MA 110:
Lecture 06

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Let S be a set of vectors (in $\mathbb{R}^{1\times n}$ or in $\mathbb{R}^{n\times 1}$). Recall:

- S is linearly dependent if a nontrivial linear combination of finitely many vectors in S is zero, i.e., there are (distinct) vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ in S and scalars $\alpha_1, \ldots, \alpha_m$, not all zero, such that $\alpha_1\mathbf{a}_1 + \cdots + \alpha_m\mathbf{a}_m = \mathbf{0}$.
- S is linearly independent if it is not linearly dependent.
- Crucial Result: If S has s vectors, and each of them is a linear combination of a (fixed) set of r vectors, with $s > r$, then S is linearly dependent.
- In particular, if S has more than n vectors, then S is linearly dependent.
- Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. The row rank of $\mathbf{A}$ is the maximum number of linearly independent row vectors of $\mathbf{A}$.
- Elementary row operations do not alter the row rank.
- row-rank($\mathbf{A}$) $=$ number of nonzero rows in a REF of $\mathbf{A}$ $=$ the number of pivots in a REF of $\mathbf{A}$.

Example

In a previous lecture, we had seen that the matrix

$$\mathbf{A} := \begin{bmatrix} 3 & 2 & 2 & -5 \\ 0.6 & 1.5 & 1.5 & -5.4 \\ 1.2 & -0.3 & -0.3 & 2.4 \end{bmatrix}$$

can be transformed to a row echelon form

$$\mathbf{A}' := \begin{bmatrix} \boxed{3} & 2 & 2 & -5 \\ 0 & \boxed{1.1} & 1.1 & -4.4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

by elementary row transformations of type I and type II. Since the number of nonzero rows of $\mathbf{A}'$ is 2, we see that the row rank of $\mathbf{A}$ is 2. This shows that the 3 row vectors of $\mathbf{A}$ are linearly dependent.

Column Rank of a Matrix

Definition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The column rank of $\mathbf{A}$ is the maximum number of linearly independent column vectors of $\mathbf{A}$.

Clearly, column-rank$(\mathbf{A})$ = row-rank$(\mathbf{A}^{\mathsf{T}})$.

Proposition

The column rank of a matrix is equal to its row rank.

Proof. Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$. Let r = row-rank$(\mathbf{A})$ and s :=column-rank$(\mathbf{A})$. We will show that $s \leq r$ and $r \leq s$. Let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ denote the row vectors of $\mathbf{A}$, so that

$$\mathbf{a}_j := \begin{bmatrix} a_{j1} & \cdots & a_{jk} & \cdots & a_{jn} \end{bmatrix} \qquad \text{for } j = 1, \ldots, m.$$

Among these m row vectors of $\mathbf{A}$, there are r row vectors $\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_r}$ which are linearly independent, and each row vector of $\mathbf{A}$ is a linear combination of these r row vectors.

For simplicity, let $\mathbf{b}_1 := \mathbf{a}_{j_1}, \ldots, \mathbf{b}_r := \mathbf{a}_{j_r}$. Then for $j = 1, \ldots, m$ and $\ell = 1, \ldots, r$, there is $\alpha_{j\ell} \in \mathbb{R}$ such that

$$\mathbf{a}_j = \alpha_{j1}\mathbf{b}_1 + \cdots + \alpha_{j\ell}\mathbf{b}_\ell + \cdots + \alpha_{jr}\mathbf{b}_r \quad \text{for } j = 1, \ldots, m.$$

If $\mathbf{b}_\ell := \begin{bmatrix} b_{\ell 1} & \cdots & b_{\ell k} & \cdots & b_{\ell n} \end{bmatrix}$ for $\ell = 1, \ldots, r$, then the above equations can be written componentwise as follows:

$$\begin{array}{ccccccccccc} a_{1k} & = & \alpha_{11}b_{1k} & + & \cdots & + & \alpha_{1\ell}b_{\ell k} & + & \cdots & + & \alpha_{1r}b_{rk} \\ \vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{jk} & = & \alpha_{j1}b_{1k} & + & \cdots & + & \alpha_{j\ell}b_{\ell k} & + & \cdots & + & \alpha_{jr}b_{rk} \\ \vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{mk} & = & \alpha_{m1}b_{1k} & + & \cdots & + & \alpha_{m\ell}b_{\ell k} & + & \cdots & + & \alpha_{mr}b_{rk} \end{array}$$

for $k = 1, \ldots, n$. These m equations show that the kth column of $\mathbf{A}$ can be written as follows:

$$\begin{bmatrix} a_{1k} \\ \vdots \\ a_{jk} \\ \vdots \\ a_{mk} \end{bmatrix} = b_{1k} \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{j1} \\ \vdots \\ \alpha_{m1} \end{bmatrix} + b_{2k} \begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{j2} \\ \vdots \\ \alpha_{m2} \end{bmatrix} + \cdots + b_{rk} \begin{bmatrix} \alpha_{1r} \\ \vdots \\ \alpha_{jr} \\ \vdots \\ \alpha_{mr} \end{bmatrix}.$$

for $k = 1, \ldots, n$. Thus each of the n columns of $\mathbf{A}$ is a linear combination of the r column vectors $\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{j1} & \cdots & \alpha_{m1} \end{bmatrix}^\mathsf{T}$, $\begin{bmatrix} \alpha_{12} & \cdots & \alpha_{j2} & \cdots & \alpha_{m2} \end{bmatrix}^\mathsf{T}, \ldots, \begin{bmatrix} \alpha_{1r} & \cdots & \alpha_{jr} & \cdots & \alpha_{mr} \end{bmatrix}^\mathsf{T}$.

By a crucial result we have proved earlier, no more than r columns of $\mathbf{A}$ can be linearly independent. Hence $s \leq r$, i.e., column-rank($\mathbf{A}$) $\leq$ row-rank($\mathbf{A}$).

Applying the above result to $\mathbf{A}^\mathsf{T}$ in place of $\mathbf{A}$, we obtain $r \leq s$. Thus $s = r$, as desired. $\square$

In view of the above result, we define the **rank** of a matrix $\mathbf{A}$ to be the common value of the row rank of $\mathbf{A}$ and the column rank of $\mathbf{A}$, and we denote it by $\boxed{\operatorname{rank}\mathbf{A}}$. Consequently, $\boxed{\operatorname{rank}\mathbf{A}^{\mathsf{T}} = \operatorname{rank}\mathbf{A}}$.

Example

Let $$\mathbf{A}' := \begin{bmatrix} \boxed{1} & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & \boxed{7} & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & \boxed{10} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since $\mathbf{A}'$ is in REF, its row rank is equal to the nonzero rows, namely 3. Also, we can easily check that the columns $\mathbf{c}_1, \mathbf{c}_4$ and $\mathbf{c}_6$ are linearly independent, and so the column rank of $\mathbf{A}'$ is at least 3. Further, each of the columns $\mathbf{c}_2 = 2\mathbf{c}_1$, $\mathbf{c}_3 = 3\mathbf{c}_1$ and $\mathbf{c}_5 = (3/7)\mathbf{c}_1 + (8/7)\mathbf{c}_4$ is linear combination of the columns $\mathbf{c}_1, \mathbf{c}_4$ and $\mathbf{c}_6$, and so by a crucial result, the column rank of $\mathbf{A}'$ is at most 3. Hence the column rank of $\mathbf{A}'$ is also 3.

Let us restate some results proved earlier in terms of the newly introduced notion of the rank of a matrix.

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$.

1. If $\mathbf{A}$ is transformed to $\mathbf{A}'$ by EROs, then $\operatorname{rank} \mathbf{A}' = \operatorname{rank} \mathbf{A}$.
2. If $\mathbf{A}$ is transformed to $\mathbf{A}'$ by EROs and $\mathbf{A}'$ is in REF, then

$$\begin{aligned} \operatorname{rank} \mathbf{A} &= \text{number of nonzero rows of } \mathbf{A}' \\ &= \text{number of pivotal columns of } \mathbf{A}'. \end{aligned}$$

(Note: Each nonzero row has exactly one pivot, and a zero row has no pivot.)

3. Let $m = n$. Then

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

(Recall: A square matrix is invertible if and only if it can be transformed to the identity matrix by EROs.)

Subspace, Basis, Dimension

From now on we shall consider only the column vectors.

Let $n \in \mathbb{N}$. We have denoted the set of all column vectors whose entries are real numbers and whose length is n by $\mathbb{R}^{n\times 1}$.

Definition

A nonempty subset V of $\mathbb{R}^{n\times 1}$ is called a vector subspace, or simply a subspace of $\mathbb{R}^{n\times 1}$ if

(i) $\mathbf{a}, \mathbf{b} \in V \implies \mathbf{a} + \mathbf{b} \in V$, and

(ii) $\alpha \in \mathbb{R}$, $\mathbf{a} \in V \implies \alpha\,\mathbf{a} \in V$.

Note: The zero column vector $\mathbf{0}$ is in every subspace V since $V \neq \emptyset$ and $\mathbf{0} = \mathbf{a} + (-1)\mathbf{a}$ for each $\mathbf{a} \in V$.

Each linear combination of elements of a subspace V is in V.

$\{\mathbf{0}\}$ is the smallest and $\mathbb{R}^{n\times 1}$ is the largest subspace of $\mathbb{R}^{n\times 1}$.

Two Important Examples of Subspaces Related to a Matrix

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$.

Definition

The **null space** of $\mathbf{A}$ is

$$\mathcal{N}(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^{n \times 1} : \mathbf{A}\mathbf{x} = \mathbf{0}\}.$$

Definition

The **column space** of $\mathbf{A}$ is

$$\mathcal{C}(\mathbf{A}) := \text{ the set of all linear combinations of columns of } \mathbf{A}.$$

We shall later find an interesting relationship between the null space $\mathcal{N}(\mathbf{A})$ and the column space $\mathcal{C}(\mathbf{A})$.

Basis and dimension of a subspace

Let V be a subspace of $\mathbb{R}^{n\times 1}$. Since every element of $\mathbb{R}^{n\times 1}$ is a linear combination of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$, we see that no linearly independent subset of V contains more than n vectors.

Definition

A subset S of V is called a basis of V if S is linearly independent and S has maximum possible number of elements among linearly independent subsets of V.

Clearly, a basis of V has at most n elements.

Definition

Let $S \subset \mathbb{R}^{n\times 1}$. The set of all linear combinations of elements of S is denoted by $\operatorname{span} S$ and called the span of S.

Proposition

Let V be a subspace of $\mathbb{R}^{n\times 1}$, and let $S \subset V$. Then S is a basis for $V \iff S$ is linearly independent and span $S = V$.

Proof. Suppose S is a basis for V. By the definition of a basis, S is linearly independent. Let S have r elements. If $\mathbf{x} \in V$, but $\mathbf{x}$ is not a linear combination of elements of S, then $S \cup \{\mathbf{x}\}$ would be a linearly independent subset of V containing $r+1$ elements, which would be a contradiction.

Conversely, suppose S is linearly independent and span $S = V$. Let S have r elements. Since every element of V is a linear combination of elements of S, no more than r elements of V can be linearly independent. Hence S is a basis for V. $\square$

Corollary

Let V be a subspace of $\mathbb{R}^{n\times 1}$ and let S, S' be two bases of V. Then S and S' have the same cardinality.

Proof. It is enough to prove that $|S| \leq |S'|$, because we can reverse the roles of S and S' and prove the other inequality. By the last proposition, every element of S is a linear combination of vectors from S'. If $|S| > |S'|$, then S would be linearly dependent. As S is a basis, it is linearly independent, which implies $|S| \leq |S'|$. □

Definition

The dimension of V is defined as the number of elements in a basis of V. It is denoted by $\dim V$.

Example: Clearly, $\dim \mathbb{R}^{n\times 1} = n$ and $\dim\{\mathbf{0}\} = 0$.

Corollary

Let V be a subspace of $\mathbb{R}^{n\times 1}$. Every linearly independent subset of V can be enlarged to a basis for V.

Proof. Let S be a linearly independent subset of V. If $\operatorname{span} S = V$, then by the previous result, S is a basis for V. Suppose $\operatorname{span} S \neq V$. Then there is $\mathbf{x}_1 \in V$ such that $\mathbf{x}_1$ is not a linear combination of elements of S. Now $S_1 := S \cup \{\mathbf{x}_1\}$ is a linearly independent subset of V. If $\operatorname{span} S_1 = V$, then as before, S_1 is a basis for V. If $\operatorname{span} S_1 \neq V$, then there is $\mathbf{x}_2 \in V$ such that $\mathbf{x}_2$ is not a linear combination of elements of S_1. Now $S_2 := S_1 \cup \{\mathbf{x}_2\}$ is a linearly independent subset of V.

This process must end after a finite number of steps since the number of elements of any linearly independent subset of V is less than or equal to $\dim V$, and $\dim V \leq n$. $\square$

Remark: All things we have defined above for column vectors can also be defined for row vectors.

Another immediate consequence of the earlier result is the unique representation of every element of a subspace in terms of the elements belonging to a basis of that subspace.

Proposition

Let $S := \{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$ be a basis for a subspace V of $\mathbb{R}^{n\times 1}$, and let $\mathbf{x} \in V$. Then there are unique $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{c}_1 + \cdots + \alpha_r\mathbf{c}_r$.

Proof. Since V is a basis for V, we obtain $V = \operatorname{span} S$, and so the vector $\mathbf{x}$ is a linear combination of vectors in S, that is, there are scalars $\alpha_1, \ldots, \alpha_r$ such that $\mathbf{x} = \alpha_1\mathbf{c}_1 + \cdots + \alpha_r\mathbf{c}_r$. Now suppose $\mathbf{x} = \beta_1\mathbf{c}_1 + \cdots + \beta_r\mathbf{c}_r$ for some $\beta_1, \ldots, \beta_r \in \mathbb{R}$. Then

$$(\alpha_1 - \beta_1)\mathbf{c}_1 + \cdots + (\alpha_r - \beta_r)\mathbf{c}_r = \mathbf{0}.$$

Since the set S is linearly independent, $\alpha_1 - \beta_1 = \cdots = \alpha_r - \beta_r = 0$, that is, $\beta_1 = \alpha_1, \ldots, \beta_r = \alpha_r$. This proves the uniqueness. $\square$

Let us recall some definitions and results from the last lecture.

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The row rank of $\mathbf{A}$ is the maximum number of linearly independent row vectors of $\mathbf{A}$. The column rank of $\mathbf{A}$ is the maximum number of linearly independent column vectors of $\mathbf{A}$. The two are equal, and we denote both by $\boxed{\operatorname{rank} \mathbf{A}}$. Further, if $\mathbf{A}'$ is any REF of $\mathbf{A}$, then the number of nonzero rows of $\mathbf{A}'$ is equal to $\operatorname{rank} \mathbf{A}'$ and it is equal to $\operatorname{rank} \mathbf{A}$.

The column space of $\mathbf{A}$ is denoted by $\mathcal{C}(\mathbf{A})$ and the null space of $\mathbf{A}$ by $\mathcal{N}(\mathbf{A})$.

Proposition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, and let $\operatorname{rank} \mathbf{A} = r$. Then $\dim \mathcal{C}(\mathbf{A}) = r$ and $\dim \mathcal{N}(\mathbf{A}) = n - r$.

Proof. Since the column rank of $\mathbf{A}$ is equal to r, there are r linearly independent columns $\mathbf{c}_{k_1}, \dots, \mathbf{c}_{k_r}$ of $\mathbf{A}$, and any other column of $\mathbf{A}$ is a linear combination of these r columns.

Let $\mathbf{x} \in \mathcal{C}(\mathbf{A})$. Then $\mathbf{x}$ is a linear combination of columns of $\mathbf{A}$, each of which in turn is a linear combination of $\mathbf{c}_{k_1}, \ldots, \mathbf{c}_{k_r}$. Thus $\mathbf{x}$ is a linear combination of $\mathbf{c}_{k_1}, \ldots, \mathbf{c}_{k_r}$. This shows that $\operatorname{span}\{\mathbf{c}_{k_1}, \ldots, \mathbf{c}_{k_r}\} = \mathcal{C}(A)$. Hence $\{\mathbf{c}_{k_1}, \ldots, \mathbf{c}_{k_r}\}$ is a basis for $\mathcal{C}(\mathbf{A})$ and $\dim \mathcal{C}(\mathbf{A}) = r$.

To find the dimension of $\mathcal{N}(\mathbf{A})$, let us transform $\mathbf{A}$ to a REF $\mathbf{A}'$ by EROs of type I and type II. Since the row rank of $\mathbf{A}$ is equal to r, the matrix $\mathbf{A}'$ has exactly r nonzero rows and exactly r pivotal columns.

Let the $n - r$ nonpivotal columns be denoted by $\mathbf{c}_{\ell_1}, \ldots, \mathbf{c}_{\ell_{n-r}}$. Then $x_{\ell_1}, \ldots, x_{\ell_{n-r}}$ are the free variables. For each $\ell \in \{\ell_1, \ldots, \ell_{n-r}\}$, there is a basic solution $\mathbf{s}_\ell$ of the homogeneous equation $\mathbf{Ax} = \mathbf{0}$, and every solution of this homogeneous equation is a linear combination of these $n - r$ basic solutions. Let S denote the set of these $n - r$ basic solutions. Then $\operatorname{span} S = \mathcal{N}(\mathbf{A})$.

We claim that the set S of the $n-r$ basic solutions is linearly independent. To see this, we note that each basic solution is equal to 1 in one of the free variables and it is equal to 0 in the other free variables. Let $\alpha_1, \ldots, \alpha_{n-r} \in \mathbb{R}$ be such that

$$\mathbf{x} := \alpha_1 \mathbf{s}_{\ell_1} + \cdots + \alpha_{n-r} \mathbf{s}_{\ell_{n-r}} = \mathbf{0}.$$

For $j = 1, \ldots, n$, let x_j denote the jth entry of $\mathbf{x}$. Then for each $\ell \in \{\ell_1, \ldots, \ell_{n-r}\}$, we see that $\alpha_\ell \cdot 1 = x_\ell = 0$. Hence S is linearly independent. Thus S is a basis for $\mathcal{N}(\mathbf{A})$ and $\dim \mathcal{N}(\mathbf{A}) = n - r$, the number of elements in S. □

The dimension of the null space $\mathcal{N}(\mathbf{A})$ of $\mathbf{A}$ is called the **nullity** of $\mathbf{A}$. Since $r + (n-r) = n$, we have proved the following **Rank-Nullity Theorem**.

Theorem

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then $\operatorname{rank} \mathbf{A} + \operatorname{nullity} \mathbf{A} = n$.