LINEAR ALGEBRA -II

B V Rajarama Bhat

Indian Statistical Institute, Bangalore

➤ You have studied vector spaces. Now we introduce a new structure on them called 'inner product'.

- ➤ You have studied vector spaces. Now we introduce a new structure on them called 'inner product'.
- ▶ This abstractly captures the notions of 'length' and 'angle'.

- ➤ You have studied vector spaces. Now we introduce a new structure on them called 'inner product'.
- ▶ This abstractly captures the notions of 'length' and 'angle'.
- ▶ Once we have an inner product we can talk about the distance between elements of the vector space. This allows us to define convergence of a sequence vectors.

- You have studied vector spaces. Now we introduce a new structure on them called 'inner product'.
- ▶ This abstractly captures the notions of 'length' and 'angle'.
- ▶ Once we have an inner product we can talk about the distance between elements of the vector space. This allows us to define convergence of a sequence vectors.
- ► The notion of inner product also allows us to define as to when one vector is 'orthogonal' to another.

▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .

- ▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .
- ▶ We recall that for any complex number z = a + ib with $a, b \in \mathbb{R}$, $\overline{z} := a ib$.

- ▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .
- We recall that for any complex number z = a + ib with $a, b \in \mathbb{R}$, $\overline{z} := a ib$.
- For any two complex numbers z, w we have $\overline{(zw)} = \overline{z}.\overline{w}$.

- ▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .
- We recall that for any complex number z = a + ib with $a, b \in \mathbb{R}$, $\overline{z} := a ib$.
- For any two complex numbers z, w we have $\overline{(zw)} = \overline{z}.\overline{w}$.
- ▶ For $z \in \mathbb{C}$ as above,

$$|z| := (z\overline{z})^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}}.$$

- ▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .
- We recall that for any complex number z = a + ib with $a, b \in \mathbb{R}$, $\overline{z} := a ib$.
- For any two complex numbers z, w we have $\overline{(zw)} = \overline{z}.\overline{w}$.
- ▶ For $z \in \mathbb{C}$ as above,

$$|z| := (z\overline{z})^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}}.$$

Recall that any complex number $z \neq 0$ has the unique polar decomposition as $z = re^{i\theta}$ where r = |z| and $0 \leq \theta < 2\pi$.

- ▶ In the following the field \mathbb{F} would be either \mathbb{R} or \mathbb{C} .
- We recall that for any complex number z = a + ib with $a, b \in \mathbb{R}$, $\overline{z} := a ib$.
- For any two complex numbers z, w we have $\overline{(zw)} = \overline{z}.\overline{w}$.
- ▶ For $z \in \mathbb{C}$ as above,

$$|z| := (z\overline{z})^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}}.$$

- Recall that any complex number $z \neq 0$ has the unique polar decomposition as $z = re^{i\theta}$ where r = |z| and $0 \leq \theta < 2\pi$.
- ▶ We have |z| = 0 if and only if z = 0. Further, |zw| = |z||w| and $|z + w| \le |z| + |w|$ for all $z, w \in \mathbb{C}$.

ightharpoonup Consider a vector space V over a field $\mathbb F$ (which is either $\mathbb R$ or $\mathbb C$.

- ▶ Consider a vector space V over a field \mathbb{F} (which is either \mathbb{R} or \mathbb{C} .
- An inner product between two vectors x, y in V, usually denoted by $\langle x, y \rangle$ would be a scalar (an element of \mathbb{F}).

- ▶ Consider a vector space V over a field \mathbb{F} (which is either \mathbb{R} or \mathbb{C} .
- An inner product between two vectors x, y in V, usually denoted by $\langle x, y \rangle$ would be a scalar (an element of \mathbb{F}).
- ▶ The main examples we have in mind are the following:

- ▶ Consider a vector space V over a field \mathbb{F} (which is either \mathbb{R} or \mathbb{C} .
- An inner product between two vectors x, y in V, usually denoted by $\langle x, y \rangle$ would be a scalar (an element of \mathbb{F}).
- ▶ The main examples we have in mind are the following:
- ▶ Definition 7.1: For $n \in \mathbb{N}$, consider the vector space \mathbb{R}^n . The standard inner product on \mathbb{R}^n is defined by:

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n x_j y_j.$$

- ▶ Consider a vector space V over a field \mathbb{F} (which is either \mathbb{R} or \mathbb{C} .
- An inner product between two vectors x, y in V, usually denoted by $\langle x, y \rangle$ would be a scalar (an element of \mathbb{F}).
- ▶ The main examples we have in mind are the following:
- ▶ Definition 7.1: For $n \in \mathbb{N}$, consider the vector space \mathbb{R}^n . The standard inner product on \mathbb{R}^n is defined by:

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n x_j y_j.$$

▶ We note that $\langle x, x \rangle \ge 0$ for every x in \mathbb{R}^n and $\langle x, x \rangle = 0$ if and only if x = 0.

\mathbb{C}^n . as an inner product space

▶ Definition 7.2: For $n \in \mathbb{N}$, consider the vector space \mathbb{C}^n . The standard inner product on \mathbb{C}^n is defined by:

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n \overline{x_j} y_j.$$

\mathbb{C}^n . as an inner product space

▶ Definition 7.2: For $n \in \mathbb{N}$, consider the vector space \mathbb{C}^n . The standard inner product on \mathbb{C}^n is defined by:

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n \overline{x_j} y_j.$$

▶ Here also we note that $\langle x, x \rangle \ge 0$ for every x in \mathbb{C}^n and $\langle x, x \rangle = 0$ if and only if x = 0.

▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .

- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

such that

• (i) $\langle x, cy + dz \rangle = c \langle x, y \rangle + d \langle x, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in second variable.)



- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

- (i) $\langle x, cy + dz \rangle = c \langle x, y \rangle + d \langle x, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in second variable.)
- ▶ (ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$. (Conjugate Symmetry.)

- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

- (i) $\langle x, cy + dz \rangle = c \langle x, y \rangle + d \langle x, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in second variable.)
- ▶ (ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$. (Conjugate Symmetry.)
- (iii) $\langle x, x \rangle \ge 0$ for all x in V. (Positivity.)

- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

- (i) $\langle x, cy + dz \rangle = c \langle x, y \rangle + d \langle x, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in second variable.)
- ▶ (ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$. (Conjugate Symmetry.)
- (iii) $\langle x, x \rangle \ge 0$ for all x in V. (Positivity.)
- (iv) $\langle x, x \rangle = 0$ if and only if x = 0. (Definiteness.)

- ▶ Definition 7.3: Let V be a vector space over a field \mathbb{F} where \mathbb{F} is either \mathbb{R} or \mathbb{C} .
- ► An inner product on *V* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

- (i) $\langle x, cy + dz \rangle = c \langle x, y \rangle + d \langle x, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in second variable.)
- ▶ (ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in V$. (Conjugate Symmetry.)
- (iii) $\langle x, x \rangle \ge 0$ for all x in V. (Positivity.)
- (iv) $\langle x, x \rangle = 0$ if and only if x = 0. (Definiteness.)
- Some authors take inner product as linear in the first variable. It is a matter of convention. A vector space with a specified inner product is called an inner product space.



It is clear that standard inner products on \mathbb{R}^n and \mathbb{C}^n are inner products.

- It is clear that standard inner products on \mathbb{R}^n and \mathbb{C}^n are inner products.
- **Example 7.4**: For $n \in \mathbb{N}$, fix scalars a_1, a_2, \ldots, a_n .

- ▶ It is clear that standard inner products on \mathbb{R}^n and \mathbb{C}^n are inner products.
- **Example 7.4:** For $n \in \mathbb{N}$, fix scalars a_1, a_2, \ldots, a_n .
- ► Then

$$\left\langle \left(\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right), \left(\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array}\right) \right\rangle = \sum_{j=1}^n a_j x_j y_j,$$

- ▶ It is clear that standard inner products on \mathbb{R}^n and \mathbb{C}^n are inner products.
- **Example 7.4**: For $n \in \mathbb{N}$, fix scalars a_1, a_2, \ldots, a_n .
- ► Then

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n a_j x_j y_j,$$

▶ is an inner-product on \mathbb{R}^n if and only if $a_j > 0$ for every j.

- ▶ It is clear that standard inner products on \mathbb{R}^n and \mathbb{C}^n are inner products.
- **Example 7.4**: For $n \in \mathbb{N}$, fix scalars a_1, a_2, \ldots, a_n .
- ► Then

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{j=1}^n a_j x_j y_j,$$

- ▶ is an inner-product on \mathbb{R}^n if and only if $a_i > 0$ for every j.
- Note that if $a_j \geq 0$, then conditions (i)-(iii) of the inner product are satisfied but the definiteness may not be satisfied. In such cases, $\langle \cdot, \cdot \rangle$ is known as semi-inner product.

▶ If the field $\mathbb{F} = \mathbb{R}$, from linearity and (conjugate) symmetry of the inner-product we get also linearity in the first variable.

- ▶ If the field $\mathbb{F} = \mathbb{R}$, from linearity and (conjugate) symmetry of the inner-product we get also linearity in the first variable.
- ▶ That is, $\langle cx + dy, z \rangle = c \langle x, z \rangle + d \langle y, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in first variable in the real case.)

- ▶ If the field $\mathbb{F} = \mathbb{R}$, from linearity and (conjugate) symmetry of the inner-product we get also linearity in the first variable.
- ▶ That is, $\langle cx + dy, z \rangle = c \langle x, z \rangle + d \langle y, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in first variable in the real case.)
- ▶ On the other hand if the field $\mathbb{F} = \mathbb{C}$, from linearity and conjugate symmetry of the inner-product we get anti-linearity in the first variable.

- ▶ If the field $\mathbb{F} = \mathbb{R}$, from linearity and (conjugate) symmetry of the inner-product we get also linearity in the first variable.
- ▶ That is, $\langle cx + dy, z \rangle = c \langle x, z \rangle + d \langle y, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$ (Linearity in first variable in the real case.)
- ▶ On the other hand if the field $\mathbb{F} = \mathbb{C}$, from linearity and conjugate symmetry of the inner-product we get anti-linearity in the first variable.
- ▶ That is, $\langle cx + dy, z \rangle = \bar{c}\langle x, z \rangle + \bar{d}\langle y, z \rangle$, for all $x, y, z \in V, c, d \in \mathbb{F}$.

Another example

▶ Take $\mathbb{F} = \mathbb{C}$. Fix $m, n \in \mathbb{N}$. Then we know that $M_{m,n}(\mathbb{C})$ is a vector space of dimension mn over \mathbb{C} .

Another example

- ▶ Take $\mathbb{F} = \mathbb{C}$. Fix $m, n \in \mathbb{N}$. Then we know that $M_{m,n}(\mathbb{C})$ is a vector space of dimension mn over \mathbb{C} .
- ▶ For $X, Y \in M_{m,n}(\mathbb{C})$ take

$$\langle X, Y \rangle = \operatorname{trace}(X^*Y)$$

where $(X^*)_{jk} = \overline{x_{kj}}, 1 \le k \le m; 1 \le j \le n$.

Another example

- ▶ Take $\mathbb{F} = \mathbb{C}$. Fix $m, n \in \mathbb{N}$. Then we know that $M_{m,n}(\mathbb{C})$ is a vector space of dimension mn over \mathbb{C} .
- ▶ For $X, Y \in M_{m,n}(\mathbb{C})$ take

$$\langle X, Y \rangle = \operatorname{trace}(X^*Y)$$

where $(X^*)_{jk} = \overline{x_{kj}}, 1 \le k \le m; 1 \le j \le n$.

▶ Then $\langle \cdot, \cdot \rangle$ is an inner product on $M_{m,n}(\mathbb{C})$.

Another example

- ▶ Take $\mathbb{F} = \mathbb{C}$. Fix $m, n \in \mathbb{N}$. Then we know that $M_{m,n}(\mathbb{C})$ is a vector space of dimension mn over \mathbb{C} .
- ▶ For $X, Y \in M_{m,n}(\mathbb{C})$ take

$$\langle X, Y \rangle = \operatorname{trace}(X^*Y)$$

where $(X^*)_{jk} = \overline{x_{kj}}, 1 \le k \le m; 1 \le j \le n$.

- ▶ Then $\langle \cdot, \cdot \rangle$ is an inner product on $M_{m,n}(\mathbb{C})$.
- ► Proof: We have,

$$\langle X, Y \rangle = \operatorname{trace}(X^*Y)$$

$$= \sum_{j=1}^{n} (X^*Y)_{jj}$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} (X^*)_{jk} (Y)_{kj}$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} \overline{x_{kj}} y_{kj}$$

Now it is clear that this is essentially the standard inner product on \mathbb{C}^{mn} .

- Now it is clear that this is essentially the standard inner product on \mathbb{C}^{mn} .
- Similarly,

$$\langle X, Y \rangle = \operatorname{trace}(X^t Y),$$

is an inner product on $M_{m,n}(\mathbb{R})$.

$$\mathbb{R}^2$$
.

$$(\langle x, x \rangle)^{\frac{1}{2}} = (x_1^2 + x_2^2)^{\frac{1}{2}}$$

is the distance of point x from the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$$(\langle x, x \rangle)^{\frac{1}{2}} = (x_1^2 + x_2^2)^{\frac{1}{2}}$$

is the distance of point x from the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

More generally the distance between any two points $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is

$$((x_1-y_1)^2+(x_2-y_2)^2))^{\frac{1}{2}}=(\langle x-y,x-y\rangle)^{\frac{1}{2}}.$$

$$(\langle x, x \rangle)^{\frac{1}{2}} = (x_1^2 + x_2^2)^{\frac{1}{2}}$$

is the distance of point x from the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

More generally the distance between any two points $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is

$$((x_1-y_1)^2+(x_2-y_2)^2))^{\frac{1}{2}}=(\langle x-y,x-y\rangle)^{\frac{1}{2}}.$$

This suggests the following definitions.

The norm on an inner product space

▶ Definition 7.5: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then the norm of a vector $x \in V$ is defined as

$$||x|| = (\langle x, x \rangle)^{\frac{1}{2}}.$$

The norm on an inner product space

▶ Definition 7.5: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then the norm of a vector $x \in V$ is defined as

$$||x|| = (\langle x, x \rangle)^{\frac{1}{2}}.$$

▶ If x, y are vectors in V, the distance of y from x is defined as

$$d(x,y) = \|y - x\|.$$

The norm on an inner product space

▶ Definition 7.5: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then the norm of a vector $x \in V$ is defined as

$$||x|| = (\langle x, x \rangle)^{\frac{1}{2}}.$$

ightharpoonup If x, y are vectors in V, the distance of y from x is defined as

$$d(x,y) = \|y - x\|.$$

► The 'distance function' $d: V \times V \rightarrow \mathbb{R}$ is also known as metric.



▶ Theorem 7.6: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\| \cdot \|$ be the associated norm on V. Then

- ▶ Theorem 7.6: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\| \cdot \|$ be the associated norm on V. Then
- ▶ (i) $||x|| \ge 0$ for all $x \in V$ and ||x|| = 0 if and only if x = 0.

- ▶ Theorem 7.6: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\| \cdot \|$ be the associated norm on V. Then
- \blacktriangleright (i) $||x|| \ge 0$ for all $x \in V$ and ||x|| = 0 if and only if x = 0.
- $(ii) ||ax|| = |a|||x||, \quad \forall a \in \mathbb{F} \text{ and } x \in V.$

- ▶ Theorem 7.6: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\| \cdot \|$ be the associated norm on V. Then
- ▶ (i) $||x|| \ge 0$ for all $x \in V$ and ||x|| = 0 if and only if x = 0.
- $(ii) ||ax|| = |a|||x||, \forall a \in \mathbb{F} \text{ and } x \in V.$
- $(iii) ||x + y|| \le ||x|| + ||y||, \forall x, y \in V.$

▶ Proof. (i) This is clear, as $\langle x, x \rangle \ge 0$ for all x and $\langle x, x \rangle = 0$ if and only if x = 0.

- ▶ Proof. (i) This is clear, as $\langle x, x \rangle \ge 0$ for all x and $\langle x, x \rangle = 0$ if and only if x = 0.
- ► (ii) We have,

$$\langle ax, ax \rangle = \bar{a}.a\langle x, x \rangle = |a|^2 \langle x, x \rangle$$

from anti-linearity of the inner product in the first variable and linear in the second variable. Now (ii) is immediate.

- ▶ Proof. (i) This is clear, as $\langle x, x \rangle \ge 0$ for all x and $\langle x, x \rangle = 0$ if and only if x = 0.
- ► (ii) We have,

$$\langle ax, ax \rangle = \bar{a}.a\langle x, x \rangle = |a|^2 \langle x, x \rangle$$

from anti-linearity of the inner product in the first variable and linear in the second variable. Now (ii) is immediate.

▶ (iii) This is a consequence of Cauchy-Schwarz inequality and we will prove it in the next class.

- ▶ Proof. (i) This is clear, as $\langle x, x \rangle \ge 0$ for all x and $\langle x, x \rangle = 0$ if and only if x = 0.
- ► (ii) We have,

$$\langle ax, ax \rangle = \bar{a}.a\langle x, x \rangle = |a|^2 \langle x, x \rangle$$

from anti-linearity of the inner product in the first variable and linear in the second variable. Now (ii) is immediate.

- ▶ (iii) This is a consequence of Cauchy-Schwarz inequality and we will prove it in the next class.
- ► END OF LECTURE 7.