ETMAG

LECTURE 14

- EROS cont-d.
- Systems of Equations
- Kronecker-Capelli Theorem
- Determinant

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. (skipped)

Note. Since the rank of a row echelon matrix is the number of its nonzero rows, to calculate the rank of a matrix we row reduce the matrix to a row echelon one and count its nonzero rows.

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*.

Example.

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

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

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

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

FAQ.

1. Can we do several EROS in one step, like we 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:

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

which, if correct, would mean that the rank of every matrix is 0.

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 then, 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.

SYSTEMS OF LINEAR EQUATIONS

A system of linear equations

$$(*)\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ ... + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ ... + a_{2,n}x_n = b_2 \\ \cdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \ ... + a_{m,n}x_n = b_m \end{cases}$$

can be represented as a single matrix equation $AX = B$, where A = $[a_{i,j}]$ is called the coefficient matrix,

$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$. X and B are single-column matrices.

The system of linear equations $(*)$ can also be represented as a vector equation

$$x_1 \underbrace{\begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix}}_{C_1} + x_2 \underbrace{\begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix}}_{C_2} + \ \dots + x_n \underbrace{\begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix}}_{C_n} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{B}$$

Which means we are trying to find coefficients to express B as a linear combination of columns of A. This can only be done if

$$\text{span}\{C_1, C_2, \dots, C_n\} = \text{span}\{C_1, C_2, \dots, C_n, B\}.$$

The matrix with columns $C_1, C_2, \dots, C_n$ and B is called *the augmented matrix* of the system $(*)$ and is denoted by $[A|B]$.

Theorem. (Kronecker, Cappelli)

A system $AX = B$ of linear equations has a solution iff

$$r(A) = r([A|B]).$$

Proof. The vector-oriented approach from the previous slide together with properties of the *span* operation is proof enough.

Remark.

Interchanging equations, multiplying equations by non-zero numbers and adding equations one to another do not affect the set of solutions of a system of equations. EROS are exactly these operations except that they are performed on rows of a matrix, rather than on equations. This suggests a strategy for solving a system of equations. Start with a system $(*)$, represent it as its augmented matrix [A|B], row-reduce the matrix to a row echelon (or row canonical) matrix [E|C], return to the language of equations.

Consider the system

$$\begin{cases} 2x + 4y - z = 11 \\ -4x - 3y + 3z = -20 \\ 2x + 4y + 2z = 2 \end{cases}$$

Its augmented matrix is

$$\begin{bmatrix} 2 & 4 & -1 & 11 \\ -4 & -3 & 3 & -20 \\ 2 & 4 & 2 & 2 \end{bmatrix} \sim (r_3 - r_1, r_2 + 2r_1) \begin{bmatrix} 2 & 4 & -1 & 11 \\ 0 & 5 & 1 & 2 \\ 0 & 0 & 3 & -9 \end{bmatrix}.$$

Clearly, the rank of both A and $[A|B]$ is 3 which means the system is solvable. Let us reduce $[A|B]$ to a row-canonical matrix.

$$\sim (r_1 + \tfrac{1}{3} r_3, r_2 - \tfrac{1}{3} r_3, \tfrac{1}{3} r_3) \begin{bmatrix} 2 & 4 & 0 & 8 \\ 0 & 5 & 0 & 5 \\ 0 & 0 & 1 & -3 \end{bmatrix} \sim \left(r_1 - \tfrac{4}{5} r_2, \tfrac{1}{5} r_2\right)$$

$$\begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{bmatrix} \sim (\tfrac{1}{2} r_1) \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{bmatrix},$$ which is the matrix of

$x = 2, y = 1\ and\ z = -3.$

Definition.

A system of equations $AX = B$ is called *homogeneous* iff $B = \Theta$.

Fact.

Every homogeneous system of linear equations has a solution, namely $x_1 = 0, x_2 = 0, \dots, x_n = 0$. Any other solution (if there is one) is called a *non-trivial* or *non-zero* solution.

Theorem.

Let $AX = \Theta$ be a homogeneous system of m linear equations with n unknowns. Then the set $W = \{v \in \mathbb{K}^n | Av = \Theta\}$ of all solutions of the system is a subspace of the vector space $\mathbb{K}^n$. Moreover,

$\dim(W) = n - r(A)$.

Proof. (of the first statement)

Take $u, v \in W$. This means that $Au = \Theta$ and $Av = \Theta$. Since matrix multiplication is distributive over addition, we have $A(u + v) = Au + Av = \Theta + \Theta = \Theta$ i.e., $u + v \in W$.

Similarly, we prove that for every $k \in \mathbb{K}$ we have $A(ku) = k(Au) = k\Theta = \Theta$.

We skip the proof of the second statement. QED

Example.

$$\begin{cases} x + y - z = 0 \\ 2x - 3y + z = 0 \\ x - 4y + 2z = 0 \end{cases} \quad A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -3 & 1 \\ 1 & -4 & 2 \end{bmatrix} \sim r_2 - 2r_1, r_3 - r_1 \sim$$

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & -5 & 3 \\ 0 & -5 & 3 \end{bmatrix} \sim r_3 - r_2 \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$. The rank of the last matrix is 2. Hence the dimension of the solution space is $3 - 2 = 1$.

We shall find a basis for the space reducing the matrix further.

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & 0 \end{bmatrix} \sim \frac{1}{-5} r_2 \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & \frac{-3}{5} \\ 0 & 0 & 0 \end{bmatrix} \sim \mathrm{r}_1\text{-}\mathrm{r}_2 \sim \begin{bmatrix} 1 & 0 & \frac{-2}{5} \\ 0 & 1 & \frac{-3}{5} \\ 0 & 0 & 0 \end{bmatrix}.$$

In the language of equations this reads

$$\begin{cases} x + \quad \frac{-2}{5} z = 0 \\ \quad y - \frac{3}{5} z = 0 \\ \quad 0z = 0 \end{cases} .$$

The bottom equation really says, "z may be anything you like" and the top two say $x = \frac{2}{5} z$ and $y = \frac{3}{5} z$. Hence every vector (x, y, z) belonging to the solution space looks like $\left(\frac{2}{5} z, \frac{3}{5} z, z\right) =$

$z(\frac{2}{5}, \frac{3}{5}, 1)$ and the set $\{(\frac{2}{5}, \frac{3}{5}, 1)\}$ is a one-element basis for the space.

Theorem.

Let $AX = B$ be an arbitrary system of linear equations. Let U be the solution set and let $v_0 \in U$. Then,

$$U = v_0 + W = \{v_0 + w | w \in W\},$$

where W is the solution space of the corresponding homogeneous system $AX = \Theta$.

Proof.

Each vector $v = v_0 + w$ from $v_0 + W$ is a solution to $AX = B$. Indeed, $A(v_0 + w) = Av_0 + Aw = B + \Theta = B$. Hence, $v_0 + W \subseteq U$

Moreover, for every vector $z \in U$ we may put $t = z - v_0$ so that $z = v_0 + t$. Then

$At = A(z - v_0) = Az - Av_0 = B - B = \Theta$

which means $t \in W$. Hence, $U \subseteq v_0 + W$. QED

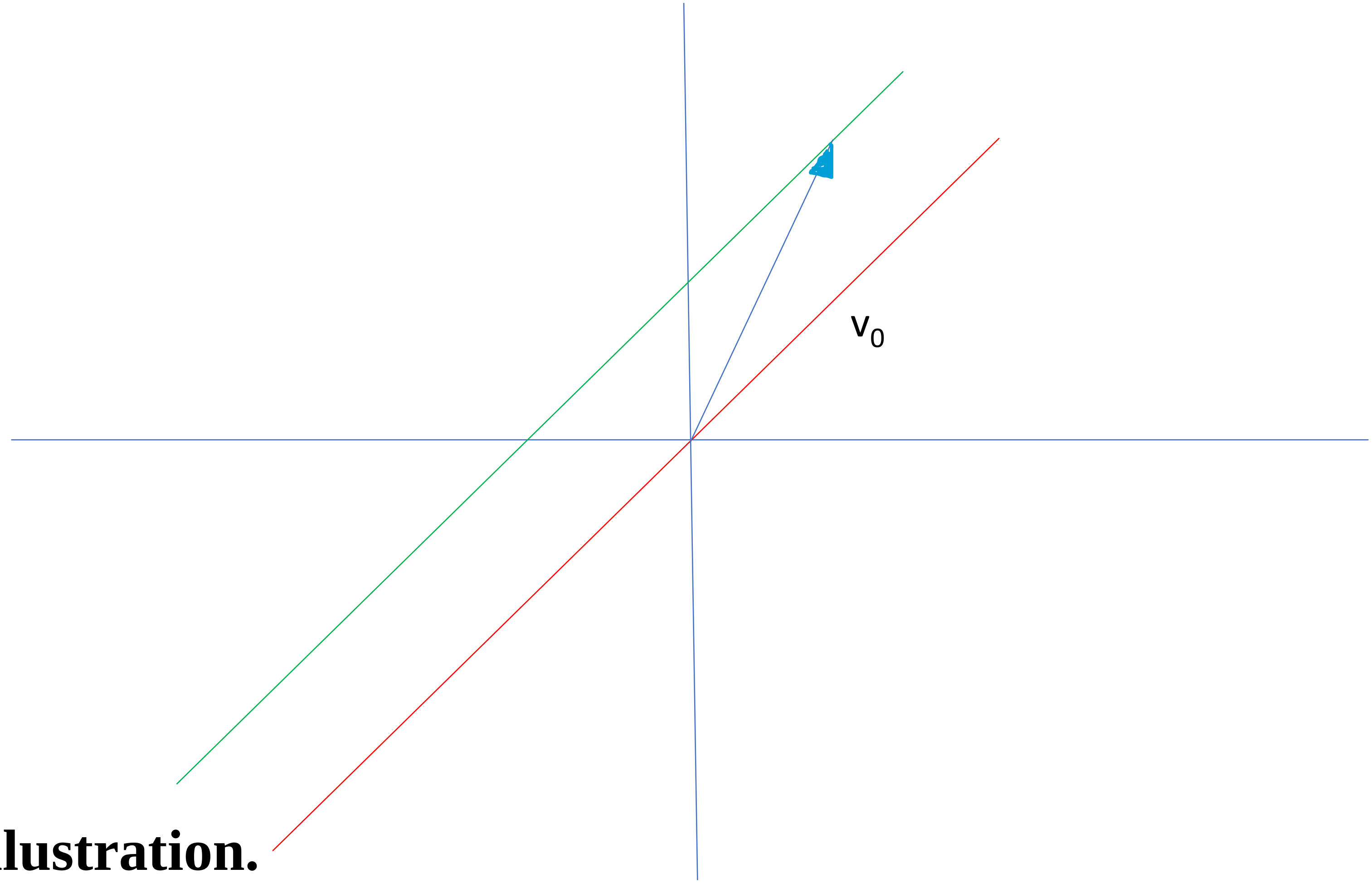

Illustration.

(1) $\{-x + y = 1$ (a system of equation, one equation two unknowns)

(2) $\{-x + y = 0$ (the corresponding homogeneous system)

v_0 – a solution of (1)

Definition.

Determinant (*det* for short) is a function defined on the set of all square matrices ($n\times n$ matrices, n=1,2, …) over a field $\mathbb{K}$ into $\mathbb{K}$. The definition is inductive with respect to n:

1. if $n = 1$, ($A = [a_{1,1}]$) then $\det(A) = a_{1,1}$
2. if $n > 1$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det(A_{i,1})$$

where $A_{i,j}$ denotes the matrix obtained from A by the removal of row number i and column j. $\det(A)$ is also denoted by $|A|$.

The formula is known as *Laplace expansion on column 1*.

Notice that we only use the symbol $A_{i,j}$ in the case $j = 1$.

Example.

1. Find $det\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$.

$$\det(A) = \sum_{i=1}^{2} (-1)^{i+1} a_{i,1} \det\left(A_{i,1}\right) = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$$

In particular, $\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = 2 \cdot (-2) - 1 \cdot 3 = -7$

Example.

2. $det\begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} = (-1)^{1+1}a\, det\begin{bmatrix} q & r \\ y & z \end{bmatrix} + (-1)^{2+1}p\, det\begin{bmatrix} b & c \\ y & z \end{bmatrix}$

$+ (-1)^{3+1}x\, det\begin{bmatrix} b & c \\ q & r \end{bmatrix} = a(qz - ry) - p(bz - cy) + x(br - qc) =$

$aqz + pyc + xbr - cqx - rya - zbp$. The last formula is known as the *Sarrus Rule.*

BEWARE !. It only works for 3×3 matrices.

$$\begin{array}{c|ccc|c} + & a & b & c & - \\ + & p & q & r & - \\ + & x & y & z & - \\ & a & b & c & \\ & p & q & r & \end{array}$$

Theorem.

For every $j = 1,2, \ldots, n$ and for every $n \times n$ matrix A

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

Proof. Omitted.

Remark. The theorem says that instead of Laplace expansion on column 1 we can do Laplace expansion on column *j*, for every *j*.

Example.

1. Find $det\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ by *Laplace expansion* on column 2.

$$\det(A) = \sum_{i=1}^{2} (-1)^{i+2} a_{i,2} \det\left(A_{i,2}\right) = -a_{1,2}a_{2,1} + a_{2,2}a_{1,1}$$

Theorem. (determinant versus transposition)

For every matrix A $\det(A) = \det(A^T)$

Proof. Omitted.

Remark. The theorem says (indirectly) that instead of Laplace expansion on columns we can do Laplace expansion on rows.

Example.

1. Find $det\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ by *Laplace expansion* on row 1.

$$\det(A) = \sum_{i=1}^{2} (-1)^{1+i} a_{1,i} \det\left(A_{1,i}\right) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

Theorem. (*determinant* versus EROS)

For every matrix A

1. If $A \sim (r_i \leftrightarrow r_j)\ B$ then $\det(B) = -\det(A)$ ($i \neq$ j)
2. If $A \sim (r_i \leftarrow cr_i)\ B$ then $\det(B) = c\det(A)$
3. If $A \sim (r_i \leftarrow r_i + r_j)\ B$ and $i \neq$ j then $\det(B) = \det(A)$
4. Combining 3 with 2 we get

   If $A \sim (r_i \leftarrow r_i + cr_j)B$ and $i \neq$ j then $\det(B) = \det(A)$.

Proof. Omitted.

Remark. Thanks to the transposition law the theorem applies to column operations as well.

Remark.
"$A \sim (r_i \leftarrow cr_i)\ B$" means "$B$ has been obtained from A by replacing r_i of A with cr_i".

Theorem. (determinant versus not-quite-matrix-addition)

Suppose $s \in \{1,2,\dots,n\}$ and $A[i,j] = B[i,j] = C[i,j]$ for every i,j such that $j \neq s$ and $C[i,s] = A[i,s] + B[i,s]$. Then $\det(C) = \det(A) + \det(B)$.

Proof..

$$\det\begin{bmatrix} c_{1,1} & \cdots & a_{1,s}+b_{1,s} & \cdots & c_{1,n} \\ c_{2,1} & \cdots & a_{2,s}+b_{2,s} & \cdots & c_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{n,1} & \cdots & a_{m,s}+b_{n,s} & \cdots & c_{n,n} \end{bmatrix} =$$

By Laplace
expansion
on column s

$$\sum_{i=1}^{n}(-1)^{i+s}(a_{i,s}+b_{i,s})\det(C_{i,s}) = \sum_{i=1}^{n}(-1)^{i+s}a_{i,s}\det(C_{i,s}) + \sum_{i=1}^{n}(-1)^{i+s}b_{i,s}\det(C_{i,s}) = \det(A) + \det(B).$$

Warning. This is NOT about determinant of the sum of two matrices being equal to the sum of their determinants; **that is not true.** This is about determinant of a matrix whose ONE column is the sum of two vectors.

Theorem. (other properties of *det*)

For every $n \times n$ matrices A and B

1. $\det(A) \neq 0$ iff $r(A) = n$, in other words, rows of A are linearly independent
2. If for every i,j such that $i > j$ $a_{i,j} = 0$ (only 0's below the main diagonal, triangular matrix) then $\det(A) = a_{1,1} a_{2,2} \ldots a_{n,n}$
3. In particular, $det(I_{n,n}) = 1$
4. $\det(AB) = \det(A) \det(B)$

Proof. Omitted.

Part 2 suggests a strategy for calculation of determinants of large matrices: row-reduce the matrix to a triangular form.