Lecture 3

Linear independence

Suppose we choose $v_1 \dots v_k \in V$ to construct a subspace $W \subset V$ by spanning those vectors, $W = \mathrm{span}(v_1 \dots v_k)$. Are all $v_1 \dots v_k$ essential ? In other words, can we get the same W by dropping some of the original vectors ?

Sometimes yes : $\mathrm{span}\left(\binom{1}{0}, \binom{0}{1}, \binom{1}{1}\right) = \mathrm{span}\left(\binom{1}{0}, \binom{0}{1}\right)$, vector $\binom{1}{1}$ is redundant here.

Sometimes no : dropping $\binom{1}{0}$ or $\binom{0}{1}$ will change $W = \mathrm{span}\left(\binom{1}{0}, \binom{0}{1}\right)$.

both vectors are essential.

The redundant vectors can be identified using the fundamental notion of linear independence / dependence.

Def The vectors $v_1 \dots v_k \in V$ are called linearly dependent if $\exists\, \alpha_1 \dots \alpha_k$ (not all zero!), such that $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$.

The vectors are linearly independent if they are not linearly dependent (i.e. $\forall\, \alpha_1 \dots \alpha_k$ (not all zero), $\alpha_1 v_1 + \dots + \alpha_k v_k \neq 0$)

Example 1 Let $\mathbb{T}^{(2)} = \mathrm{span}(1, \cos x, \sin x, \cos^2 x, \cos x \sin x, \sin^2 x)$ be the space of quadratic trigonometric polynomials. The generating vectors are dependent since $\sin^2 x + \cos^2 x = 1$: $\underline{(-1)} \cdot 1 + \underline{0} \cdot \cos x + \underline{0} \cdot \sin x + \underline{1} \cdot \cos^2 x + \underline{0} \cdot \cos x \sin x + \underline{1} \cdot \sin^2 x = 0$

Example 2 $\mathbb{T}^{(2)} = \mathrm{span}(\cos x, \sin x, \cos^2 x, \cos x \sin x, \sin^2 x)$ and the corresponding vectors are linearly independent (HW).

Remark General strategy for proving linear independence : assume that $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$ and show that this leads to $\alpha_1 = \dots = \alpha_k = 0$ (trivial linear combination.)

Let us now focus on linear (in)dependence of $v_1 \dots v_n \in \mathbb{R}^m$

Theorem 3 Let $v_1 \dots v_n \in \mathbb{R}^m$ and $A = [v_1 \dots v_n]$ is the $m \times n$ matrix

1. $v_1 \dots v_n$ are lin. dep. $\Leftrightarrow$ $\exists\, \alpha \neq 0, \alpha \in \mathbb{R}^n$: $A\alpha = 0$ $\Leftrightarrow$ $\mathrm{rank}\, A < n$
2. $v_1 \dots v_n$ are lin. ind. $\Leftrightarrow$ $A\alpha = 0$ only if $\alpha = 0$ $\Leftrightarrow$ $\mathrm{rank}\, A = n$
3. $v \in \mathrm{span}(v_1 \dots v_n)$ $\Leftrightarrow$ $A\alpha = v$ has at least one solution. $\Leftrightarrow$ $\mathrm{rank}\,[A|v] = \mathrm{rank}\,[A]$

This is basically reformulation of the definitions of lin. dep/ind. and span on the matrix language.

Let us look at the homogeneous linear system $A\alpha = 0$ more carefully. It has m equations and n unknowns. Recall that

$$\exists \alpha \neq 0 \; : \; A\alpha = 0 \iff \operatorname{rank} A < n$$

(if $\operatorname{rank} A = n \Rightarrow$ unique sol. $\alpha = 0$)

1. If $n > m \Rightarrow \operatorname{rank} A < n \Rightarrow v_1 \dots v_n \in \mathbb{R}^m$ are linearly dependent
2. If $n \leq m$, then $v_1 \dots v_n \in \mathbb{R}^m$ are lin. ind. (dep.) $\iff \operatorname{rank} A = n$ $(\operatorname{rank} A < n)$

<u>When $v_1 \dots v_n$ span all of $\mathbb{R}^m$?</u>

This happens when $\forall v \in \mathbb{R}^m$ the system $A\alpha = v$ has at least one solution. This happens if and only if the row echelon form of A does not contain all zero rows, i.e. if <u>$\operatorname{rank} A = m$</u>.

<u>Remark</u> A necessary (but not sufficient) condition for $\operatorname{span}(v_1 \dots v_n) = \mathbb{R}^m$ is $n \geq m$.

Our discussion of linear dependence/independence and span gets us very close to the fundamental notion of basis.

The notion of basis is absolutely fundamental in linear algebra and its applications, including differential equations, statistics, signal and image processing, data compression, etc. It formalizes the idea of the most "economical" set of vectors that spans a given vector space V.

If we want to span V by $v_1 \dots v_k \in V$, then k must be large enough (recall that to span $\mathbb{R}^n$, $k \geq n$).

On the other hand, if k is too large, then $v_1 \dots v_k$ will be linearly dependent, and, therefore, some of the vectors will be redundant, making the set $v_1 \dots v_k$ not "economical."

Def 1 Let V be a vector space. A collection of vectors $v_1, \dots, v_n \in V$ is called a basis of V if

1. $\operatorname{span}(v_1 \dots v_n) = V$ (we have enough vectors)
2. $v_1 \dots v_n$ are linearly independent (just enough)

Ex 1 The standard basis of $\mathbb{R}^n$ is $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$

The standard basis consists of n vectors. It turns out that so does every basis of $\mathbb{R}^n$

Th 4 Every basis of $\mathbb{R}^n$ consists of exactly n vectors.

Proof: Let $v_1 \dots v_k$ be a basis of $\mathbb{R}^n$. Since $\operatorname{span}(v_1 \dots v_k) = \mathbb{R}^n \Rightarrow k \geq n$.
On the other hand, if $k > n \Rightarrow v_1 \dots v_k$ are linearly dependent

Hence, $k = n$ $\square$

How to check whether $v_1 \dots v_n \in \mathbb{R}^n$ form a basis?

Th 5 A set of vectors $v_1 \dots v_n \in \mathbb{R}^n$ form a basis of $\mathbb{R}^n$ $\iff$ $\operatorname{rank} A = n$

$A = [v_1, \dots, v_n]$

(we established this [redacted] [redacted] when discussed when $v_1 \dots v_k \in \mathbb{R}^n$ are lin. ind & span $\mathbb{R}^n$)

The fact that every basis of $\mathbb{R}^n$ consists of the same number of vectors is not special for $\mathbb{R}^n$, it is true for every vector space

Th 6 Let $v_1 \dots v_n$ be a basis in a vector space V.
Then every basis of V consists of n vectors.

Proof: Suppose $u_1 \dots u_m$ is another basis of V and $m \neq n$.
Without loss of generality, assume $m > n$. We will show that $\{u_j\}$ are lin. dep.
Since $v_1 \dots v_n$ is a basis, $u_j \in \operatorname{span}(v_1 \dots v_n) \Rightarrow u_j = \sum_{i=1}^{n} \alpha_{ij} v_i$ for $j = \overline{1,m}$.
Let us form a linear combination of $u_1 \dots u_m$ and see when it is zero.

$$\sum_{j=1}^{m} \beta_j u_j = \sum_{j=1}^{m} \sum_{i=1}^{n} \beta_j \alpha_{ij} v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \alpha_{ij} \beta_j \right) v_i = 0 \underset{\text{since } \{v_i\} \text{ are lin. ind.}}{\iff} \sum_{j=1}^{m} \alpha_{ij} \beta_j = 0 \quad \forall i = \overline{1,n}$$

Note that $\sum_{j=1}^{m} \alpha_{ij} \beta_j = 0 \quad \forall i = \overline{1,n}$ is a homogeneous system of n equations with m unknowns : $\beta_1 \dots \beta_m$. But we know $\left(\begin{smallmatrix}\text{since}\\ m>n\end{smallmatrix}\right)$ $\underset{n}{}\underbrace{\boxed{\ \alpha\ }}_{m}\ \boxed{\beta} = 0$ that in this case $\exists$ a non-trivial solution $\beta \neq 0$

$\Rightarrow \exists$ non-trivial lin. combination $\sum_{j=1}^{m} \beta_j u_j = 0 \Rightarrow u_1 \dots u_m$ are lin. dep. $\Rightarrow$ we can't have $m>n$ $\square$

(similarly, we can't have $m<n$)

This result motivates the following definition

<u>Def 2</u> : The number of vectors in a basis of a vector space V is called the <u>dimension</u> of V.

<u>Example 2</u> : The standard basis of the space $\mathbb{P}^{(n)}$ of all polynomials of deg $\leq n$ is $1, x, x^2, \dots, x^n$. The dimension is thus $\dim \mathbb{P}^{(n)} = n+1$.

Now we can formulate precisely what we mean when we say that a basis optimally (economically) represents a vector space.

<u>Th 4</u> : Let V be an n-dimensional vector space. Then

1. Every set of $m>n$ vectors is linearly dependent. $\leftarrow$ we just saw this in the proof of Th. 3.
2. No set of $m<n$ vectors spans V. $\leftarrow$ If $\exists$ such a set, then $v_1 \dots v_n$ - the basis of V - are lin. dep. (same argument as in Th 3.)

According to definition, to check whether a set of vectors is a basis of V, we need to check that a) vectors are lin. ind. b) they span V.

If we know that $\dim V = n$ and we have n vectors (only sets of n vectors can be potentially bases) then we need to check just one of the two conditions, the other will be fullfilled automatically.

<u>Th 5</u> Let V be an n-dimensional vector space.

1. Vectors $v_1 \dots v_n$ form a basis $\Longleftrightarrow$ $v_1 \dots v_n$ are lin. ind
2. Vectors $v_1 \dots v_n$ form a basis $\Longleftrightarrow$ $\text{span}(v_1 \dots v_n) = V$

Proof : In both cases, we need to prove only one direction : $\Leftarrow$

1. Suppose $v_1 \dots v_n$ are lin. ind. We need to show that $\text{span}(v_1 \dots v_n) = V$
Assume $\text{span}(v_1 \dots v_n) \neq V \Rightarrow \exists u \in V : u \notin \text{span}(v_1 \dots v_n)$. Consider a set of $(n+1)$ vectors : $(v_1 \dots v_n, u)$. We claim they are lin. ind. Indeed, if $\alpha_1 v_1 + \dots + \alpha_n v_n + \beta u = 0 \Rightarrow \beta = 0$ (otherwise $u \in \text{span}(v_1 \dots v_n)$), but then all $\alpha_i = 0$ since v_i are lin. ind.
Contradiction with Th 4. 1) $\Rightarrow \text{span}(v_1 \dots v_n) = V$ and $v_1 \dots v_n$ form a basis of V.

2. Suppose now that $v_1 \dots v_n$ span V. We need to show that $v_1 \dots v_n$ are lin. ind.
Assume they are not, i.e. $\exists$ non-trivial lin. combination
$\alpha_1 v_1 + \dots + \alpha_n v_n = 0$, where not all α_i are zero. Assume $\alpha_n \neq 0$ (does not matter which α_i)
$\Rightarrow$ $v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \dots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1} \in \operatorname{span}(v_1 \dots v_{n-1})$, in other words, v_n is redundent.
Thus, $\operatorname{span}(v_1 \dots v_{n-1}) = V$, which contradicts to Th 4. 2)
So, $v_1 \dots v_n$ are lin. ind, and, therefore, form a basis of V. $\square$

In Th5, we assumed that $\dim V$ is known and made some statements about the properties of sets of vectors. Suppose now, that we don't know $\dim V$, but we do know that a certain set of vectors is lin. ind. or that it spans V.
Can we say anything about $\dim V$?

Th 6 Let V be a vector spase, and $v_1 \dots v_n \in V$.

1. If $v_1 \dots v_n$ are lin. ind. $\Rightarrow \dim V \geq n$
2. If $\operatorname{span}(v_1 \dots v_n) = V \Rightarrow \dim V \leq n$

Follows directly from Th. 4.

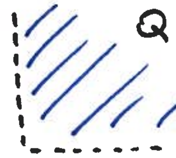
Let us finish the discussion of bases with the following useful characterization.

Th 7 The vectors $v_1 \dots v_n$ form a basis of V $\Leftrightarrow$
$\forall v \in V \ \exists! \alpha_1 \dots \alpha_n : v = \sum_{i=1}^{n} \alpha_i v_i$ (any vector $v \in V$ can be written uniquely as a linear combination of $v_1 \dots v_n$)

Proof: ($\Rightarrow$) Since $v_1 \dots v_n$ is a basis, $\operatorname{span}(v_1 \dots v_n) = V$,
and therefore $\forall v \in V \ \exists \alpha_1 \dots \alpha_n : v = \sum_{i=1}^{n} \alpha_i v_i$ (there exists at least one representation as a lin. combination)
Suppose $\exists$ another representation of v as a lin. combination of $v_1 \dots v_n$: $v = \sum \tilde{\alpha}_i v_i$.
But then $\sum_{i=1}^{n} (\alpha_i - \tilde{\alpha}_i) v_i = 0 \Rightarrow \alpha_i - \tilde{\alpha}_i = 0$ (since $v_1 \dots v_n$ are independent)
Thus, the representation $v = \sum_{i=1}^{n} \alpha_i v_i$ is unique.

($\Leftarrow$) We know that $\operatorname{span}(v_1 \dots v_n) = V$ and need to show that $v_1 \dots v_n$ are lin. ind.
Suppouse they are not: $\exists \alpha_1 \dots \alpha_n$ (not all zero):
$\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ Let say $\alpha_n \neq 0 \Rightarrow v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \dots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1}$
On the other hand, $v_n = 0 \cdot v_1 + \dots + 0 \cdot v_{n-1} + 1 \cdot v_n$.
And we constructed two different representations of v_n as a lin. combination of $v_1 \dots v_n$. Contradiction! $\Rightarrow v_1 \dots v_n$ are lin. ind. $\square$

Def3 $(\alpha_1 \dots \alpha_n)$ are called the coordinates of v in the basis $v_1 \dots v_n$.



Ex 3: Recall our "weird" example of a vector space $(x_1,y_1)\oplus(x_2 y_2) = (x_1x_2, y_1y_2)$, $\alpha\cdot(x,y) = (x^\alpha, y^\alpha)$. Let us find its dimension and a basis. It "feels" like Q should be 2-dim. Since the vector space operations $\oplus$ and $\cdot$ are based on (ordinary) multiplication, its reasonable to use 1 and e to construct a basis. So, let us define $v_1 = (1, e)$ and $v_2 = (e, 1)$.

Let $(x,y)\in Q$ be any vector, let us try to represent it as a lin. comb. of v_1 and v_2

$$(x,y) = \alpha v_1 \oplus \beta v_2 = (1, e^\alpha)\oplus(e^\beta, 1) = (e^\beta, e^\alpha).$$

This equation has unique solution: $\begin{cases}\alpha = \log y\\ \beta = \log x\end{cases} \Rightarrow v_1, v_2$ is indeed a basis and (α,β) are the coord. of (x,y) in that basis.

Thus, $\dim Q = 2$.

The Fundamental Matrix Subspaces

Let $A\in M_{m\times n}$ be an $m\times n$ matrix. There are 4 $\underbrace{\text{fundamental subspaces}}_{}$ associated with A: kernel (null), image (range), cokernal, and coimage.

many results assoc. with A can be formul. in terms of these spaces.

We can view A as a (linear) map $A: \mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$.

Def 4 The kernel of A is the subspace $\ker A \subset \mathbb{R}^n$;

$$\boxed{\ker A = \{x\in\mathbb{R}^n \mid Ax = 0\}}$$

Let us check that $\ker A$ is indeed a subspace. Let $x, y\in\ker A$ and $\alpha,\beta\in\mathbb{R}$

$\Rightarrow \quad A(\alpha x+\beta y) = \alpha Ax + \beta Ay = 0 \quad\Rightarrow\quad \alpha x+\beta y\in\ker A.$

The fact that $\ker A$ is a subspace can be re-expressed as the following

Superposition Principle (for homogeneous systems)

If $x_1 \dots x_k$ are solutions of $Ax = 0 \Rightarrow \sum_{i=1}^{k}\alpha_i x_i$ is also a solution for any $\alpha_1\dots\alpha_n\in\mathbb{R}$.

Why is this principle important?

Because it tells us that if we want to find a general solution of $Ax=0$, we need to find a basis in $\ker A$, and then any solution is a lin. combination of the basis elements.

Def 5 The image of A is the subspace $\operatorname{im} A\subset\mathbb{R}^m$,

$$\boxed{\operatorname{im} A = \{Ax \mid x\in\mathbb{R}^n\}}$$

Let us also check that $\operatorname{im} A$ is a subspace. Let $b_1, b_2\in\operatorname{im}A \Rightarrow \exists\, x_1, x_2\in\mathbb{R}^n$: $Ax_1 = b_1$, $Ax_2 = b_2$

$\Rightarrow \forall\,\alpha,\beta \quad A(\alpha x_1 + \beta x_2) = \alpha b_1 + \alpha b_2 \Rightarrow \alpha b_1 + \beta b_2 \in \operatorname{im} A.$

- The definition of $\operatorname{im} A$ is a bit implicit, but it has a very simple interpretation. Let $a_1 \dots a_n$ be the columns of A, then $\operatorname{im} A$ consists of all vectors $b = Ax = [a_1 \dots a_n]\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 a_1 + \dots + x_n a_n \,, \ \forall x$

  Therefore, $\boxed{\operatorname{im} A = \operatorname{span}(a_1, \dots, a_n)}$

- The image of A is important because it is exactly the set of vectors b, for which the linear system $Ax = b$ has solutions.

  It turns out that to find the general solution of $Ax = b$, all we need to know is one particular solution and $\ker A$.

<u>Th 8</u> Let x^* be a solution of $Ax = b$ (we assume $b \in \operatorname{im} A$).

Then $x \in \mathbb{R}^n$ is a solution $\Longleftrightarrow$ $x = x^* + x_0$, where $x_0 \in \ker A$.

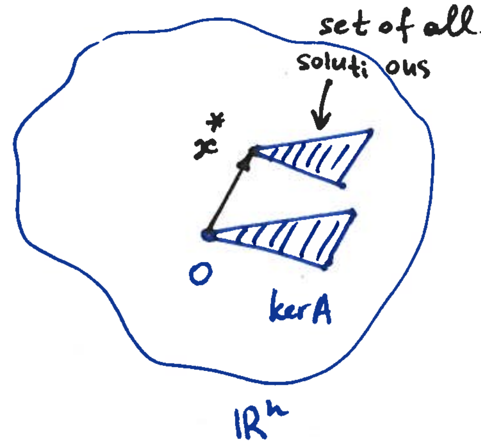
<u>Proof</u> $(\Leftarrow)$ is easy: $A(x^* + x_0) = \underbrace{Ax^*}_{b} + \underbrace{Ax_0}_{0} = b$.

$(\Rightarrow)$ If x is a solution $\Rightarrow Ax = b$. We also know that $Ax^* = b$

Therefore, $A(x - x^*) = 0 \Rightarrow x - x^* \equiv x_0 \in \ker A$ $\square$

- The existence of a solution of $Ax = b$ depends on whether $b \in \operatorname{im} A$ or $b \notin \operatorname{im} A$.
- But if $b \in \operatorname{im} A$, a solution exists and its uniqueness does not depend on a particular b:

$$\boxed{\text{Solution of } Ax = b \text{ is unique} \Longleftrightarrow \ker A = \{0\}}$$



The superposition principle lies at the heart of linearity. For homogeneous systems: it allows to construct new solutions of $Ax=0$ by simply taking linear combinations of old solutions. For inhomogeneous systems $Ax=b$ this of course does not work, and the s.p. takes the following form:

Superposition Principle (for inhomogeneous systems)

Let $x_1^*, \dots, x_k^*$ be particular solutions of $Ax=b_1, \dots, Ax=b_k$

$\Rightarrow$ $x^* = \sum_{i=1}^{k} \alpha_i x_i^*$ is a particular solution of $Ax = \sum_{i=1}^{k} \alpha_i b_i$ (here we assume that $b_i \in \operatorname{im} A$)

$\Rightarrow$ The general solution of $Ax = \sum_{i=1}^{k} \alpha_i b_i$ is then $x = x^* + x_0$, $x_0 \in \ker A$.

Useful interpretation

In physical applications, $b_1, \dots, b_k$ usually represent external forces, and the solutions $x_1^* \dots x_k^*$ represent the corresponding responses of the physical system. The s.p. tells us that we know how the system responds to the individual excitations, then we also know how it will respond to any combination of the exications. The precise details of the system are irrelevant: all is important is its linearity

Cokernel & Coimage

We introduced two (out of four) fundamental subspaces associated with any matrix $A \in \mathbb{M}_{m \times n}$: $\ker A \subset \mathbb{R}^n$ and $\operatorname{im} A \subset \mathbb{R}^m$.

Every system $Ax=b$ has a syster (or a brother if you wish), called adjoint.

Def The adjoint to $Ax=b$ is $A^T y = c$. Here, $A \in \mathbb{M}_{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A^T \in \mathbb{M}_{n \times m}$, $y \in \mathbb{R}^m$, $c \in \mathbb{R}^n$

The remaining two fundamental subspaces of A are the two known subspaces of A^T.

Def The cokernel of A is $\operatorname{coker} A = \ker A^T = \{ y \in \mathbb{R}^m \mid A^T y = 0 \} \subset \mathbb{R}^m$

The coimage of A is $\operatorname{coim} A = \operatorname{im} A^T = \{ A^T y \mid y \in \mathbb{R}^m \} \subset \mathbb{R}^n$

Recall that $\operatorname{im} A = \operatorname{span}(a_1, \dots, a_n)$, where a_i is the i^{th} column of A.

Therefore $\operatorname{coim} A = \operatorname{span}(a^1, \dots, a^m)$, where a^i is the i^{th} row of A.

Remark:

- $\operatorname{im} A$ is also called column space
- $\operatorname{coim} A$ is also called row space