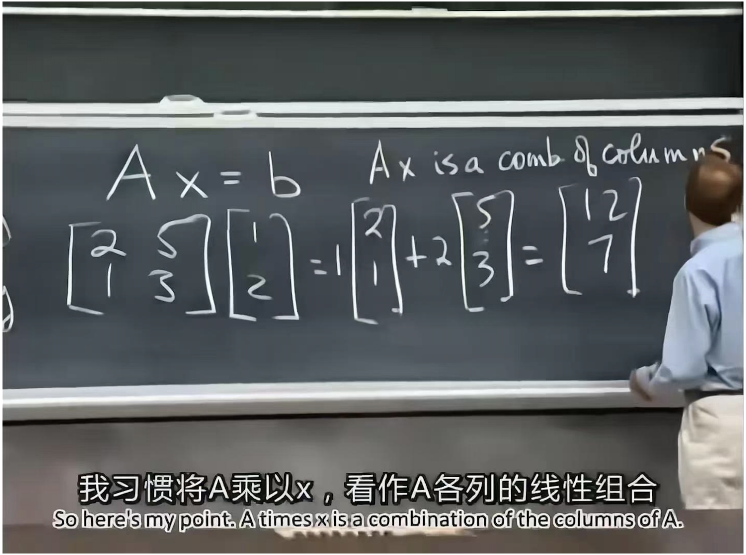
Linear Independence and Basis

Wang Xiao

Shanghai University of Finance and Economics

March 23, 2026


A x = b
A x is a comb of columns
我习惯将A乘以x，看作A各列的线性组合
So here's my point. A times x is a combination of the columns of A.

Given vectors $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$, it is natural to ask whether $\mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$. That is, do there exist scalars $c_1, .., c_k$ so that $\mathbf{v} = c_1\mathbf{v}_1 + ... + c_k\mathbf{v}_k$? This is in turn a question of whether a certain inhomogeneous system of linear equations has a solution. We are often interested in the question: is that solution unique?

Proposition

Let $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$ *and let* $V = \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$. *An arbitrary vector* $\mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$ *has a unique expression as a linear combination of* $\mathbf{v}_1, ..., \mathbf{v}_k$ *if and only if the zero vector has a unique expression as a linear combination of* $\mathbf{v}_1, ..., \mathbf{v}_k$; *i.e.,*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k = \mathbf{0} \Rightarrow c_1 = ... = c_k = 0.$$

Proof: 1. linear comb. of $\mathbf{0}$ is unique $\Rightarrow$ linear comb. of any $\mathbf{v}$ is unique.

Suppose for some $\mathbf{v} \in V$ there are two different expressions

$$\mathbf{v} = c_1\mathbf{v}_1 + ... + c_k\mathbf{v}_k$$

and

$$\mathbf{v} = d_1\mathbf{v}_1 + ... + d_k\mathbf{v}_k.$$

Then subtracting, we obtain

$$\mathbf{0} = (c_1 - d_1)\mathbf{v}_1 + ... + (c_k - d_k)\mathbf{v}_k,$$

and so the zero vector has a nontrivial representation as a linear combination of $\mathbf{v}_1, ..., \mathbf{v}_k$, since different expressions means that at least one difference $(c_i - d_i)$ is nonzero. This means there is a nontrivial expression of $\mathbf{0}$, contradiction. □

Proof: 2. Linear comb. of any nonzero $\mathbf{v}$ is unique $\Rightarrow$ linear comb. of $\mathbf{0}$ is also unique.

Conversely, suppose there is a nontrivial linear combination

$$\mathbf{0} = s_1\mathbf{v}_1 + ... + s_k\mathbf{v}_k.$$

Then, given any vector $\mathbf{v} \in V$, we can express $\mathbf{v}$ as a linear combination of $\mathbf{v}_1, ..., \mathbf{v}_k$ in several ways: for instance, adding

$$\mathbf{v} = c_1\mathbf{v}_1 + ... + c_k\mathbf{v}_k$$

and

$$\mathbf{0} = s_1\mathbf{v}_1 + ... + s_k\mathbf{v}_k,$$

we obtain another formula for $\mathbf{v}$, namely,

$$\mathbf{v} = (c_1 + s_1)\mathbf{v}_1 + ... + (c_k + s_k)\mathbf{v}_k,$$

contradiction. This completes the proof.

This discussion leads us to make the following concept

Definition

The set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is called *linear independent* if

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k = \mathbf{0} \Rightarrow c_1 = ... = c_k = 0,$$

i.e., if the only way of expressing the zero vector as a linear combination of $\mathbf{v}_1, ..., \mathbf{v}_k$ is the trivial linear combination $0\mathbf{v}_1 + ... + 0\mathbf{v}_k$.

Importantly, so if you are to prove a set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is linear independent, you should write

Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k = \mathbf{0}$. I must show that $c_1 = ... = c_k = 0$.

Example

We wish to decide whether the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^4$$

form a linearly independent set.

Example

Suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Show that if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, the so is $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{w}\}$.

Proposition

Suppose $\mathbf{v}_1, .., \mathbf{v}_k \in \mathbb{R}^n$ *form a linearly independent set, and suppose* $\mathbf{x} \in \mathbb{R}^n$*. Then* $\{\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{x}\}$ *is linearly independent if and only if* $\mathbf{x} \notin \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$*.*

Proof

We can prove the contrapositive: Supposing that $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$ form a linearly independent set,

$$\{\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{v}\} \text{ is linearly dependent iff } \mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k).$$

Suppose that $\mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$. Then $\mathbf{v} = c_1\mathbf{v}_1 + ... + c_k\mathbf{v}_k$ for scalars $c_1, ..., c_k$, so

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k + (-1)\mathbf{v} = \mathbf{0},$$

which implies that $\{\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{v}\}$ is linearly dependent.

Now suppose $\{\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{v}\}$ is linearly dependent. This means that there are scalars $c_1, ..., c_k, c$, not all 0, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c\mathbf{v} = \mathbf{0}.$$

Apparently, c cannot equal to 0, otherwise linearly independence of $\mathbf{v}_1, ..., \mathbf{v}_k$ would end up with an contraction. So dividing by c, we have

$$\mathbf{v} = -\frac{1}{c}(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k)$$

which tell us that $\mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$.

Definition

Let $V \subset \mathbb{R}^n$ be a subspace. The set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is called a *basis* for V if

i. $\mathbf{v}_1, ..., \mathbf{v}_k$ span V; i.e., $V = \text{Span}(\mathbf{v}_1, ..., \mathbf{v}_k)$, and

ii. $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is linearly independent.

Example

Let $\mathbf{e}_1 = (1, ..., 0)^\top, \mathbf{e}_2 = (0, 1, 0, ..., 0)^\top, ..., \mathbf{e}_n = (0, ..., 0, 1)^\top \in \mathbb{R}^n$. Then $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$, called the *standard basis*. To check this, we must establish that properties (i) and (ii) hold for $V = \mathbb{R}^n$.

Example

Consider the plane given by $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 - x_2 + 2x_3 = 0\} \subset \mathbb{R}^3$, and two vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and } \mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Let c_1, c_2 be two real numbers,

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}$$

which gives

$$\begin{bmatrix} c_1 - 2c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $\{\mathbf{v}_1, \mathbf{v}_2\}$ forms a basis of V.

Corollary

Let $V \subset \mathbb{R}^n$ *be a subspace, and let* $\mathbf{v}_1, \ldots \mathbf{v}_k \in V$*. Then* $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *is a basis for* V *if and only if every vector of* V *can be written uniquely as a linear combination of* $\mathbf{v}_1, \ldots, \mathbf{v}_k$*.*

Definition

When we write $\mathbf{v} = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k$, we refer to $c_1, \ldots, c_k$ as the *coordinates* of $\mathbf{v}$ with respect to the ordered basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

Example

Consider the three vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

Take a general vector $\mathbf{b} \in \mathbb{R}^3$, find a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Forming the augmented matrix and row reducing, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 2 & 1 & 0 & b_2 \\ 1 & 2 & 2 & b_3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2b_1 - b_3 \\ 0 & 1 & 0 & -4b_1 + b_2 + 2b_3 \\ 0 & 0 & 1 & 3b_1 - b_2 - b_3 \end{array}\right]$$

Thus an arbitrary vector $\mathbf{b} \in \mathbb{R}^3$ can be written in the form

$$\mathbf{b} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

where

$$\begin{aligned} c_1 &= 2b_1 - b_3 \\ c_2 &= -4b_1 + b_2 + 2b_3 \\ c_3 &= 3b_1 - b_2 - b_3. \end{aligned}$$

This process also gives a standard way of finding coordinates of $\mathbf{b}$ with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

The following is an important fact.

Proposition

Let A be an $n \times n$ matrix. Then A is nonsingular if and only if its column vectors form a basis for $\mathbb{R}^n$.

Proof.

Denote the column vectors of A by $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$. Using corollary above, we are to prove that A is nonsingular if and only if every vector in $\mathbb{R}^n$ can be written uniquely as a linear combination of $\mathbf{a}_1, ..., \mathbf{a}_n$. □

The following theorem tells us that every subspace has a basis.

Theorem

Any subspace $V \subset \mathbb{R}^n$ other than the trivial subspace has a basis.

Proof.

Since $V \neq \{\mathbf{0}\}$, we can choose a nonzero vector $\mathbf{v}_1 \in V$. If $\mathbf{v}_1$ spans V, then we know $\{\mathbf{v}_1\}$ is a basis for V. If not, choose $\mathbf{v}_2 \notin \text{Span}(\mathbf{v}_1)$. Previous proposition asserts that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent. If $\mathbf{v}_1, \mathbf{v}_2$ span V, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ will be a basis for V. If not, choose $\mathbf{v}_3 \notin \text{Span}(\mathbf{v}_1, \mathbf{v}_2)$. We know that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and hence will form a basis for V if the three vectors span V. We continue in this fashion, and we are guaranteed that the process will terminate in at most n steps because once we have $n+1$ vectors in $\mathbb{R}^n$, they must be linearly dependent. $\square$

The following result shows that all bases for a given subspace have one thing in common: They all consist of the same number of elements.

Proposition

Let $V \subset \mathbb{R}^n$ be a subspace, let $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a basis for V, and let $\mathbf{w}_1, ..., \mathbf{w}_l \in V$. If $l > k$, then $\{\mathbf{w}_1, ..., \mathbf{w}_l\}$ must be linearly dependent.

Proof.

Each vector in V can be written uniquely as a linear combination of $\mathbf{v}_1, ..., \mathbf{v}_k$. So let's write each vector $\mathbf{w}_1, ..., \mathbf{w}_l$ as such:

$$\mathbf{w}_1 = a_{11}\mathbf{v}_1 + ... + a_{k1}\mathbf{v}_k$$

$$\vdots$$

$$\mathbf{w}_l = a_{1l}\mathbf{v}_1 + ... + a_{kl}\mathbf{v}_k.$$

We now form the $k \times l$ matrix $A = [a_{ij}]$. This gives the matrix equation □

$$\begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_k \\ | & | & & | \end{bmatrix} A = \begin{bmatrix} | & | & & | \\ \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_l \\ | & | & & | \end{bmatrix}.$$

Since $l > k$, there cannot be a pivot in every column of A, and so there is a nonzero vector $\mathbf{c}$ satisfying

$$A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix} = \mathbf{0}.$$

Using the original equation and associativity, we have

$$\begin{bmatrix} | & | & & | \\ \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_l \\ | & | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_k \\ | & | & & | \end{bmatrix} \left(A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix} \right) = \mathbf{0}.$$

This is, we have found a nontrivial linear combination

$$c_1\mathbf{w}_1 + ... + c_l\mathbf{w}_l = \mathbf{0},$$

which means that $\{\mathbf{w}_1, ..., \mathbf{w}_l\}$ is linearly dependent, as we claimed.
This proposition leads directly to the main result.

Theorem

Let $V \subset \mathbb{R}^n$ be a subspace, and let $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ and $\{\mathbf{w}_1, .., \mathbf{w}_l\}$ be two bases for V. Then we have $k = l$.

Proof.

Since $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ forms a basis for V and $\{\mathbf{w}_1, ..., \mathbf{w}_l\}$ is known to be linearly independent, we use proposition to conclude that $l < k$. Since $\{\mathbf{w}_1, ..., \mathbf{w}_l\}$ is a basis for V and $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is known to be linearly independent, so we have that $k \leq l$. Both inequalities hold only if they are equal. □

Definition

The *dimension* of a subspace $V \subset \mathbb{R}^n$ is the number of vectors in any basis for V. We denote the dimension of V by $\dim V$. By convention, $\dim\{\mathbf{0}\} = 0$.

Four fundamental subspaces

Definition (Nullspace)

Let A be an $m \times n$ matrix. The *nullspace* of A is the set of solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$:

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

Definition (Column space)

Let A be an $m \times n$ matrix with column vectors $\mathbf{a}_1, ..., \mathbf{a}_n \in \mathbb{R}^m$. We define the *column space* of A to be the subspace of $\mathbb{R}^m$ spanned by the column vectors:

$$\mathbf{C}(A) = \mathrm{Span}(\mathbf{a}_1, ..., \mathbf{a}_n) \subset \mathbb{R}^m.$$

Proposition

Let A be an $m \times n$ matrix. Let $\mathbf{b} \in \mathbb{R}^m$. Then $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. That is,

$$\mathbf{C}(A) = \{\mathbf{b} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{b} \textit{ is consistent}\}.$$

Proof.

By definition, $\mathbf{C}(A) = \text{Span}(\mathbf{a}_1, ..., \mathbf{a}_n)$, and so $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b}$ is a linear combination of the vectors $\mathbf{a}_1, ..., \mathbf{a}_n$, i.e., $\mathbf{b} = x_1\mathbf{a}_1 + ... + x_n\mathbf{a}_n$ for some scalars $x_1, ..., x_n$. We conclude that $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. Finally, the system $A\mathbf{x} = \mathbf{b}$ is consistent provided it has a solution. □

Definition (Row space)

Let A be an $m \times n$ matrix with row vectors $\mathbf{A}_1, ..., \mathbf{A}_m \in \mathbb{R}^n$. We define the *row space* of A to be the subspace of $\mathbb{R}^n$ spanned by the row vectors $\mathbf{A}_1, ..., \mathbf{A}_m$:

$$\mathbf{R}(A) = \text{Span}(\mathbf{A}_1, ..., \mathbf{A}_m) \subset \mathbb{R}^n.$$

Noting that $\mathbf{R}(A) = \mathbf{C}(A^\top)$, it is natural to complete the rest as follows:

Definition (Left nullspace)

We define the *left nullspace* of the $m \times n$ matrix A to be

$$\mathbf{A}(A^\top) = \{\mathbf{x} \in \mathbb{R}^m : A^\top \mathbf{x} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x}^\top A = \mathbf{0}^\top\}.$$

Proposition

Let A be an $m \times n$ matrix. Then $\mathbf{N}(A) = \mathbf{R}(A)^\perp$

Proof.

If $\mathbf{x} \in \mathbf{N}(A)$, then $\mathbf{x}$ is orthogonal to each row vector $A_1, \dots, A_m$ of A. Then $\mathbf{x}$ is orthogonal to every vector in $\mathbf{R}(A)$ and is therefore an element of $\mathbf{R}(A)^\perp$. Thus $\mathbf{N}(A)$ is a subset of $\mathbf{R}(A)^\perp$, and so we need to show that $\mathbf{R}(A)^\perp$ is a subset of $\mathbf{N}(A)$. If $\mathbf{x} \in \mathbf{R}(A)^\perp$, this means that $\mathbf{x}$ is orthogonal to every vector in $\mathbf{R}(A)$, so $\mathbf{x}$ is orthogonal to each of the row vector $A_1, \dots, A_m$. But this means that $A\mathbf{x} = \mathbf{0}$, so $\mathbf{x} \in \mathbf{N}(A)$. □

Since $\mathbf{C}(A) = \mathbf{R}(A^\top)$, when we substitute $A^\top$ for A, we have

Proposition

Let A be an $m \times n$ matrix. Then $\mathbf{N}(A^\top) = \mathbf{C}(A)^\perp$.

Proposition

Let A be an $m \times n$ matrix. Then $\mathbf{C}(A) = \mathbf{N}(A^\top)^\perp$.

Proof.

Since $\mathbf{C}(A)$ and $\mathbf{N}(A^\top)$ are orthogonal subspaces, we have that $\mathbf{C}(A) \subset \mathbf{N}(A^\top)^\perp$. On the other hand, there is a system of constraint equations

$$\mathbf{c}_1 \cdot \mathbf{b} = ... = \mathbf{c}_k \cdot \mathbf{b} = 0$$

that give necessary and sufficient conditions for $\mathbf{b} \in \mathbb{R}^m$ to belong to $\mathbf{C}(A)$. Setting $V = \mathrm{Span}(\mathbf{c}_1, ..., \mathbf{c}_k) \subset \mathbb{R}^m$, this means that $\mathbf{C}(A) = V^\perp$. Since each such vector $\mathbf{c}_j$ is an element of $\mathbf{C}(A)^\perp = \mathbf{N}(A^\top)$, we conclude that $V \subset \mathbf{N}(A^\top)$. It follows that $\mathbf{N}(A^\top)^\perp \subset V^\perp = \mathbf{C}(A)$. Combining the two inclusions, we have $\mathbf{C}(A) = \mathbf{N}(A^\top)^\perp$. □

The following theorem summarizes geometric relations of the pairs of the four fundamental spaces.

Theorem

Let A be an $m \times n$ matrix. Then

1. $\mathbf{R}(A)^{\perp} = \mathbf{N}(A)$
2. $\mathbf{N}(A)^{\perp} = \mathbf{R}(A)$
3. $\mathbf{C}(A)^{\perp} = \mathbf{N}(A^{\top})$
4. $\mathbf{N}(A^{\top})^{\perp} = \mathbf{C}(A)$