Notes 5 – Reduced matrices and their rank

Solving a linear system by applying suitable ERO's is called *Gaussian elimination.* In this lecture we shall describe it more carefully, and present some variations. First we need to know exactly what type of matrices we want to achieve by operating on the rows.

L5.1 Echelon forms. The idea is to convert a matrix into a steplike sequence of rows like an arrangement of toy soldiers ready for battle. We shall make this precise using the

Definition. *An entry of a matrix is called a* marker *if it is the first nonzero entry of a row (starting from the left).*

Consider the condition

(M1) there is *at most* (meaning *not more than)* one marker in each column.

This is important for two reasons. First, because it turns out that the equations defined by the rows of a matrix satisfying (M1) are independent in a sense that we shall make precise in a moment. Second, because of the following result that becomes self-evident in the light of further practice:

Proposition. *Given any matrix A, one may apply a sequence of ERO's of type (i) so as to obtain a matrix B satisfying* (M1).

In solving the system, the order of the rows is immaterial. It is common practice to 'tidy up' by applying ERO's of type (iii) to permute the rows so that

(M2) moving down the rows from the top, the markers move from left to right, and

(M3) all the null rows are at the bottom.

It is a consequence of (M1) and (M2) that all the entries (in the same column) underneath a given marker are zero.

Definition. *A matrix satisfying* (M1)–(M3) *is called* step-reduced.

𝔚𝔞𝔯𝔫𝔦𝔫𝔤. This is sometimes referred to as 'row echelon form', but is a stronger notion than 'reduced' in the sense of the Greco-Valabrega text. We prefer to work with step-reduced matrices as they are easy to spot visually: the main null part has an approximately triangular form, and the markers represent 'corner soldiers'.

Once a matrix is step-reduced, its markers take on a greater significance and are also called *pivots;* often we shall $\boxed{\text{box}}$ them. We shall call a column *marked* if it contains a marker, and otherwise *unmarked.* It is sometimes convenient to suppose in addition that

(M4) each marker equals 1.

This can be quickly achieved using ERO's of type (ii), but may have the adverse effect of introducing fractions elsewhere. A more stringent condition is that

(M5) each marker is the only nonzero entry in its column (above as well as below).

Definition. *A matrix satisfying* (M1)–(M5) *is said to be* super-reduced.

A step-reduced matrix can be made super-reduced by a process called *backwards reduction.* Start with the last (bottom right) marker and subtract multiples of its row to remove all the entries above it. Then do the same with the second-to-last marker and so forth. (The word 'super' is a pun since it means both 'above' and 'extra'; some authors use the terminology 'reduced row echelon form' or RREF.)

Example. Consider again

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 5 & 8 & 13 & 21 \\ 34 & 55 & 89 & 144 \end{pmatrix}.$$

We have already step-reduced it to

$$B = \begin{pmatrix} \boxed{1} & 1 & 2 & 3 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

To super-reduce B, we merely perform the operation $\mathbf{r}_1 \mapsto \mathbf{r}_1 - \mathbf{r}_2$ so as to obtain

$$C = \begin{pmatrix} \boxed{1} & 0 & 1 & 1 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

When a matrix is super-reduced, it is possible to read off immediately the solution to the original system. In the above example, we get the equations

$$\begin{cases} x_1 \quad + x_3 + x_4 &= 0 \\ \quad x_2 + x_3 + 2x_4 &= 0, \end{cases}$$

whence

$$x_1 = -s - t, \quad x_2 = -s - 2t,$$

slightly more effectively than before.

L5.2 Linear independence. Converting a matrix to row echelon form has the effect of making the transformed equations 'independent'. We formulate this notion.

Definition. *Let* $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ *be a finite subset of* $\mathbb{R}^n$ *(meaning either in* $\mathbb{R}^{1,n}$ *or* $\mathbb{R}^{n,1}$*). The set is called* linearly independent (LI) *if the equation*

$$x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k = \mathbf{0} \tag{1}$$

admits only the trivial solution $x_1 = \cdots = x_k = 0$.

One often says '$\mathbf{u}_1, \cdots, \mathbf{u}_k$ are linearly independent', though strictly speaking being LI is a property of a set or list and not its individual elements. The order of the elements is immaterial, and any duplication prevents the list from being LI.

A singleton set $\{\mathbf{v}\}$ is LI iff $\mathbf{v} \neq \mathbf{0}$, and no set that contains $\mathbf{0}$ can be LI. A set $\{\mathbf{u}, \mathbf{v}\}$ is LI iff neither vector is a multiple (including zero times) the other. More generally,

Lemma. *A set* $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ *is LI iff no one element in it can be expressed as a linear combination of the others.*

Proof. Suppose that the set is LI, but that one of the elements is a LC of the others. For sake of argument, suppose that $\mathbf{u}_k \in \mathscr{L}\{\mathbf{u}_1, \cdots, \mathbf{u}_{k-1}\}$, so that $\mathbf{u}_k = a_1\mathbf{u}_1 + \cdots + a_{k-1}\mathbf{u}_{k-1}$ for some $a_i \in \mathbb{R}$. But then

$$a_1\mathbf{u}_1 + \cdots + a_{k-1}\mathbf{u}_{k-1} + (-1)\mathbf{u}_k = 0,$$

contradicting (1). The converse is similar. QED

Proposition. *Let B be a step-reduced matrix. Then its nonzero rows are LI.*

Proof. We refer to the example

$$\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array} \begin{pmatrix} 0 & \boxed{2} & 3 & 0 & 4 & 6 \\ 0 & 0 & \boxed{7} & 3 & 8 & 1 \\ 0 & 0 & 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B \tag{2}$$

assuming a linear relation

$$a_1\mathbf{r}_1 + \cdots + a_r\mathbf{r}_r = \mathbf{0}$$

between the nonzero rows. Perform the addition column by column, high school fashion. The marker of $\mathbf{r}_1$ is the only nonzero entry in its column, so $2a_1 = 0$. Passing to the next marker, $3a_1 + 7a_2 = 0$, so $a_2 = 0$ and so on. QED

> **Exercise.** The same conclusion holds if B satisfies just the first condition (M1), namely that there is at most one marker in each column. Indeed, once B satisfies (M1) we can make it step-reduced without changing the *unordered set* of rows.

The columns of the matrix in (2) are not LI. Even if we forget about $\mathbf{c}_1$ and $\mathbf{c}_4$, we have

$$\mathbf{c}_6 = 2\mathbf{c}_5 - \tfrac{15}{7}\mathbf{c}_3 - k\mathbf{c}_2,$$

for some $k \in \mathbb{R}$. This illustrates the

Proposition. *Let B be a step-reduced matrix. A column $\mathbf{c}_j$ of B is unmarked iff it is a LC of the previous marked columns (or null if $j = 1$).*

The point is that we can always express an unmarked column as a LC of the previous marked ones by finding the coefficients one at a time starting from the bottom.

Corollary. *If B is step-reduced then its* marked *columns are LI.*

Thus, *the markers of a step-reduced matrix 'mark out' an independent set of both rows and columns.* Whilst there may be unmarked columns in any position, row reduction ensures that all the unmarked rows are null. If *every* column of a step-reduced matrix C is marked, then the set of columns is LI and (from the column vector form of the system) $C\mathbf{x} = \mathbf{0}$ has only the trivial solution. Here is an example that makes this clear:

$$C = \begin{pmatrix} \boxed{1} & 2 & 3 & 1 \\ 0 & \boxed{1} & 2 & 1 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \left\{ \begin{array}{rcl} x + 2y + 3z + t & = & 0 \\ y + 2z + t & = & 0 \\ z + t & = & 0 \\ t & = & 0. \end{array} \right. \tag{3}$$

L5.3 The rank. This is defined to measure the *number* of independent rows or columns of *any* matrix, in a way we shall make precise. We first restrict to reduced matrices.

Definition. *Let B be step-reduced. Its* rank *is the number of markers, or equivalently nonzero rows. This number is denoted by* rank B, rk B *or* $r(B)$.

Example. The matrices in (2) and (3) have rank 3 and 4 respectively.

If B has size $m{\times}n$ then obviously $r(B) \leqslant \min\{m,n\}$. If $r(B)$ does not achieve this minimum, we can think of B as 'defective'. The importance of the rank derives from the

Theorem. *Suppose B and C are both step-reduced and that $B \sim C$. Then the markers of B and C occur in exactly the same positions. In particular, $r(B) = r(C)$.*

This enables us to define the *rank* of an arbitrary matrix A to be the rank of any reduced matrix B row equivalent to A. For if $A \sim B$ and $A \sim C$ with B and C step-reduced then

$$B \sim A \sim C \quad \Rightarrow \quad B \sim C,$$

and their ranks are equal.

Proof. It follows from what we have seen that a column $\mathbf{c}_j$ is unmarked iff

$$\mathbf{c}_j \in \mathscr{L}\{\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}\}. \tag{4}$$

This translates into a linear relation

$$a_1\mathbf{c}_1 + \ldots + a_{j-1}\mathbf{c}_{j-1} - \mathbf{c}_j = 0,$$

and asserts that $B\mathbf{x} = \mathbf{0}$ has a solution $\mathbf{x} = (a_1, \ldots, a_{j-1}, -1, 0 \ldots, 0)^T$ or equivalently one with $x_j = 1$ and $x_{j+1} = \cdots = x_n = 0$. But $B\mathbf{x} = \mathbf{0}$ and $C\mathbf{x} = \mathbf{0}$ are equivalent systems of equations, they have *identical* solutions and (from the column vector form) *identical* linear relations between their respective columns. QED

It is not hard to deduce an even stronger result of theoretical importance, namely

Corollary. *Any matrix is row equivalent to a* unique *super-reduced one.*

L5.4 Further exercises.

1. Use ERO's to reduce the matrices

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 8 & 10 & 12 \\ 1 & 1 & 1 \\ 9 & 7 & 6 \end{pmatrix}, \quad \begin{pmatrix} 4 & 2 & 1 & 0 \\ 3 & 7 & 0 & 5 \\ 0 & 8 & 4 & 2 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix},$$

and find their ranks.

2. Given $A = \begin{pmatrix} 0 & -5 & 1 \\ -1 & -1 & 0 \\ 0 & -5 & -3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 4 \end{pmatrix}$, decide which of the following are valid:

$$r(A) = r(B), \qquad r(A) = r(AB), \qquad r(B) = r(AB).$$

Find matrices $A, B \in \mathbb{R}^{3,3}$ for which $r(A) > r(A+B) > r(B)$.

3. Let $A = \begin{pmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 2 & 0 & 0 & a \end{pmatrix}$. Find the value of $r(A)$ as $a \in \mathbb{R}$ varies.