Linear Algebra & Geometry

LECTURE 10

- Matrices
- Elementary row operations (EROS)
- Rank of a matrix

Fact.

The $n \times n$ matrix I defined as $I(s,t) = \begin{cases} 1 \ \ if \ \ s = t \\ 0 \ \ if \ \ s \neq t \end{cases}$ is the identity element for matrix multiplication in $\mathbb{F}^{n \times n}$.

Indeed, for every matrix A, $AI(i,j) = \sum_{t=1}^{n} A(i,t)I(t,j) = A(i,j)$ because $A(i,t)I(t,j) = A(i,j)$ when $t = j$ and 0 otherwise. Hence $AI = A$. A similar argument proves that $IA = A$.

The existence of the identity element for matrix multiplication in $\mathbb{F}^{n \times n}$ raises the question of invertibility. We will address the question soon enough.

All entries of the identity matrix I are equal to 0 except for the diagonal entries which are all equal to 1. The term *diagonal entries* of a square $n{\times}n$ matrix A refers to the main diagonal of the $n{\times}n$ matrix, i.e., the line connecting the top-left with the bottom-right corner of the matrix. The line consists of all elements of the form $A(i,i)$, i=1,2, … , n.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ the } 3 \times 3 \text{ identity matrix.}$$

Theorem.

For every $n \times k$ matrix A and for every two $k \times s$ matrices B and C, $A(B+C) = AB + AC$, i.e. matrix multiplication is distributive with respect to matrix addition.

Theorem.

For every $n \times k$ matrix A, $k \times s$ matrix B and $s \times l$ matrix C, $A(BC) = (AB)C$, i.e. matrix multiplication is associative.

We postpone proving the last two theorems until we introduce linear mappings and their matrix representations. Then, the theorems follow easily from some general facts.

Definition.

Transposition is another matrix-specific operation. If A is an $m\times n$ matrix then "*A transposed*" is the $n\times m$ matrix A^{T} such that for each i and j ($1\leq i\leq n, 1\leq j\leq m$) $A^T(i,j) = A(j,i)$.

In other words, the first row of A becomes the first column of A^T and so on:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1,n} & a_{2.n} & \cdots & a_{n,m} \end{bmatrix}$$

Definition.

If $A = A^T$ then A is said to be *symmetric*.

Example.

$$[1\ 3\ 4]^T = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix},$$

$$([1\ 3\ 4]^T)^T = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}^T = [1\ 3\ 4]$$

Fact. (obvious)

For every matrix A, $(A^T)^T = A$

For every two matrices of matching sizes, $(A + B)^T = A^T + B^T$

Fact. (less obvious but easy enough)

For every two matrices A and B such that AB exists

$$(AB)^T = B^T A^T$$

Proof.

$(AB)^T(j,i) = (AB)(i,j) =$

$$\sum_{s=1}^{n} A(i,s)B(s,j) = \sum_{s=1}^{n} A^T(s,i)B^T(j,s) = \sum_{s=1}^{n} B^T(j,s)A^T(s,i) =$$

$B^T A^T(j,i)$. QED

Definition.

Let A be an $n \times k$ matrix. We say that A is a *row echelon* matrix iff

(a) if r_i is a nonzero row of A then r_{i-1} is also a nonzero row, $i = 2,3, \ldots$

(b) if $a_{i,j}$ is the first nonzero entry in r_i and $a_{i-1,p}$ is the first nonzero entry in r_{i-1} then $p < j$

If, in addition,

(c) the first nonzero entry in each nonzero row is equal to 1

(d) the first nonzero entry in each nonzero row is the only nonzero entry in its column

then A is called a *row canonical* matrix.

Example.

$$A = \begin{bmatrix} 0 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 3 & 4 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 7 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition.

The following transformations of a matrix are called *elementary row operations (EROS):*

1. $r_i \leftrightarrow r_j$ - replacing row r_i with r_j and vice versa (row swapping)
2. $r_i \leftarrow cr_i$ - replacing row r_i with r_i scaled by a nonzero constant c. In practice, we abbreviate the symbol to cr_i
3. $r_i \leftarrow r_i + r_j$ - replacing row r_i with the sum of r_i and r_j (adding of r_j to r_i). Usually, we write simply $r_i + r_j$.
4. $r_i \leftarrow r_i + cr_j$ - replacing row r_i with the sum of r_i and the multiple of r_j by a constant c. We just write $r_i + cr_j$ for short.

Notice that 4 is a composition of 2 and 3. Namely, we do cr_j, then $r_i + r_j$ (here r_i denotes the "new" row j, after scaling) and finally $c^{-1}r_j$ to convert row j to its original form.

Definition.

Matrices A and B are said to be *row-equivalent* iff A can be transformed into B by a (finite) number of elementary row operations. We denote row-equivalence by $A \sim B$.

Proposition.

The relation of row-equivalence is an equivalence relation on $\mathbb{F}^{n \times m}$.

Theorem.

Every matrix A is row-equivalent to some row-canonical matrix B. In other words, every matrix can be row-reduced to a row-canonical matrix.

Proof. (Row-reduction algorithm)

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}.$$

Step 1: Swap rows (if necessary) to get a non-zero number in the top position, (1,1). If all entries in column 1 are 0 do this for column 2 instead, if … . Multiply row 1 (new row 1) by $a_{1,1}^{-1}$ (new $a_{1,1}$ inverse). Use operation 4 to eliminate (turn into 0) every entry below the resulting 1)

This results in a new matrix, say $A_1 = \begin{bmatrix} 1 & b_{1,2} & \dots & b_{1,n} \\ 0 & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & b_{m,2} & \dots & b_{m,n} \end{bmatrix}$.

Step 2: Apply Step 1 to the matrix A_1 without row 1, i.e., to the matrix $\begin{bmatrix} 0 & b_{2,2} & \dots & b_{2,n} \\ 0 & b_{3,2} & \dots & b_{3,n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & b_{m,2} & \dots & b_{m,n} \end{bmatrix}$. This results in something like

$\begin{bmatrix} 1 & b_{1,2} & b_{1,3} & \dots & b_{1,n} \\ 0 & 1 & c_{2,3} & \dots & c_{2,n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & c_{m,3} & \dots & c_{m,n} \end{bmatrix}$. Repeat step 2 until you obtain something like this

Repeat step 2 until you obtain a matrix like this

$$\begin{bmatrix} 1 & b_{1,2} & b_{1,3} & \dots & \dots & b_{1,n} \\ 0 & 1 & c_{2,3} & \dots & \dots & c_{2,n} \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 & ? \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix}$$

This is a row echelon matrix with a leading 1 in each non-zero row.

Step 3. Apply operation 4 to eliminate all non-zeroes above the leading 1's. Thanks to the form of the nonzero rows, this will not destroy any zero to the left of the leading 1's.

Example.

$$A = \begin{bmatrix} 0 & 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & 6 & 4 & 2 \end{bmatrix} r_1 \leftrightarrow r_4 \begin{bmatrix} 0 & 2 & 6 & 4 & 2 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{bmatrix} \frac{r_1}{2} \begin{bmatrix} 0 & 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{bmatrix}$$

$$r_2 - r_1, r_3 - 2r_1 \begin{bmatrix} 0 & 1 & 3 & 4 & 1 \\ 0 & 0 & -2 & -1 & -1 \\ 0 & 0 & -5 & -4 & -2 \\ 0 & 0 & 2 & 1 & 1 \end{bmatrix} 2r_2 - r_3 \begin{bmatrix} 0 & 1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & -5 & -4 & -2 \\ 0 & 0 & 2 & 1 & 1 \end{bmatrix}$$

$$r_3 + 5r_2, r_4 - 2r_2 \begin{bmatrix} 0 & 1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 6 & -2 \\ 0 & 0 & 0 & -3 & 1 \end{bmatrix} r_4 + \frac{1}{2}r_3, \frac{1}{6}r_3 \begin{bmatrix} 0 & 1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$r_1 - 3r_2, r_2 - 2r_3 \begin{bmatrix} 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} r_1 + 2r_3 \begin{bmatrix} 0 & 1 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition.

The *row rank* of an $n \times m$ matrix A, $r(A)$, is the dimension of the subspace of $\mathbb{F}^m$ spanned by rows of A.

Theorem.

For every two matrices A and B, if $A \sim B$ then $r(A) = r(B)$.

Proof. (outline)

We prove this by showing that each elementary row operation preserves the very space spanned by rows of the matrix hence, it also preserves its dimension.

Note. Since the rank of any row echelon matrix is clearly the number of its nonzero rows, the theorem provides a method for calculating the rank of the matrix - row reduce the matrix to a row echelon one and count its nonzero rows.

In particular, the rank of the matrix from the last example is 3

Theorem.

For every matrix A, $r(A) = r(A^T)$.

We skip the proof .

Note. We could just as well define column rather than row operations and column rank of a matrix. In view of the last theorem the row and the column rank of every matrix is the same, so we just use the term *rank*.

FAQ.

1. Can we do several EROS in one step, like you did in the example?

It depends. A common mistake is to do something like $r_1 - r_2$ and $r_3 - r_1$ in one go. What is wrong with this? Row r_1 is modified by the first operation which means in the second one you should use the new r_1. On the other hand, if you first do $r_3 - r_1$ and then $r_1 - r_2$ it's ok. In extreme cases, people might row-reduce any matrix to nil, like this:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \sim_{r_1 - r_2, r_2 - r_1} \begin{bmatrix} a - c & b - d \\ c - a & d - b \end{bmatrix} \sim_{r_1 + r_2, r_2 + r_1} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

In short, when in doubt do one ERO at a time.

FAQ.

2. Can we mix EROS with ECOS?

It depends. You must avoid doing row and column operations in one transformation: writing $A \sim (r_1 - r_3,\, c_4 + c_1)\, B$ is asking for trouble because after $r_1 - r_3$ columns c_4 and c_1 are not what they were, a row operation affects all columns (a row contains one entry from each column).

3. OK, can we mix EROS with ECOS but using only EROS or only ECOS within a single transformation?

It depends. If you calculate a determinant (soon to be introduced) it's ok. If you calculate the rank of a matrix – no worries.
But if you are solving a system of equations – beware. Row operations correspond to operations on equations (side-to-side addition and the like) which preserve the solution set of a system. Column operations would mean adding coefficients of one unknown to coefficients of another – that makes no sense at all.