & -\frac{3}{10} & -\frac{2}{5} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right).$$

Add $-\frac{3}{2}$ times the second row to the first row:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{5} & -\frac{3}{5} & \frac{1}{5} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right).$$

From this last augmented matrix, we read that:

$$\boldsymbol{A}^{-1} = \frac{1}{5}\begin{pmatrix} 2 & -3 & 1 \\ 1 & 1 & -2 \\ -2 & 3 & 4 \end{pmatrix}.$$