

Module 8: Linear Algebra II

Operational math bootcamp



Statistical Sciences
UNIVERSITY OF TORONTO

University of Toronto

July 21, 2025

Outline

Last time:

- Vector spaces and subspaces
- Linear independence and bases
- Linear maps, null space, range

Today:

- Inverses of linear maps
- Matrices as linear maps
- Determinants *of matrices*
- Inner product spaces



Recall

Theorem

Suppose $\mathbf{u}_1, \dots, \mathbf{u}_n$ is a basis for U and $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis for V . Then there exists a unique linear map $T: U \rightarrow V$ such that $T\mathbf{u}_j = \mathbf{v}_j$ for $j = 1, \dots, n$.

or $V \rightarrow U$

Definition (Injective and surjective)

Let $T: U \rightarrow V$. T is *injective* if $T\mathbf{u} = T\mathbf{v}$ implies $\mathbf{u} = \mathbf{v}$, if and only if $\text{null } T = \{\mathbf{0}\}$. T is *surjective* if $\forall \mathbf{v} \in V, \exists \mathbf{u} \in U$ such that $\mathbf{v} = T\mathbf{u}$, i.e. if $\text{range } T = V$.

Theorem (Rank Nullity Theorem)

Let $T: U \rightarrow V$ be a linear transformation, where U and V are finite-dimensional vector spaces. Then

$$\text{rank } T + \text{nullity } T = \dim U.$$

$$\dim(\text{range}(T)) + \dim(\ker(T)) = \dim U$$



Definition (Product of linear maps)

Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. We define the product $TS \in \mathcal{L}(U, W)$ for $\mathbf{u} \in U$ as $TS(\mathbf{u}) = T(S(\mathbf{u}))$.
 $S(\mathbf{u}) \in V, \quad T(S(\mathbf{u})) \in W$

Definition

A linear map $T: U \rightarrow V$ is *invertible* if there exists a linear map $S: V \rightarrow U$ such that ST is the identity map on U and TS is the identity map on V . Such a map S is called the *inverse* of T .

If T is invertible, we denote the inverse ^{S} by T^{-1} . This is justified by the fact that the inverse is unique, as the next proposition shows.

'and' is necessary.

e.g. $T: \mathbb{Q} \rightarrow \mathbb{R}, T(q) = q$

doesn't exist a linear map S s.t. $ST = I_U$
even without linear constraints, we cannot find
a S s.t. $ST = I_U$ and $TS = I_V$

$\downarrow$
 $S(x) = \begin{cases} x, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}, ST = I_{\mathbb{Q}} \text{ but } TS(2) = 0 \neq 2$
 $\Rightarrow TS \neq I_{\mathbb{R}}$

$\triangle$ If both U and V are finite dimensional, ($T: U \rightarrow V$)
and $\dim U = \dim V$, then we have

$$ST = I_U \Leftrightarrow TS = I_V$$

$T: U \rightarrow V$ is invertible if $\exists S: V \rightarrow U$
 s.t. $ST = \text{Id}_U$ and $TS = \text{Id}_V$
 denotes the inverse as T^{-1}

Proposition

Any invertible linear map has a unique inverse.

Proof. Let $T: U \rightarrow V$ is invertible

Let S and S' be inverses of $T \Rightarrow ST = S'T = \text{Id}_U$, $TS = Ts' = \text{Id}_V$

For any $v \in V$

$$Sv = \text{Id}_U(\underbrace{Sv}_{\in U}) = (S'T)(Sv) = S'((TS)(v)).$$

$$= S'(\text{Id}_V(v))$$

$$= S'v$$

Thus $S = S'$

$$T: U \rightarrow V$$

Theorem

A linear map is invertible if and only if it is injective and surjective.

See proof in the book.

$$ST = Id_U \quad TS = Id_V$$

Definition

An invertible linear map is called an *isomorphism*. If there exists an isomorphism from one vector space to another, we say that the vector spaces are isomorphic.

$$U \text{ and } V$$

$\exists$ invertible linear map.

Theorem

Two finite-dimensional vector spaces U and V over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.

Proof. ($\Rightarrow$) Suppose U and V are isomorphic.

Then exists an invertible map $T: U \rightarrow V$

T is invertible $\Leftrightarrow T$ is injective and surjective (by previous theorem)

$$\Leftrightarrow \ker T = \{0\} \text{ and } \text{range}(T) = V$$

$$\Rightarrow \dim(\ker T) = 0 \text{ and } \dim(\text{range}(T)) = \dim V$$

By rank-nullity theorem,

$$\dim U = \dim(\ker T) + \dim(\text{range}(T))$$

$$= 0 + \dim V = \dim V$$

Basis

($\Leftarrow$)

Suppose $\dim U = \dim V = n$

Then we have a basis $\{u_1, u_2, u_3, u_4, \dots, u_n\}$ of U

and a basis $\{v_1, v_2, \dots, v_n\}$ of V

Define $T: U \rightarrow V$ by $Tu_i = v_i$ for $i=1, \dots, n$

Then T is invertible, i.e. $\exists S: V \rightarrow U$

s.t. $TS = \text{Id}$, $ST = \text{Id}$.

(Need check)!

Linear maps and matrices

(A matrix can be a linear map)

Example

Let $A \in M_{m \times n}$ be a fixed matrix. Then, we can define a linear map $T_A: \mathbb{F}^n \rightarrow \mathbb{F}^m$ via $T_A(\mathbf{v}) = A\mathbf{v}$, where we recall matrix vector multiplication $(A\mathbf{v})_i = \sum_{k=1}^n A_{ik}v_k$ for $i = 1, \dots, m$.
 $\forall \mathbf{v} \in \mathbb{F}^n$, it is a vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$

Next we will see that we can use matrices to represent linear maps between finite dimensional vector spaces.

For any linear map between finite-dimensional vector spaces,
they can be represented by matrices.
By basis.

Definition

Let $T \in \mathcal{L}(U, V)$ where U and V are vector spaces. Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ be bases for U and V respectively. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ with entries A_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n$ defined by

$$T\mathbf{u}_k = \underbrace{A_{1k}\mathbf{v}_1 + \dots + A_{mk}\mathbf{v}_m}_{\text{linear combination of basis}}, \quad \begin{array}{l} \forall \mathbf{u}_k \in U \\ T\mathbf{u}_k \in V \Rightarrow T\mathbf{u}_k \text{ can be written} \\ \text{as linear combination of} \\ \text{basis.} \end{array}$$

i.e. the k th column of A is the scalars needed to write $T\mathbf{u}_k$ as a linear combination of the basis of V :

$$T\mathbf{u}_k = \sum_{i=1}^m A_{ik}\mathbf{v}_i$$

Note that since a linear map $T \in \mathcal{L}(U, V)$ is uniquely determined by its image on a basis of U , we see that once we pick basis of U and V its matrix representation is uniquely determined. — given a fixed T

Example

Let $D \in \mathcal{L}(\mathbb{P}_4(\mathbb{R}), \mathbb{P}_3(\mathbb{R}))$ be the differentiation map, $Dp = p'$. Find the matrix of D with respect to the standard bases of $\mathbb{P}_3(\mathbb{R})$ and $\mathbb{P}_4(\mathbb{R})$.

Standard basis: $1, x, x^2, x^3, (x^4)$

$D \mathbf{T}(u_1) = 0 = \vec{0}$ u_1, u_1, u_1, u_1, u_5 $(u_i = x^{i-1})$

$D \mathbf{T}(u_2) = 1 = u_1$

$D \mathbf{T}(u_3) = 2x = 2u_2$

$D \mathbf{T}(u_4) = 3x^2 = 3u_3$

$D \mathbf{T}(u_5) = 4x^3 = 4u_4$

The matrix is:

$$M(D) = \begin{matrix} & \begin{matrix} D u_1 & D u_2 & D u_3 & D u_4 & D u_5 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{matrix} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}_{4 \times 5} \end{matrix} \in \mathbb{R}^{4 \times 5}$$

- Observe that if we choose bases $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ for U, V and represent $T \in \mathcal{L}(U, V)$ as a matrix $\mathcal{M}(T)$, then the corresponding map can be obtained by just working with the coordinates of vectors in U, V with respect to the chosen basis
 $\forall u \in U, u$ can be written as linear combination of $\mathbf{u}_i$
- If $\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$, then the coordinates of $T(\mathbf{u})$ with respect to $\mathbf{v}_1, \dots, \mathbf{v}_m$ can be obtained by the matrix vector multiplication $\mathcal{M}(T)\alpha$, where α is the $n \times 1$ matrix with entries α_i

$$T(u) \in V \Rightarrow T(u) = \sum_{i=1}^m \beta_i \mathbf{v}_i, \text{ what's correlation between } \beta_i, \alpha_i, T$$

α as $n \times 1$ matrix with entries α_i

β as follow,

$$\beta = \mathcal{M}(T) \cdot \alpha$$

it's just matrix multiplication.

Example

If we want to find the derivative of $p = x^4 + 12x^3 - 5x^2 + 7$ with respect to the standard monomial basis of $\mathbb{P}_4(\mathbb{R})$, we use $\mathcal{M}(D)$ from the previous example to obtain

$$\mathcal{M}(D)\alpha = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 7 \\ 0 \\ -5 \\ 12 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -10 \\ 36 \\ 4 \end{pmatrix}.$$

Thus, translating back into the monomial basis of $\mathbb{P}_3(\mathbb{R})$ gives

$$D(p) = -10x + 36x^2 + 4x^3.$$

Other points

- Looking at matrices as representations of linear maps gives us an intuitive explanation for why we do matrix multiplication the way we do! In fact, we want matrix multiplication to represent composition of linear maps
- We can use matrices to solve linear systems.

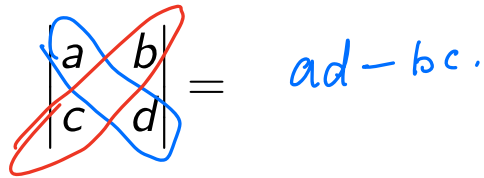
Determinants

Determinant

- The determinant is a function from $M_{n \times n} \rightarrow \mathbb{F}$, i.e. it is a function from the entries of a square matrix to a real or complex number.
- Notation: $\det(A) = |A|$
- The determinant has applications in solving linear systems, computing eigenvalues, etc

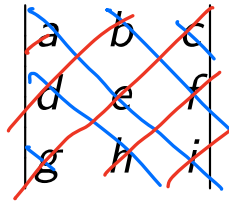
Example: 2×2 matrix

The determinant of a 2×2 matrix is


$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Example: 3×3 matrix

There is a **trick** for finding the determinant of a 3 by 3 matrix:



$$|A| = aei + bfg + cdh - ceg - fha - idb$$

Cofactor expansion

For other $n \times n$ matrices, one can compute the determinant using **cofactor expansion**.

Definition (Cofactor expansion)

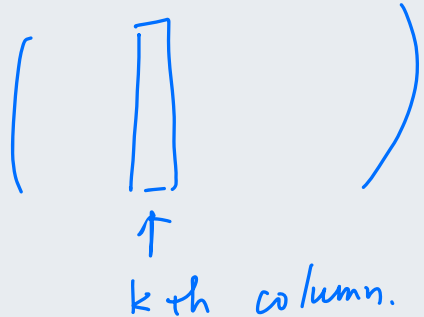
Let $A = \{a_{j,k}\}_{j,k=1}^n$ be a $n \times n$ matrix. Let $M_{j,k}$ denote the determinant of the $(n-1) \times (n-1)$ matrix obtained by removing the j^{th} row and the k^{th} column of A . For each row $j = 1, \dots, n$

$$|A| = \sum_{k=1}^n a_{j,k} (-1)^{j+k} M_{j,k}.$$



Similarly, for each column $k = 1, \dots, n$

$$|A| = \sum_{j=1}^n a_{j,k} (-1)^{j+k} M_{j,k}.$$



The numbers $C_{j,k} = (-1)^{j+k} M_{j,k}$ are called *cofactors*.

Similarly, for each column $k = 1, \dots, n$

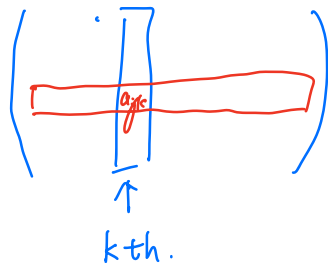
$$|A| = \sum_{j=1}^n a_{j,k} (-1)^{j+k} M_{j,k}.$$



The numbers $C_{j,k} = (-1)^{j+k} M_{j,k}$ are called *cofactors*.

Statistical Sciences

$$A_{n \times n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}, \quad \text{fix } k$$



$\forall j$, delete k th column and j th row,
we'll get a sub-matrix,
and the $\det(\sim) = M_{j,k}$

$$|A| = \sum_{j=1}^n a_{j,k} M_{j,k} (-1)^{j+k}$$

$\det$ of $n \times n$ matrix = combination of $\det$ of
 $(n-1) \times (n-1)$ matrix

"Idea of Induction"

Proposition

The determinant of a diagonal matrix or triangular matrix is the product of the entries on the diagonal.

By induction

Sketch of proof.

Let $A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ be upper triangular matrix (WLOG).

By induction. $n=1$, $A = (a_{11})$, $\det A = a_{11}$.

Suppose for $k-1$, we have $\det A = \prod_{i=1}^{k-1} a_{ii}$.

For n , by cofactor expansion w.r.t 1st column

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & & \\ 0 & a_{2k} & \dots & a_{kn} \end{pmatrix}$$

↑
expand with column.

△ we can also expand by k th. row

$$\det A = a_{11} (-1)^{1+1} \det M_{1,1}$$

$$= \det \begin{pmatrix} a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ 0 & a_{k1} & \dots & a_{kn} \end{pmatrix} = \sum_{i=2}^n a_{ii} \quad \text{by induction hypothesis.}$$

Thus, we can conclude ~

Inverse of a matrix

$C_{j,k} = M_{j,k}$ in the above.

Theorem

Let A be an $n \times n$ invertible matrix and let $C = \{C_{j,k}\}_{j,k=1}^n$ be its cofactor matrix. Then

$$A^{-1} = \frac{1}{|A|} C^T$$

Note: The matrix A is invertible if and only if the linear map represented by the matrix is an isomorphism.

is invertible.

Different .

Cramer's rule

$$A\mathbf{x} = \mathbf{b}$$

Corollary

Suppose A is an $n \times n$ invertible matrix. The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution given by

$$x_i = \frac{|A_i|}{|A|}, \quad i = 1, \dots, n,$$

where A_i is the matrix obtained by replacing the i^{th} column of A with $\mathbf{b}$.

$$A = (a_1, \dots, \textcircled{a_i}, \dots, a_n)$$

↓ replace

$$A_i = (a_1, \dots, \textcircled{b}, \dots, a_n)$$

Transpose of a matrix

Definition

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix, denoted A^T , defined entry-wise as $\{A_{j,k}^T\} = \{A_{k,j}\}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$ (i.e. the rows of A are the columns of A^T and the columns of A are the rows of A^T)

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}_{3 \times 2}, (A^T)_{2 \times 3} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Properties of the determinant

$$\Rightarrow \text{consider } \frac{1}{|A|} C^T, \quad A \cdot \left(\frac{1}{|A|} C^T \right) = I$$

Proposition

$|A| \neq 0$ if and only if A is invertible

$\Leftarrow$ prove if $|A|=0$, A is not invertible.

Proposition

for matrix,
not for the linear map

Let A be an $n \times n$ real matrix.

- ① If A has a zero column, then $|A| = 0$.
- ② If A has two equal columns, then $|A| = 0$.
- ③ If one column of A is a multiple of another, then $|A| = 0$.
- ④ $|AB| = |A||B|$
- ⑤ $|\alpha A| = \alpha^n |A|$ for $\alpha \in \mathbb{F}$ $\Rightarrow$ linearity
- ⑥ $|A^T| = |A|$

① + ⑥ $\Rightarrow$ If A has a zero row, $|A|=0$



Inner product spaces

Complex numbers

Recall that for a complex number $z = a + ib$, we define the following:

- Real part: $\operatorname{Re}(z) = a$,
- Imaginary part: $\operatorname{Im}(z) = b$,
- Complex conjugate: $\bar{z} = a - ib$,
- Modulus: $|z| = \sqrt{\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2} = \sqrt{a^2 + b^2}$

We have $|z|^2 = z\bar{z}$ and $\operatorname{Re}(z) = \frac{z + \bar{z}}{2}$.

$$\operatorname{Im}(z) = \frac{z - \bar{z}}{2i}$$

Definition

Let V be an $\mathbb{F}$ -vector space. A function $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{F}$ is called *inner product* on V if the following holds:

- ① (Conjugate) symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ for all $\mathbf{x}, \mathbf{y} \in V$, where $\bar{a}$ denotes the complex conjugate for $a \in \mathbb{C}$.
↙
If $\mathbb{F}$ is $\mathbb{R}$, then $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- ② Linearity in the first argument: $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{F}$
- ③ Positive definiteness: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$

A vector space equipped with an inner product is called an *inner product space*.

$$\begin{aligned}
 \textcircled{1} + \textcircled{2} &\Rightarrow \langle \mathbf{z}, \alpha \mathbf{x} + \beta \mathbf{y} \rangle \stackrel{\textcircled{1}}{=} \overline{\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle} \stackrel{\textcircled{2}}{=} \overline{\alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle} \\
 &= \bar{\alpha} \overline{\langle \mathbf{x}, \mathbf{z} \rangle} + \bar{\beta} \overline{\langle \mathbf{y}, \mathbf{z} \rangle}
 \end{aligned}$$

Not linear in general argument

Example

- Standard inner product on $\mathbb{R}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
- Standard inner product on $\mathbb{C}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i \bar{y}_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$
 $= \mathbf{x}^T \bar{\mathbf{y}}$
- On the space of polynomials $\mathbb{P}_n(\mathbb{R})$: $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^1 p(x)q(x)dx$ for $\mathbf{p}, \mathbf{q} \in \mathbb{P}_n(\mathbb{R})$

Proposition

Let V be an inner product space. Then $\mathbf{x} = \mathbf{0}$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in V$.

Proof. ($\Rightarrow$) Suppose $\mathbf{x} = \mathbf{0}$. Then by linearity

$$\langle \mathbf{0}, \mathbf{y} \rangle = \langle \mathbf{0} + \mathbf{0}, \mathbf{y} \rangle = \langle \mathbf{0}, \mathbf{y} \rangle + \langle \mathbf{0}, \mathbf{y} \rangle \Rightarrow \langle \mathbf{0}, \mathbf{y} \rangle = 0 \text{ for any } \mathbf{y} \in V.$$

($\Leftarrow$) Take $\mathbf{y} = \mathbf{x}$

Then $\langle \mathbf{x}, \mathbf{x} \rangle = 0$

But definition, $\mathbf{x} = \mathbf{0}$

Cauchy-Schwarz Inequality

Proposition

Let V be an inner product space. Then

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}$$

for all $\mathbf{x}, \mathbf{y} \in V$.

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle$$

Proof. Consider $\langle tx+y, tx+y \rangle$. for $t \in \mathbb{R}$

Then by positive definiteness

$$0 \leq \langle tx+y, tx+y \rangle$$

$$= \langle tx, tx \rangle + \langle y, y \rangle + \langle tx, y \rangle + \langle y, tx \rangle$$

$$= \underbrace{t \cdot \bar{t}}_{|t|^2} \langle x, x \rangle + \langle y, y \rangle + t \langle x, y \rangle + \bar{t} \langle y, x \rangle$$

$$= |t|^2 \langle x, x \rangle + t \langle x, y \rangle + \bar{t} \langle y, x \rangle + \langle y, y \rangle \quad (\text{holds for any } t)$$

$$\text{Let } t = - \frac{\overline{\langle x, y \rangle}}{\langle x, x \rangle}$$

Then we will have

$$0 \leq \frac{| \langle x, y \rangle |^2}{\langle x, x \rangle} - \frac{| \langle x, y \rangle |^2}{\langle x, x \rangle} - \frac{| \langle x, y \rangle |^2}{\langle x, x \rangle} + \langle y, y \rangle$$

$$\Rightarrow \frac{| \langle x, y \rangle |^2}{\langle x, x \rangle} \leq \langle y, y \rangle$$

$$\Rightarrow | \langle x, y \rangle | \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}$$

Proposition

Let V be an inner product space. Then $\langle \cdot, \cdot \rangle$ induces a norm on V via $\|x\| = \sqrt{\langle x, x \rangle}$ for all $x \in V$.

Proof. ① positive definiteness

$$\|x\| \geq 0 \quad \text{since } \langle x, x \rangle \geq 0$$

$$\|x\| = 0 \quad \text{iff } x = 0 \quad \Leftrightarrow \quad \langle x, x \rangle = 0 \quad \text{iff } x = 0$$

② linearity

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \bar{\alpha} \langle x, x \rangle} = |\alpha| \|x\|$$

③ Triangle inequality

$$(\|x\| + \|y\|)^2 - \|x+y\|^2$$

$$= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 - \langle x+y, x+y \rangle$$

$$= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 - \left(\underbrace{\langle x, x \rangle}_{\|x\|^2} + \underbrace{\langle x, y \rangle + \langle y, x \rangle}_{2\operatorname{Re} \langle x, y \rangle} + \underbrace{\langle y, y \rangle}_{\|y\|^2} \right)$$

$$= \underbrace{2(\|x\|\|y\| - \operatorname{Re} \langle x, y \rangle)}_{(*)} \quad (*)$$

By Cauchy-Schwarz,

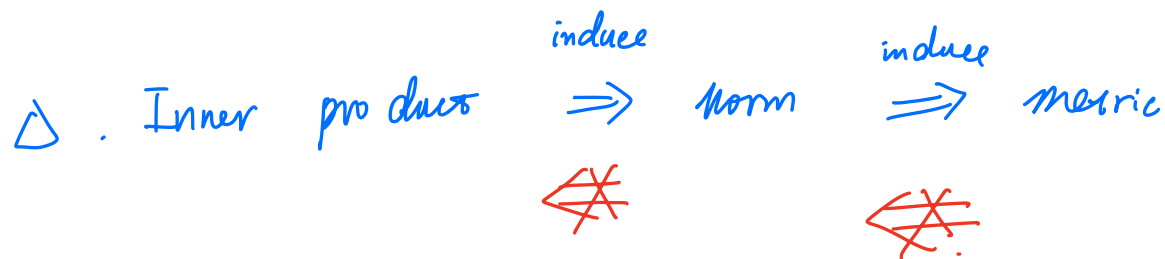
$$\operatorname{Re} \langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\|\|y\|$$

Therefore $(*) \geq 0$

Note: With this identification the Cauchy-Schwarz inequality can be restated as:
 $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$.

Example

The norm introduced by the standard inner product on $\mathbb{R}^n$ is the Euclidean distance.



Counter example:

$\mathbb{R}^n$ and ℓ_1 norm.

Define $\|(x_1, \dots, x_n)\|_1 = \sum_{i=1}^n |x_i|$

Note that if a norm is induced from an inner product, then

$$\forall x, y \in V, \quad \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

$$\|x+y\|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle$$

Let $x = (1, 0)$, $y = (0, 1)$

$$\text{then } x+y = (1, 1) \Rightarrow \|x+y\|_1 = 2$$

$$x-y = (1, -1) \Rightarrow \|x-y\|_1 = 2$$

$$\Rightarrow \|x+y\|_1^2 + \|x-y\|_1^2 = 2^2 + 2^2 = 8 \quad (\text{LHS})$$

$$2\|x\|_1^2 + 2\|y\|_1^2 = 2 + 2 = 4 \quad (\text{RHS})$$

Supplement: vector spaces and closure

A vector space starts with two closed operations

Let $\mathbb{F}$ be a field. We specify two operations on a nonempty set V :

$\mathbb{R}$ or $\mathbb{C}$

$$+ : V \times V \longrightarrow V,$$

$$(u, v) \longmapsto u + v,$$

$$\cdot : \mathbb{F} \times V \longrightarrow V,$$

$$(\alpha, v) \longmapsto \alpha v.$$

closure is encoded.

Closure is already encoded in the codomain

Because both arrows end in V ,

$$\underbrace{u, v \in V} \implies \underbrace{u + v \in V}, \quad \alpha \in \mathbb{F}, v \in V \implies \alpha v \in V.$$

Thus closure need not be repeated as two additional conditions.

Here “multiplication” means *scalar multiplication*; a general vector space does not require a product of two vectors.

The vector space: **addition**

A nonempty set V with the two operations on the previous slide is a vector space over $\mathbb{F}$ if, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,

- ① $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- ② $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
- ③ there exists $\mathbf{0} \in V$ such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$;
- ④ for every $\mathbf{v} \in V$, there exists $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

Important

The symbol $\mathbf{0}$ denotes the additive identity in V ; it need not be the real number 0.

The vector space: scalar multiplication

In addition, for all $\alpha, \beta \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in V$,

⑤ $1\mathbf{v} = \mathbf{v};$

⑥ $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v});$

⑦ $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v};$

⑧ $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}.$

Together with the declared codomains $V \times V \rightarrow V$ and $\mathbb{F} \times V \rightarrow V$, these are the usual eight vector space conditions.

Vector space versus subspace: what must be checked?

Defining a vector space structure on V

Specify operations $V \times V \rightarrow V$ and $\mathbb{F} \times V \rightarrow V$, then verify the eight conditions.
Closure is part of specifying the operations.

Showing that $W \subseteq V$ is a subspace

A lot of condition is also inherited.

The operations are inherited from V , but their outputs are initially known only to lie in V . We must check that they remain in W :

$$\boxed{0 \in W}, \quad \underline{u, v \in W \implies u + v \in W}, \quad \underline{\alpha \in \mathbb{F}, u \in W \implies \alpha u \in W}.$$

Equivalently, for nonempty W , one may use the one-step test

$$\alpha u + \beta v \in W \quad \text{for all } u, v \in W, \alpha, \beta \in \mathbb{F}.$$

All other vector space conditions are inherited from V .

References

Axler S. *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Springer, 2015. Available from:
<https://link.springer.com/book/10.1007/978-3-319-11080-6>

Treil S. *Linear Algebra Done Wrong*. 2017. Available from:
<https://www.math.brown.edu/streil/papers/LADW/LADW.html>