Eigenvalue eigenvector

Instructor: Dr. Avijit Pal

Linear algebra- II (IC152)

- In $\mathbb{R}^2$, given two vectors $\mathbf{x} = (x_1, x_2),\ \mathbf{y} = (y_1, y_2)$, we know the inner product $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2$.
- Note that for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$, this inner product satisfies the conditions:
  1. $\mathbf{x} \cdot (\mathbf{y} + \alpha\mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \alpha\mathbf{x} \cdot \mathbf{z}$
  2. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
  3. $\mathbf{x} \cdot \mathbf{x} \geq 0$ and $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
- Thus, we are motivated to define an inner product on an arbitrary vector space.

Definition (Inner product)

Let $V(\mathbb{F})$ be a vector space over $\mathbb{F}$. An inner product over $V(\mathbb{F})$, denoted by $\langle\ ,\ \rangle$, is a map, $\langle\ ,\ \rangle\ : V \times V \longrightarrow \mathbb{F}$ such that for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in \mathbb{F}$

1. $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w}\rangle = a\langle \mathbf{u}, \mathbf{w}\rangle + b\langle \mathbf{v}, \mathbf{w}\rangle$, $\langle \mathbf{u}, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}\rangle}$, the complex conjugate of $\langle \mathbf{u}, \mathbf{v}\rangle$,
2. $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ for all $\mathbf{u} \in V$ and equality holds if and only if $\mathbf{u} = \mathbf{0}$.

Definition (Inner product space)

Let V be a vector space with an inner product $\langle\ ,\ \rangle$. Then $(V, \langle\ ,\ \rangle)$ is called an inner product space, in short denoted by IPS.

Example 1

- Let $V = \mathbb{R}^n$ be the real vector space of dimension n. Given two vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ of V, we define

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \mathbf{u}\mathbf{v}^t.$$

- We will show that $\langle \ , \ \rangle$ is an inner product.
- One can easily check from definition that $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
- $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ for all $\mathbf{u} \in V$, because $\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{i=1}^{n} u_i^2$.
- Also, $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \sum_{i=1}^{n} u_i^2 = 0$ which is equivalent to $\mathbf{u} = 0$.
- Note that for any $\mathbf{w} = (w_1, \ldots, w_n)$ of V, and $\alpha, \beta \in \mathbb{R}$, we have

$$\begin{aligned} \langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle &= \alpha \sum_{i=1}^{n} u_i w_i + \beta \sum_{i=1}^{n} v_i w_i \\ &= \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle. \end{aligned}$$

Example 2

- Let $V = \mathbb{C}^n$ be a complex vector space of dimension n. Then for $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in V, we define

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1\overline{v_1} + u_2\overline{v_2} + \cdots + u_n\overline{v_n} = \mathbf{u}\mathbf{v}^*$$

- Check that $\langle \; , \; \rangle$ is an inner product.

Remark

Note that in parts 1 *and* 2 *of above Example , the inner products are* $\mathbf{u}\mathbf{v}^t$ *and* $\mathbf{u}\mathbf{v}^*$ *, respectively. This occurs because the vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are row vectors. In general,* $\mathbf{u}$ *and* $\mathbf{v}$ *are taken as column vectors and hence one uses the notation* $\mathbf{u}^t\mathbf{v}$ *or* $\mathbf{u}^*\mathbf{v}$ *.*

Example 3

- Let $V = \mathbb{R}^2$ and let $A = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$.
- Define
$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}A\mathbf{y}^t = 4x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2.$$
- Check that $\langle \ , \ \rangle$ is an inner product.
- Observe that the matrix A is real symmetric matrix.
- xAx^t is a quadratic form of a real symmetric matrix A which we will discuss later.

Example 4

- Consider the set $M_{n\times n}(\mathbb{R})$ of all real square matrices of order n. For $A, B \in M_{n\times n}(\mathbb{R})$ we define $\langle A, B\rangle = tr(AB^t)$.
- Then

$$\langle A+B, C\rangle = tr\big((A+B)C^t\big) = tr(AC^t) + tr(BC^t) = \langle A, C\rangle + \langle B, C\rangle.$$

- $\langle A, B\rangle = tr(AB^t) = tr(\ (AB^t)^t\ ) = tr(BA^t) = \langle B, A\rangle.$
- Let $A = (a_{ij})$. Then

$$\langle A, A\rangle = tr(AA^t) = \sum_{i=1}^{n}(AA^t)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2$$

  and therefore, $\langle A, A\rangle > 0$ for all non-zero matrices A.
- So, it is clear that $\langle A, B\rangle$ is an inner product on $M_{n\times n}(\mathbb{R})$.

Definition (Length/Norm of a Vector)

For $\mathbf{u} \in V$, we define the length (norm) of $\mathbf{u}$, denoted $\|\mathbf{u}\|$, by $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$, the positive square root.

- A very useful and a fundamental inequality concerning the inner product is due to Cauchy and Schwartz. The next theorem gives the statement and a proof of this inequality.

Theorem (Cauchy-Schwartz inequality)

Let $V(\mathbb{F})$ be an inner product space. Then for any $\mathbf{u}, \mathbf{v} \in V$

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|.$$

The equality holds if and only if the vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. Further, if $\mathbf{u} \neq \mathbf{0}$, then

$$\mathbf{v} = \langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|} \rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Outline of the proof

- If $\mathbf{u} = \mathbf{0}$, then the inequality holds.
- Let $\mathbf{u} \neq \mathbf{0}$. Note that $\langle \lambda\mathbf{u} + \mathbf{v}, \lambda\mathbf{u} + \mathbf{v} \rangle \geq 0$ for all $\lambda \in \mathbb{F}$.
- In particular, for $\lambda = -\dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2}$, we get

$$\begin{aligned}
0 &\leq \langle \lambda\mathbf{u} + \mathbf{v}, \lambda\mathbf{u} + \mathbf{v} \rangle \\
&= \lambda\bar{\lambda}\|\mathbf{u}\|^2 + \lambda\langle \mathbf{u}, \mathbf{v} \rangle + \bar{\lambda}\langle \mathbf{v}, \mathbf{u} \rangle + \|\mathbf{v}\|^2 \\
&= \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \frac{\overline{\langle \mathbf{v}, \mathbf{u} \rangle}}{\|\mathbf{u}\|^2} \|\mathbf{u}\|^2 - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \langle \mathbf{u}, \mathbf{v} \rangle - \frac{\overline{\langle \mathbf{v}, \mathbf{u} \rangle}}{\|\mathbf{u}\|^2} \langle \mathbf{v}, \mathbf{u} \rangle + \|\mathbf{v}\|^2
\end{aligned}$$

- Or, in other words

$$|\langle \mathbf{v}, \mathbf{u} \rangle|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2$$

and the proof of the inequality is over.

- Observe that if $\mathbf{u} \neq \mathbf{0}$ then the equality holds if and only of $\lambda \mathbf{u} + \mathbf{v} = \mathbf{0}$ for $\lambda = -\frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2}$.
- That is, $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.
- We leave it for the reader to prove

$$\mathbf{v} = \langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|} \rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Definition (Angle between two vectors)

Let V be a real vector space. Then for every $\mathbf{u}, \mathbf{v} \in V$, by the Cauchy-Schwartz inequality, we have

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \ \|\mathbf{v}\|} \leq 1.$$

Remark

We know that $\cos : [0, \pi] \longrightarrow [-1,\ 1]$ *is an one-one and onto function. Therefore, for every real number* $\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \ \|\mathbf{v}\|}$, *there exists a unique* $\theta,\ 0 \leq \theta \leq \pi$, *such that*

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \ \|\mathbf{v}\|}.$$

Observations

- The real number θ with $0 \leq \theta \leq \pi$ and satisfying

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}$$

  is called the angle between the two vectors $\mathbf{u}$ and $\mathbf{v}$ in V.
- The vectors $\mathbf{u}$ and $\mathbf{v}$ in V are said to be orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is called mutually orthogonal if $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $1 \leq i \neq j \leq n$.

Definition (Orthogonal Complement)

Let W be a subspace of a vector space V with inner product $\langle \, , \, \rangle$. Then the subspace

$$W^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}$$

is called the orthogonal complement of W in V.

Theorem

Let V be an inner product space. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a set of non-zero, mutually orthogonal vectors of V.

- *Then the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is linearly independent.*
- $\| \sum_{i=1}^{n} \alpha_i \mathbf{u}_i \|^2 = \sum_{i=1}^{n} |\alpha_i|^2 \|\mathbf{u}_i\|^2;$
- *Let $\dim(V) = n$ and also let $\|\mathbf{u}_i\| = 1$ for $i = 1, 2, \ldots, n$. Then for any $\mathbf{v} \in V$,*

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i.$$

In particular, $\langle \mathbf{v}, \mathbf{u}_i \rangle = 0$ for all $i = 1, 2, \ldots, n$ if and only if $\mathbf{v} = \mathbf{0}$.

Outline of proof

- Consider the set of non-zero, mutually orthogonal vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$. Suppose there exist scalars $c_1, c_2, \ldots, c_n$ not all zero, such that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n = \mathbf{0}.$$

- Then for $1 \leq i \leq n$, we have

$$0 = \langle \mathbf{0}, \mathbf{u}_i \rangle = \langle \sum_{i=1}^{n} c_i u_i, u_i \rangle = \sum_{j=1}^{n} c_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = c_i$$

  as $\langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$ for all $j \neq i$ and $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$.
- This gives a contradiction to our assumption that some of the c_i 's are non-zero.
- This establishes the linear independence of a set of non-zero, mutually orthogonal vectors.

- For the second part, using $= \langle \mathbf{u}_i, \mathbf{u}_j \rangle \begin{cases} 0 & \text{if } i \neq j \\ \|\mathbf{u}_i\|^2 & \text{if } i = j \end{cases}$ for $1 \leq i, j \leq n$, we have

$$\begin{aligned} \|\sum_{i=1}^{n} \alpha_i \mathbf{u}_i|^2 &= \langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \sum_{i=1}^{n} \alpha_i \mathbf{u}_i \rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle \mathbf{u}_i, \sum_{j=1}^{n} \alpha_j \mathbf{u}_j \rangle \\ &= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \overline{\alpha_j} \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\ &= \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} \langle \mathbf{u}_i, \mathbf{u}_i \rangle \\ &= \sum_{i=1}^{n} |\alpha_i|^2 \|\mathbf{u}_i\|^2 \end{aligned}$$

- For the third part, observe from the first part, the linear independence of the non-zero mutually orthogonal vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$.
- Since $\dim(V) = n$, they form a basis of V. Thus, for every vector $\mathbf{v} \in V$, there exist scalars $\alpha_i,\ 1 \leq i \leq n$, such that $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_n$.
- Hence,

$$\langle \mathbf{v}, \mathbf{u}_j \rangle = \langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \alpha_j.$$

- Therefore, we have obtained the required result.

Definition (Orthonormal Set)

Let V be an inner product space. A set of non-zero, mutually orthogonal vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in V is called an orthonormal set if $\|\mathbf{v}_i\| = 1$ for $i = 1, 2, \ldots, n$.
If the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is also a basis of V, then the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is called an orthonormal basis of V.

- In view of Theorem (7), we inquire into the question of extracting an orthonormal basis from a given basis.
- In the next lecture, we describe a process (called the Gram-Schmidt Orthogonalisation process) that generates an orthonormal set from a given set containing finitely many vectors.

- Consider the vector space $\mathbb{R}^2$ with the standard inner product. Then the standard ordered basis

$$\mathcal{B} = \big((1,0),(0,1)\big)$$

is an orthonormal set. Also, the basis

$$\mathcal{B}_1 = \big(\frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1)\big)$$

is an orthonormal set.

- Let $\mathbb{R}^n$ be endowed with the standard inner product. Then check that the standard ordered basis

$$(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$$

is an orthonormal set.

Remark

The last part of the above theorem can be rephrased as "suppose $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of an inner product space V. Then for each $\mathbf{u} \in V$ the numbers $\langle \mathbf{u}, \mathbf{v}_i \rangle$ for $1 \leq i \leq n$ are the coordinates of $\mathbf{u}$ with respect to the above basis".
That is, let $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ be an ordered basis. Then for any $\mathbf{u} \in V$,

$$[\mathbf{u}]_{\mathcal{B}} = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{u}, \mathbf{v}_n \rangle)^t.$$

Thank You