Notes 3 – Linear systems of equations

In this lecture, we shall introduce linear systems, interpret them in terms of matrices and vectors, and define linear combinations of vectors.

L3.1 Introduction. Here is an example of a so-called *linear system:*

$$\left\{\begin{array}{rcl} x_1 + x_2 + 2x_3 + 3x_4 &=& 3 \\ 5x_1 + 8x_2 + 13x_3 + 21x_4 &=& 7 \\ 34x_1 + 55x_2 + 89x_3 + 144x_4 &=& 4 \end{array}\right. \tag{1}$$

This one consists of 3 equations in 4 unknowns. The problem is to determine *all* possible values of these unknowns x_1, x_2, x_3, x_4 that solve all 3 equations simultaneously. Each equation might be expected to impose a single constraint, and since there are fewer equations than unknowns we might guess that it is not possible to specify the unknowns uniquely. However, without checking the numbers on the right, it is conceivable that the 3 equations are *inconsistent* and that there are *no* solutions.

The situation is best illustrated with pairs of equations, each in 2 unknowns. Consider the four separate systems

$$\begin{array}{llll} (a) & \left\{\begin{array}{l} x + 2y = 0 \\ 3x + 4y = 0 \end{array}\right. & (b) & \left\{\begin{array}{l} x + 2y = 7 \\ 3x + 4y = 8 \end{array}\right. \\ (c) & \left\{\begin{array}{l} x + 2y = 0 \\ 2x + 4y = 0 \end{array}\right. & (d) & \left\{\begin{array}{l} x + 2y = 7 \\ 2x + 4y = 8 \end{array}\right. \end{array} \tag{2}$$

Those on the left are called *homogeneous* because the numbers on the right are all zero, whereas (b) and (d) are *inhomogeneous* or *nonhomogeneous.* Any homogeneous system always has at least one solution, namely the one in which all the unknowns are assigned the value 0; this is called the *trivial solution.*

It is easy to check that in cases (a) and (b) the two equations are independent and that there is a unique solution of the system. For (a), it is the trivial solution $x = 0 = y$; for (b) it is $x = -6$, $y = 13/2$, or expressed more neatly $(x, y) = (-6, \frac{13}{2})$.

In (c) it is obvious that the second equation is completely redundant; it is merely twice the first. In this case, we can assign *any* value to (say) y and then declare that $x = -2y$; we say that y is a *free variable* and that the general solution depends on one *free parameter.* In a sense, the system (c) is 'underdetermined'.

In (d), the two equations are incompatible; the first would imply that $2x + 4y = 14$ and we get $14 = 8$. This means that there is no solution; the system is called *inconsistent.* By contrast, homogeneous equations are always consistent.

To sum up, we can have no solutions, a unique solution (one and only one value for each unknown) or infinitely many solutions. We shall see that the same is true for a linear system of arbitrary size. With this knowledge, and without further examination, we can be confident that (1) has either infinitely many solutions or not at all; it cannot have say exactly four solutions!

L3.2 Matrix form. Let us begin with an arbitrary linear system of the form

$$\left\{\begin{array}{rcl} a_{11}x_1 + \ldots\ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \ldots\ldots + a_{2n}x_n & = & b_2 \\ \ldots\ldots & & \\ \ldots\ldots & & \\ a_{m1}x_1 + \ldots\ldots + a_{mn}x_n & = & b_m \end{array}\right.$$

We shall always use

m to denote the number of equations, and

n to denote the number of unknowns or variables.

We now see that the notation is tailored to that of matrices; indeed the system can be rewritten in the succinct matrix form

$$A\mathbf{x} = \mathbf{b} \tag{3}$$

where

$$A = \begin{pmatrix} a_{11} & \cdot & \cdot & a_{1n} \\ \cdot & & & \cdot \\ a_{m1} & \cdot & \cdot & a_{mn} \end{pmatrix} \in \mathbb{R}^{m,n}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \in \mathbb{R}^{n,1}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \cdot \\ b_m \end{pmatrix} \in \mathbb{R}^{m,1}.$$

The system is homogeneous iff $\mathbf{b} = \mathbf{0}$ is the null vector. A *solution* of the system is now understood as meaning a column vector $\mathbf{x}$ of length n that satisfies (3). The problem is to find *all* such vectors.

Observe that in the matrix form (3), the left-hand side of each equation is translated into a *row* of A. We shall normally solve such a system by operating on the rows of A, but first we show how the inverse matrix can sometimes be used.

Example. Consider the linear system (3), and suppose that $m = n$ and that A is *invertible.* This means that we can find a matrix A^{-1} such that $A^{-1}A = I_n$. Then

$$A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} \quad \Rightarrow \quad (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b} \quad \Rightarrow \quad \mathbf{x} = A^{-1}\mathbf{b},$$

and the system is solved uniquely. Thus, *a linear system with the same number of equations and variables whose associated matrix is invertible has a unique solution.* Applying this method to the generic 2×2 system

$$\left\{\begin{array}{rcl} ax + by & = & p \\ cx + dy & = & q. \end{array}\right.$$

gives

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} dp - bq \\ -cp + aq \end{pmatrix}.$$

The solution is neatly expressed as

$$x = \frac{\left|\begin{pmatrix} p & b \\ q & d \end{pmatrix}\right|}{\left|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right|}, \qquad y = \frac{\left|\begin{pmatrix} a & p \\ c & q \end{pmatrix}\right|}{\left|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right|}.$$

It is a special case of *Cramer's rule,* whereby each unknown is obtained by susbstituting a column of A by $\mathbf{b}$, taking the determinant, and then dividing by $\det A$.

L3.3 Linear combinations. Let A be the matrix of left-hand coefficients defined by a linear system. We can instead emphasize the role played by the *columns* $\mathbf{c}_1,\dots\mathbf{c}_n$ of A by rewriting the system as

$$x_1\begin{pmatrix}a_{11}\\ \cdot\\ a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\ \cdot\\ a_{m2}\end{pmatrix}+\cdots+x_n\begin{pmatrix}a_{1n}\\ \cdot\\ a_{mn}\end{pmatrix}=\begin{pmatrix}b_1\\ \cdot\\ b_m\end{pmatrix}.$$

Equivalently,

$$x_1\mathbf{c}_1+x_2\mathbf{c}_2+\cdots+x_n\mathbf{c}_n=\mathbf{b}. \tag{4}$$

This is called the *column vector form* of the system. In this interpretation, the simultaneous nature of the m equations translates into a relation between the column vectors of length m involving the coefficients x_i. For example

$$-6\begin{pmatrix}1\\3\end{pmatrix}+\tfrac{13}{2}\begin{pmatrix}2\\4\end{pmatrix}=\begin{pmatrix}7\\8\end{pmatrix}$$

is the solution of example (2)(b) in these terms.

This motivates the

Definition. *Fix* n*, and let* $\mathbf{u}_1,\dots,\mathbf{u}_k$ *be finitely many vectors of length* n *(either all in* $\mathbb{R}^{1,n}$ *or all in* $\mathbb{R}^{n,1}$*). A* linear combination (LC) *of these vectors is any vector of the form*

$$a_1\mathbf{u}_1+\cdots+a_k\mathbf{u}_k$$

with $a_1,\dots,a_k\in\mathbb{R}$*. The set of all such linear combinations is written* $\mathscr{L}\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$.

Thus $\mathscr{L}\{\mathbf{u}_1,\cdots,\mathbf{u}_k\}$ is the set of vectors 'generated' by the $\mathbf{u}_i$. Often it is called their *span* and written $\langle\mathbf{u}_1,\dots,\mathbf{u}_k\rangle$. It is an example of a *subspace,* something that we shall study in a future lecture. It does not depend on the order in which the $\mathbf{u}_i$ are written; it is a function of the unordered set $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$, which mathematicians usually write with curly brackets.

Solving a linear system then amounts to trying to express the given vector $\mathbf{b}$ as a LC of the columns manufactured from the left-hand coefficients. A solution exists iff

$$\mathbf{b}\in\mathscr{L}\{\mathbf{c}_1,\dots,\mathbf{c}_n\}.$$

Whilst the rows of A represent the equations, *it is linear combinations of the columns that characterize the solutions*. In the study of linear systems, one is constantly torn between favouring the rows of the associated coefficient matrix, or the columns.

Exercise. Let $\mathbf{v}=\begin{pmatrix}1\\2\end{pmatrix}$, $\mathbf{w}=\begin{pmatrix}2\\4\end{pmatrix}$. Show that $\mathscr{L}\{\mathbf{v},\mathbf{w}\}=\mathscr{L}\{\mathbf{v}\}$ and that $\begin{pmatrix}x\\y\end{pmatrix}\in\mathscr{L}\{\mathbf{v}\}$ iff $y=2x$. The fact that $\begin{pmatrix}7\\8\end{pmatrix}\notin\mathscr{L}\{\mathbf{v}\}$ explains why the system (d) had no solution.

Example. Consider the row vectors $\mathbf{i}=(1,0,0)$, $\mathbf{j}=(0,1,0)$, $\mathbf{k}=(0,0,1)$. Then

$$\mathscr{L}\{\mathbf{i}\}=\{(x,0,0)\,:\,x\in\mathbb{R}\},$$
$$\mathscr{L}\{\mathbf{i},\mathbf{j}\}=\{(x,y,0)\,:\,x,y\in\mathbb{R}\}=\{(x,y,z)\,:\,z=0\}.$$

The last line shows a linear combination of vectors characterized by an equation, something we shall see over and over again. Note also that $\mathscr{L}\{\mathbf{i},\mathbf{0}\}=\mathscr{L}\{\mathbf{i}\}$ and $\mathscr{L}\{\mathbf{i},\mathbf{i}+\mathbf{j}\}=\mathscr{L}\{\mathbf{i},\mathbf{j}\}$.

L3.4 Further exercises.

1. Determine which of the following homogeneous systems admit *only* the trivial solution:

$$\begin{cases} 3x + y - z = 0 \\ x + y - 3z = 0 \\ x + y = 0, \end{cases} \qquad \begin{cases} -4x + 2y + z = 0 \\ 3x - 5y + z = 0 \\ 3x + y - 2z = 0, \end{cases} \qquad \begin{cases} -2x_1 + x_2 + x_3 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ x_1 + x_2 - 2x_3 = 0. \end{cases}$$

2. Find *all* the solutions of the linear systems

$$\begin{cases} 3x + y - z = 0 \\ x + y - 3z = 1 \\ x + y = -1, \end{cases} \qquad \begin{cases} -4x + 2y + z = 0 \\ 3x - 5y + z = 1 \\ 3x + y - 2z = -1, \end{cases} \qquad \begin{cases} -2x_1 + x_2 + x_3 = 0 \\ x_1 - 2x_2 + x_3 = 1 \\ x_1 + x_2 - 2x_3 = -1. \end{cases}$$

3. Given the row vectors $\mathbf{v}_1 = (a, b, c)$, $\mathbf{v}_2 = (1, 1, 0)$, $\mathbf{v}_3 = (0, 1, -1)$, $\mathbf{w} = (2, 3, -1)$, consider the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{w}. \tag{5}$$

Determine whether there exist $a, b, c \in \mathbb{R}$ such that

(i) equation (5) has a unique solution (x_1, x_2, x_3),

(ii) equation (5) has no solution,

(iii) equation (5) has infinitely many solutions.