

MA 110: Lecture 07

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Review of last lecture

We have discussed the following important notions.

- Rank of a matrix
- Vector subspaces (of $\mathbb{R}^{n \times 1}$).
- Basis of a subspace
- Dimension of a subspace
- Span of a subset of a vector subspace.
- Null space and the column space of a matrix.

And we proved several important results such as:

- Characterizations of a basis of a vector subspace
- Rank-Nullity Theorem
- Fundamental Theorem for Linear Systems:

Remark: The notion of a vector subspace of $\mathbb{R}^{1 \times m}$ is defined similarly. The corresponding notions of basis, dimension, span, etc. are defined in an identical manner.

Let us restate two earlier results which are in conformity with the Rank-Nullity Theorem. Let \mathbf{A} be an $n \times n$ matrix. Then

$$\mathbf{A} \text{ is invertible} \iff \text{nullity } \mathbf{A} = 0 \iff \text{rank } \mathbf{A} = n.$$

Further, $\text{rank } \mathbf{A} = n \iff \mathcal{C}(\mathbf{A}) = \mathbb{R}^{n \times 1}$.

We are now in a position to state and prove a comprehensive result regarding solutions of a system of m linear equations in n unknowns that we started with, namely

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

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

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \quad (m)$$

As usual, we write this as $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}, \mathbf{x} := [x_1 \cdots x_n]^T \text{ and } \mathbf{b} := [b_1 \cdots b_m]^T.$$

Theorem (Fundamental Theorem for Linear Systems: FTLS)

Let $m, n \in \mathbb{N}$ and \mathbf{A} be an $m \times n$ matrix with real entries. Suppose $\text{rank } \mathbf{A} = r$.

(i) **Homogeneous Linear System** : $\mathbf{Ax} = \mathbf{0}$ (H)

The solution space $\{\mathbf{x} \in \mathbb{R}^{n \times 1} : \mathbf{Ax} = \mathbf{0}\}$ of (H) is a subspace of $\mathbb{R}^{n \times 1}$ of dimension $n - r$.

In particular, $r = n$ if and only if $\mathbf{0}$ is the only solution of (H).

If $r < n$, then there are linearly independent solutions $\mathbf{x}_1, \dots, \mathbf{x}_{n-r}$ of (H) and every solution of (H) is a unique linear combination of these $\mathbf{x}_1, \dots, \mathbf{x}_{n-r}$.

(ii) **General Linear System** : $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{b} \in \mathbb{R}^{m \times 1}$ (G)

(G) has a solution if and only if $\text{rank}[\mathbf{A}|\mathbf{b}] = r$. In this case, let \mathbf{x}_0 be a particular solution of (G). If \mathbf{x} is a solution of (G), then $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$, where \mathbf{x}_h is a solution of (H) above.

Proof. (i) The solution space of the homogeneous linear system (H) is just the nullspace $\mathcal{N}(\mathbf{A})$ of \mathbf{A} , and we have seen that its dimension, that is, the nullity of \mathbf{A} , is equal to $n - r$.

We note that $r = n \iff n - r = 0$, that is, the dimension of $\mathcal{N}(\mathbf{A})$ is zero; in other words, $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$. This means that $\mathbf{0}$ is the only solution of (H).

Let now $r < n$. Then $\mathcal{N}(\mathbf{A})$ has a basis consisting of $n - r$ elements, say $\mathbf{x}_1, \dots, \mathbf{x}_{n-r}$. Hence every element of the solution space is a unique linear combination of the elements in this basis.

(ii) Let $\mathbf{b} \in \mathbb{R}^{n \times 1}$. Let $\mathbf{c}_1, \dots, \mathbf{c}_n$ be the n columns of \mathbf{A} . Then

$$\mathbf{Ax} = x_1\mathbf{c}_1 + \dots + x_n\mathbf{c}_n \quad \text{for } \mathbf{x} := \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T \in \mathbb{R}^{n \times 1}.$$

Hence $\mathbf{Ax} = \mathbf{b}$ for some $\mathbf{x} \in \mathbb{R}^{n \times 1}$ if and only if \mathbf{b} is a linear combination of the columns of \mathbf{A} , that is, $\mathbf{b} \in \mathcal{C}(\mathbf{A})$.

Since every column of \mathbf{A} is also a column of the augmented matrix $[\mathbf{A}|\mathbf{b}]$, the column space $\mathcal{C}(\mathbf{A})$ of \mathbf{A} is contained in the column space $\mathcal{C}([\mathbf{A}|\mathbf{b}])$ of $[\mathbf{A}|\mathbf{b}]$. It follows that $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ if and only if $\mathcal{C}([\mathbf{A}|\mathbf{b}]) = \mathcal{C}(\mathbf{A})$, that is, the column rank of $[\mathbf{A}|\mathbf{b}]$ is equal to the column rank of \mathbf{A} . So $\text{rank}[\mathbf{A}|\mathbf{b}] = \text{rank } \mathbf{A} = r$.

Let \mathbf{x}_0 be a particular solution of (G), that is, let $\mathbf{x}_0 \in \mathbb{R}^{n \times 1}$ satisfy $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$. Then for any $\mathbf{x} \in \mathbb{R}^{n \times 1}$, we see that $\mathbf{A}\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}$, that is, \mathbf{x} is a solution of (G) if and only if $\mathbf{x}_h := \mathbf{x} - \mathbf{x}_0$ is a solution of (H). The proof is complete. \square

Remark

The above theorem is of immense theoretical importance. It tells us precisely when solutions exist, and also describes the nature of solutions of a linear system of equations.

For example, it says that when there is a nonzero solution of a homogeneous linear system, there are infinitely many solutions.

Further, when a homogeneous system has infinitely many solutions, it says that they can be described by a one parameter family, or a two parameter family etc.

It may seem that to implement the results of the above theorem, we must first find the rank of the coefficient matrix \mathbf{A} of the linear system. This is not necessary.

We have already seen that we may directly proceed to find the solutions of the linear system by considering the augmented matrix $[\mathbf{A}|\mathbf{b}]$ and transform the coefficient matrix \mathbf{A} to a row echelon form by the Gauss Elimination Method and then use Back Substitution. This process itself reveals all possibilities.

In particular, when the rank r of \mathbf{A} is less than the number n of variables, we have shown how to construct a set S of basic solutions of an homogeneous linear system. This set S is in fact a basis of the solution space of the system. That is the reason for using the terminology 'basic solutions'.

Row Space and Column Space

Definition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The **row space** of \mathbf{A} , denoted $\mathcal{R}(\mathbf{A})$, is defined as the subspace of $\mathbb{R}^{1 \times n}$ spanned by the row vectors of \mathbf{A} .

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Here are some important observations.

- The row-rank of \mathbf{A} is precisely the dimension of $\mathcal{R}(\mathbf{A})$.
- If $\mathbf{A}' \in \mathbb{R}^{m \times n}$ is obtained from \mathbf{A} by an elementary row operation, then $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{A}')$.
- If $\mathbf{A}' \in \mathbb{R}^{m \times n}$ is in REF, then the nonzero rows of \mathbf{A}' form a basis of $\mathcal{R}(\mathbf{A}')$.
- A **basis of $\mathcal{R}(\mathbf{A})$** is given by the **nonzero rows of its REF**.
- If \mathbf{A}' is obtained from \mathbf{A} by an elementary row operation, then $\mathcal{C}(\mathbf{A})$ need not be equal to $\mathcal{C}(\mathbf{A}')$.

Lemma

The columns of \mathbf{A} corresponding to the pivotal columns of its REF form a basis of $\mathcal{C}(\mathbf{A})$.

Proof. Let \mathbf{A}' be an REF of \mathbf{A} . Suppose the columns $\mathbf{c}'_{i_1}, \dots, \mathbf{c}'_{i_r}$ of \mathbf{A}' are the pivotal columns. Let $\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_r}$ be the corresponding columns of \mathbf{A} . It is enough to prove that $\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_r}$ are linearly independent. If they were not so, then there would be constants $\alpha_{i_1}, \dots, \alpha_{i_r}$, not all zero such that

$$\alpha_{i_1} \mathbf{c}_{i_1} + \dots + \alpha_{i_r} \mathbf{c}_{i_r} = \mathbf{0}.$$

Consider the vector $\mathbf{x} = (x_1, \dots, x_n)^T$ such that $x_j = \alpha_j$ if j is one of i_1, \dots, i_r , and $x_j = 0$ other wise. Then $\mathbf{Ax} = \mathbf{0}$, thus $\mathbf{A}'\mathbf{x} = \mathbf{0}$ as well. But this means,

$$\alpha_{i_1} \mathbf{c}'_{i_1} + \dots + \alpha_{i_r} \mathbf{c}'_{i_r} = \mathbf{0},$$

which is a contradiction because $\mathbf{c}'_{i_1}, \dots, \mathbf{c}'_{i_r}$ being the pivotal columns of a matrix in REF, are linearly independent. □

Example: Consider the 5×6 matrix \mathbf{A} and its REF \mathbf{A}' given by

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then $\text{rank } \mathbf{A} = 3$. A basis of the row space $\mathcal{R}(\mathbf{A})$ is given by

$$\left\{ [1 \ 3 \ -2 \ 0 \ 2 \ 0], [0 \ 0 \ -1 \ -2 \ 0 \ -3], [0 \ 0 \ 0 \ 0 \ 0 \ 6] \right\}$$

whereas a basis for the column space $\mathcal{C}(\mathbf{A})$ is given by

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -5 \\ 0 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 0 \\ 15 \\ 18 \end{bmatrix} \right\}.$$

Determinants

You already know formulas for determinants of 1×1 , 2×2 and 3×3 matrices. Let us recall them.

$$\det [a_1] = a_1, \quad \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = a_1 b_2 - a_2 b_1 \quad \text{and}$$

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1(b_2 c_3 - b_3 c_2) - b_1(a_2 c_3 - a_3 c_2) + c_1(a_2 b_3 - a_3 b_2),$$

which is also equal to

$$a_1(b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1).$$

We shall presently give formulas for the determinant of an $n \times n$ matrix, that is, of a matrix of size n , where $n \in \mathbb{N}$, and we shall explore their use in matrix theory.

Let $n \in \mathbb{N}$ and $\mathbf{A} := \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & a_{jk} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n}.$

The determinant of \mathbf{A} is a real number defined inductively as follows. For $n := 1$, define $\det \mathbf{A} := a_{11}$. Let $n \geq 2$, and suppose we have defined the determinant of any $(n-1) \times (n-1)$ matrix. For $j, k = 1, \dots, n$, let \mathbf{A}_{jk} denote the submatrix of \mathbf{A} obtained by deleting the j th row and the k th column of \mathbf{A} , and let $M_{jk} := \det \mathbf{A}_{jk}$, called the (j, k) th **minor** of \mathbf{A} . Define

$$\det \mathbf{A} := a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{1+k}a_{1k}M_{1k} + \cdots + (-1)^{1+n}a_{1n}M_{1n}.$$

This is also known as the **expansion for the determinant of \mathbf{A} in terms of the first row of \mathbf{A} .**

An immediate consequence of our definition is the following.

Proposition

If an $n \times n$ matrix \mathbf{A} is lower triangular, then the determinant of \mathbf{A} is the product of its diagonal entries.

Proof. We prove it by induction on n . For $n = 1$, it is the definition. Suppose we know the proposition for all $(n - 1) \times (n - 1)$ matrices. Now $\det \mathbf{A} = a_{11} M_{11}$ since $a_{12} = \cdots = a_{1n} = 0$, etc. Now $\mathbf{A}_{1,1}$ is lower triangular and thus by the induction hypothesis, $M_{1,1}$ is the product of the diagonal entries of \mathbf{A} other than a_{11} . □

Next, it can be proved by induction on the size n of a matrix that $\det \mathbf{A}$ is equal to the following expansions in terms of the j th row of \mathbf{A} , and also in terms of the k th column of \mathbf{A} :

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{j+k} a_{jk} M_{jk} \quad \text{for each } j \in \{1, \dots, n\}$$

$$\det \mathbf{A} = \sum_{j=1}^n (-1)^{j+k} a_{jk} M_{jk} \quad \text{for each } k \in \{1, \dots, n\}$$

(For a proof, see Kreyszig, Appendix 4, page A81.)

Proposition

Let \mathbf{A} be a square matrix. Then $\det \mathbf{A}^T = \det \mathbf{A}$.

Proof. This is obvious if $n = 1$. Let now $n \geq 2$, and assume this property for all $(n-1) \times (n-1)$ matrices. Note that $(\mathbf{A}^T)_{jk} = (\mathbf{A}_{kj})^T$ for all $j, k = 1, \dots, n$, that is, the submatrix obtained by deleting the j th row and the k th column of \mathbf{A}^T is the same as the transpose of the submatrix obtained by deleting the k th row and the j th column of \mathbf{A} . (For example,

$$\mathbf{A} := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \implies \mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix},$$

$$\text{so that } (\mathbf{A}^T)_{21} = \begin{bmatrix} a_{21} & a_{31} \\ a_{23} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}^T = (\mathbf{A}_{12})^T.)$$

Let $\mathbf{A} := [a_{jk}]$ and $\mathbf{A}^T := [a'_{jk}]$. Then $a'_{jk} = a_{kj}$ and $M'_{jk} := \det(\mathbf{A}^T)_{jk} = \det(\mathbf{A}_{kj})^T = \det \mathbf{A}_{kj} = M_{kj}$ by the inductive hypothesis for $j, k = 1, \dots, n$. Expanding $\det \mathbf{A}^T$ in terms of its first row, and $\det \mathbf{A}$ in terms of its first column,

$$\begin{aligned} \det \mathbf{A}^T &= a'_{11}M'_{11} - a'_{12}M'_{12} + \cdots + (-1)^{1+n}a'_{1n}M'_{1n} \\ &= a_{11}M_{11} - a_{21}M_{21} + \cdots + (-1)^{n+1}a_{n1}M_{n1} \\ &= \det \mathbf{A}. \quad \square \end{aligned}$$

Corollary

If \mathbf{A} is upper triangular, then the determinant of \mathbf{A} is the product of its diagonal entries.

Proof. Let \mathbf{A} be upper triangular. Then \mathbf{A}^T is lower triangular and has the same diagonal entries as those of \mathbf{A} . \square

Let us write $\mathbf{A} := [\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_k \quad \cdots \quad \mathbf{c}_n]$ in terms of its n columns.

Crucial Properties of the **determinant function** $\mathbf{A} \mapsto \det \mathbf{A}$ from $\mathbb{R}^{n \times n}$ to \mathbb{R} :

1. Let $k \in \{1, \dots, n\}$ and $\mathbf{c}_k = \alpha \mathbf{c}'_k + \beta \mathbf{c}''_k$, where $\alpha, \beta \in \mathbb{R}$.

Then $\det \mathbf{A} = \det [\mathbf{c}_1 \ \cdots \ \alpha \mathbf{c}'_k + \beta \mathbf{c}''_k \ \cdots \ \mathbf{c}_n]$ is equal to $\alpha \det [\mathbf{c}_1 \ \cdots \ \mathbf{c}'_k \ \cdots \ \mathbf{c}_n] + \beta \det [\mathbf{c}_1 \ \cdots \ \mathbf{c}''_k \ \cdots \ \mathbf{c}_n]$.

This is proved by expanding $\det \mathbf{A}$ in terms of its k th column. In particular,

$$\det(\alpha \mathbf{A}) = \det [\alpha \mathbf{c}_1 \ \cdots \ \alpha \mathbf{c}_k \ \cdots \ \alpha \mathbf{c}_n] = \alpha^n \det \mathbf{A}.$$

2. Let $n \geq 2$. If $\mathbf{c}_k = \mathbf{c}_\ell$ for some $k \neq \ell$, then $\det \mathbf{A} = 0$, that is, if 2 columns of a matrix are identical, then its determinant is equal to 0. This is clear for $n = 2$, and for $n \geq 3$, this is proved by expanding $\det \mathbf{A}$ in terms of a column \mathbf{c}_p of \mathbf{A} , where $p \neq k$ and $p \neq \ell$, and then using induction on n .

3. $\det \mathbf{I} = 1$, that is, the determinant of the identity matrix is equal to 1. This is obvious.

Proposition

Let \mathbf{A} be a square matrix.

- (i) If two columns of \mathbf{A} are interchanged, then $\det \mathbf{A}$ gets multiplied by -1 .
- (ii) Addition of a multiple of a column to another column of \mathbf{A} does not alter $\det \mathbf{A}$.
- (iii) Multiplication of a column of \mathbf{A} by a scalar α results in the multiplication of $\det \mathbf{A}$ by α .

Proof: Let $\mathbf{A} := [\mathbf{c}_1 \ \cdots \ \mathbf{c}_k \ \cdots \ \mathbf{c}_\ell \ \cdots \ \mathbf{c}_n]$, where $k \neq \ell$.

(i) Define $\alpha := \det [\mathbf{c}_1 \ \cdots \ (\mathbf{c}_k + \mathbf{c}_\ell) \ \cdots \ (\mathbf{c}_k + \mathbf{c}_\ell) \ \cdots \ \mathbf{c}_n]$.
Then $\alpha = 0$ since the matrix has two identical columns.

On the other hand, $\alpha = \beta + \gamma$, where

$$\begin{aligned}\beta &:= \det [\mathbf{c}_1 \ \cdots \ (\mathbf{c}_k + \mathbf{c}_\ell) \ \cdots \ \mathbf{c}_k \ \cdots \ \mathbf{c}_n] \text{ and} \\ \gamma &:= \det [\mathbf{c}_1 \ \cdots \ (\mathbf{c}_k + \mathbf{c}_\ell) \ \cdots \ \mathbf{c}_\ell \ \cdots \ \mathbf{c}_n].\end{aligned}$$

In turn, $\beta := \beta_1 + \beta_2$ and $\gamma = \gamma_1 + \gamma_2$, where

$$\begin{aligned}\beta_1 &= \det [\mathbf{c}_1 \cdots \mathbf{c}_k \cdots \mathbf{c}_k \cdots \mathbf{c}_n], \\ \beta_2 &= \det [\mathbf{c}_1 \cdots \mathbf{c}_\ell \cdots \mathbf{c}_k \cdots \mathbf{c}_n], \\ \gamma_1 &= \det [\mathbf{c}_1 \cdots \mathbf{c}_k \cdots \mathbf{c}_\ell \cdots \mathbf{c}_n], \\ \gamma_2 &= \det [\mathbf{c}_1 \cdots \mathbf{c}_\ell \cdots \mathbf{c}_\ell \cdots \mathbf{c}_n].\end{aligned}$$

But $\beta_1 = 0 = \gamma_2$ since two columns are identical. Since $0 = \alpha = \beta + \gamma = \beta_2 + \gamma_1$, we see that $\gamma_1 = -\beta_2$, that is, $\det [\mathbf{c}_1 \cdots \mathbf{c}_k \cdots \mathbf{c}_\ell \cdots \mathbf{c}_n]$ is equal to $-\det [\mathbf{c}_1 \cdots \mathbf{c}_\ell \cdots \mathbf{c}_k \cdots \mathbf{c}_n]$, as desired.

(ii) Suppose α times the ℓ th column of \mathbf{A} is added to the k th column of \mathbf{A} . Then $\det [\mathbf{c}_1 \cdots (\mathbf{c}_k + \alpha \mathbf{c}_\ell) \cdots \mathbf{c}_\ell \cdots \mathbf{c}_n]$ is equal to $\det \mathbf{A} + \alpha \det [\mathbf{c}_1 \cdots \mathbf{c}_\ell \cdots \mathbf{c}_\ell \cdots \mathbf{c}_n] = \det \mathbf{A}$.

(iii) Suppose the k th column of \mathbf{A} is multiplied by α . Then $\det [\mathbf{c}_1 \cdots \alpha \mathbf{c}_k \cdots \mathbf{c}_n] = \alpha \det [\mathbf{c}_1 \cdots \mathbf{c}_k \cdots \mathbf{c}_n] = \alpha \det \mathbf{A}$.