Slide 1

Inner product

- Review: Definition of inner product.
- Norm and distance.
- Orthogonal vectors.
- Orthogonal complement.
- Orthogonal basis.

Slide 2

Definition of inner product

Definition 1 (Inner product) *Let V be a vector space over $\mathbb{R}$. An inner product $(\ ,\ )$ is a function $V \times V \to \mathbb{R}$ with the following properties*

1. $\forall\, \mathbf{u} \in V$, $(\mathbf{u}, \mathbf{u}) \geq 0$, *and* $(\mathbf{u}, \mathbf{u}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$;
2. $\forall\ \mathbf{u},\ \mathbf{v} \in V$, *holds* $(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})$;
3. $\forall\ \mathbf{u},\ \mathbf{v},\ \mathbf{w} \in V$, *and* $\forall\ a,\ b \in \mathbb{R}$ *holds*
   $(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = a(\mathbf{u}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w})$.

Notation: V together with $(\ ,\ )$ is called an inner product space.

Slide 3

Examples

- The Euclidean inner product in $\mathbb{R}^2$. Let $V = \mathbb{R}^2$, and $\{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis. Given two arbitrary vectors $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$, then

$$(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2.$$

Notice that $(\mathbf{e}_1, \mathbf{e}_1) = 1$, $(\mathbf{e}_2, \mathbf{e}_2) = 1$, and $(\mathbf{e}_1, \mathbf{e}_2) = 0$. It is also called "dot product", and denoted as $\mathbf{x} \cdot \mathbf{y}$.

- The Euclidean inner product in $\mathbb{R}^n$. Let $V = \mathbb{R}^n$, and $\{\mathbf{e}_i\}_{i=1}^n$ be the standard basis. Given two arbitrary vectors $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$ and $\mathbf{y} = \sum_{i=1}^n y_i\mathbf{e}_i$, then

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i.$$

Notice that $(\mathbf{e}_i, \mathbf{e}_j) = I_{ij}$

Slide 4

Examples

- An inner product in the vector space of continuous functions in $[0, 1]$, denoted as $V = C([0, 1])$, is defined as follows. Given two arbitrary vectors $f(x)$ and $g(x)$, introduce the inner product

$$(f, g) = \int_0^1 f(x)g(x)\,dx.$$

- An inner product in the vector space of functions with one continuous first derivative in $[0, 1]$, denoted as $V = C^1([0, 1])$, is defined as follows. Given two arbitrary vectors $f(x)$ and $g(x)$, then

$$(f, g) = \int_0^1 \left[f(x)g(x) + f'(x)g'(x)\right]\,dx.$$

Slide 5

Norm

An inner product space induces a norm, that is, a notion of length of a vector.

Definition 2 (Norm) *Let V, $(\ , \ )$ be a inner product space. The norm function, or length, is a function $V \to \mathbb{R}$ denoted as $\| \ \|$, and defined as*

$$\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})}.$$

Example:

- The Euclidean norm in $\mathbb{R}^2$ is given by

$$\|\mathbf{u}\| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{(x_1)^2 + (x_2)^2}.$$

Slide 6

Examples

- The Euclidean norm in $\mathbb{R}^n$ is given by

$$\|\mathbf{u}\| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{(x_1)^2 + \cdots + (x_n)^2}.$$

- A norm in the space of continuous functions $V = C([0, 1])$ is given by

$$\|f\| = \sqrt{(f, f)} = \sqrt{\int_0^1 [f(x)]^2 dx}.$$

For example, one can check that the length of $f(x) = \sqrt{3}x$ is 1.

Slide 7

Distance

A norm in a vector space, in turns, induces a notion of distance between two vectors, defined as the length of their difference.

Definition 3 (Distance) *Let V, $(\ ,\ )$ be a inner product space, and $\|\ \|$ be its associated norm. The distance between $\mathbf{u}$ and $\mathbf{v} \in V$ is given by*

$$dist(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Example:

- The Euclidean distance between to points $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^3$ is

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

Slide 8

Orthogonal vectors

Theorem 1 *Let V be a vector space and $\mathbf{u}$, $\mathbf{v} \in V$. Then,*

$$\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\| \quad \Leftrightarrow \quad (\mathbf{u}, \mathbf{v}) = 0.$$

Proof:

$$\|\mathbf{u} + \mathbf{v}\|^2 = (\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2(\mathbf{u}, \mathbf{v}).$$

$$\|\mathbf{u} - \mathbf{v}\|^2 = (\mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2(\mathbf{u}, \mathbf{v}).$$

then,

$$\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2 = 4(\mathbf{u}, \mathbf{v}).$$

□

Slide 9

Orthogonal vectors

Definition 4 (Orthogonal vectors) *Let V, $(\ ,\ )$ be an inner product space. Two vectors* $\mathbf{u}$, $\mathbf{v} \in V$ *are orthogonal, or perpendicular, if and only if*

$$(\mathbf{u}, \mathbf{v}) = 0.$$

We call them orthogonal, because the diagonal of the parallelogram formed by $\mathbf{u}$ and $\mathbf{v}$ have the same length.

Theorem 2 *Let V be a vector space and* $\mathbf{u}$, $\mathbf{v} \in V$ *be orthogonal vectors. Then*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Slide 10

Example

- Vectors $\mathbf{u} = [1, 2]^T$ and $\mathbf{v} = [2, -1]^T$ in $\mathbb{R}^2$ are orthogonal with the inner product $(\mathbf{u}, \mathbf{v}) = u_1 v_1 + u_2 v_2$, because,

$$(\mathbf{u}, \mathbf{v}) = 2 - 2 = 0.$$

- The vectors $\cos(x)$, $\sin(x) \in C([0, 2\pi])$ are orthogonal, with the inner product $(f, g) = \int_0^{2\pi} f g \, dx$, because

$$(\cos(x), \sin(x)) = \int_0^{2\pi} \sin(x)\cos(x)\, dx = \frac{1}{2}\int_0^{2\pi} \sin(2x)\, dx,$$

$$(\cos(x), \sin(x)) = -\frac{1}{4}\left(\cos(2x)|_0^{2\pi}\right) = 0.$$

Slide 11

Orthogonal vectors

- Review: Orthogonal vectors.
- Orthogonal projection along a vector.
- Orthogonal bases.
- Orthogonal projection onto a subspace.
- Gram-Schmidt orthogonalization process.

Slide 12

Review or orthogonal vectors

Definition 5 (Orthogonal vectors) *Let V, $(\ ,\ )$ be an inner product vector space. Two vectors $\mathbf{u}$, $\mathbf{v} \in V$ are orthogonal, or perpendicular, if and only if*

$$(\mathbf{u}, \mathbf{v}) = 0.$$

Theorem 3 *Let V, $(\ ,\ )$ be an inner product vector space.*

$$\begin{aligned} \mathbf{u}, \mathbf{v} \in V \text{ are orthogonal} \quad &\Leftrightarrow \quad \|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|, \\ &\Leftrightarrow \quad \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2. \end{aligned}$$

Slide 13

Orthogonal projection along a vector

- Fix $V, (\ ,\ )$, and $\mathbf{u} \in V$, with $\mathbf{u} \neq \mathbf{0}$.
  Can any vector $\mathbf{x} \in V$ be decomposed in orthogonal parts with respect to $\mathbf{u}$, that is, $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{x}'$ with $(\hat{\mathbf{x}}, \mathbf{x}') = 0$ and $\hat{\mathbf{x}} = c\mathbf{u}$?
  Is this decomposition unique?

Theorem 4 (Orthogonal decomposition along a vector) *$V, (\ ,\ )$, an inner product vector space, and $\mathbf{u} \in V$, with $\mathbf{u} \neq \mathbf{0}$. Then, any vector $\mathbf{x} \in V$ can be uniquely decomposed as*

$$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{x}',$$

where

$$\hat{\mathbf{x}} = \frac{(\mathbf{x}, \mathbf{u})}{\|\mathbf{u}\|^2}\,\mathbf{u}, \quad \mathbf{x}' = \mathbf{x} - \hat{\mathbf{x}}.$$

Therefore, $\hat{\mathbf{x}}$ is proportional to $\mathbf{u}$, and $(\hat{\mathbf{x}}, \mathbf{x}') = 0$.

Slide 14

Orthogonal projection along a vector

Proof: Introduce $\hat{\mathbf{x}} = c\mathbf{u}$, and then write $\mathbf{x} = c\mathbf{u} + \mathbf{x}'$. The condition $(\hat{\mathbf{x}}, \mathbf{x}') = 0$ implies that $(\mathbf{u}, \mathbf{x}') = 0$, then

$$(\mathbf{x}, \mathbf{u}) = c(\mathbf{u}, \mathbf{u}), \quad \Rightarrow \quad c = \frac{(\mathbf{x}, \mathbf{u})}{\|\mathbf{u}\|^2},$$

then

$$\hat{\mathbf{x}} = \frac{(\mathbf{x}, \mathbf{u})}{\|\mathbf{u}\|^2}\,\mathbf{u}, \quad \mathbf{x}' = \mathbf{x} - \hat{\mathbf{x}}.$$

This decomposition is unique, because, given a second decomposition $\mathbf{x} = \hat{\mathbf{y}} + \mathbf{y}'$ with $\mathbf{y} = d\mathbf{u}$, and $(\hat{\mathbf{y}}, \mathbf{y}') = 0$, then, $(\mathbf{u}, \mathbf{y}') = 0$ and

$$c\mathbf{u} + \mathbf{x}' = d\mathbf{u} + \mathbf{y}' \quad \Rightarrow \quad c = d,$$

from a multiplication by $\mathbf{u}$, and then,

$$\mathbf{x}' = \mathbf{y}'.$$

□

Slide 15

Orthogonal bases

Definition 6 *Let $V, (\ ,\ )$ be an n dimensional inner product vector space, and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be a basis of V.*

$$\begin{aligned} \text{The basis is orthogonal} &\Leftrightarrow (\mathbf{u}_i, \mathbf{u}_j) = 0, \text{ for all } i \neq j. \\ \text{The basis is orthonormal} &\Leftrightarrow \text{it is orthogonal, and in addition,} \\ & \quad \|\mathbf{u}_i\| = 1, \text{ for all } i, \end{aligned}$$

where i, $j = 1, \cdots, n$.

Slide 16

Orthogonal bases

Theorem 5 *Let $V, (\ ,\ )$ be an n dimensional inner product vector space, and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be an orthogonal basis. Then, any $\mathbf{x} \in V$ can be written as*

$$\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n,$$

with the coefficients have the form

$$c_i = \frac{(\mathbf{x}, \mathbf{u}_i)}{\|\mathbf{u}_i\|^2}, \quad i = 1, \cdots, n.$$

Proof: The set $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ is a basis, so there exist coefficients c_i such that $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$. The basis is orthogonal, so multiplying the expression of $\mathbf{x}$ by $\mathbf{u}_i$, and recalling $(\mathbf{u}_i, \mathbf{u}_j) = 0$ for all $i \neq j$, one gets,

$$(\mathbf{x}, \mathbf{u}_i) = c_i(\mathbf{u}_i, \mathbf{u}_i).$$

The $\mathbf{u}_i$ are nonzero, so $(\mathbf{u}_i, \mathbf{u}_i) = \|\mathbf{u}_i\|^2 \neq 0$, so $c_i = (\mathbf{x}, \mathbf{u}_i)/\|\mathbf{u}_i\|^2$. □

Slide 17

Orthogonal projections onto subspaces

Notice:

To write $\mathbf{x}$ in an orthogonal basis means to do an orthogonal decomposition of $\mathbf{x}$ along each basis vector.

(All this holds for vector spaces of functions.)

Theorem 6 *Let $V, (\ ,\ )$ be an n dimensional inner product vector space, and $W \subset V$ be a p dimensional subspace. Let $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ be an orthogonal basis of W.*

Then, any $\mathbf{x} \in V$ can be decomposed as

$$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{x}'$$

with $(\hat{\mathbf{x}}, \mathbf{x}') = 0$ and $\hat{\mathbf{x}} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$, where the coefficients c_i are given by

$$c_i = \frac{(\mathbf{x}, \mathbf{u}_i)}{\|\mathbf{u}_i\|^2}, \quad i = 1, \cdots, p.$$

Slide 18

Gram-Schmidt Orthogonalization process

Orthogonal bases are convenient to carry out computations. Jorgen Gram and Erhard Schmidt by the year 1900 made standard a process to compute an orthogonal basis from an arbitrary basis.

(They actually needed it for vector spaces of functions. Laplace, by 1800, used this process on $\mathbb{R}^n$.)

Slide 19

Gram-Schmidt Orthogonalization process

Theorem 7 *Let $V, (\ ,\ )$ be an inner product vector space, and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be an arbitrary basis of V. Then, an orthogonal basis of V is given by the vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, where*

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1, \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{(\mathbf{u}_2, \mathbf{v}_1)}{\|\mathbf{v}_1\|^2}\,\mathbf{v}_1, \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{(\mathbf{u}_3, \mathbf{v}_1)}{\|\mathbf{v}_1\|^2}\,\mathbf{v}_1 - \frac{(\mathbf{u}_3, \mathbf{v}_2)}{\|\mathbf{v}_2\|^2}\,\mathbf{v}_2, \\
\vdots \quad & \quad \vdots \\
\mathbf{v}_n &= \mathbf{u}_n - \sum_{i=1}^{n-1} \frac{(\mathbf{u}_n, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2}\,\mathbf{v}_i.
\end{aligned}$$

Slide 20

Least-squares approximation

- Review: Gram-Schmidt orthogonalization process.
- Least-squares approximation.
  - Definition.
  - Normal equation.
  - Examples.

Slide 21

Gram-Schmidt Orthogonalization process

Theorem 8 *Let $V, (\ ,\ )$ be an inner product vector space, and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be an arbitrary basis of V. Then, an orthogonal basis of V is given by the vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, where*

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1, \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{(\mathbf{u}_2, \mathbf{v}_1)}{\|\mathbf{v}_1\|^2}\,\mathbf{v}_1, \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{(\mathbf{u}_3, \mathbf{v}_1)}{\|\mathbf{v}_1\|^2}\,\mathbf{v}_1 - \frac{(\mathbf{u}_3, \mathbf{v}_2)}{\|\mathbf{v}_2\|^2}\,\mathbf{v}_2, \\
\vdots\ & \quad \vdots \\
\mathbf{v}_n &= \mathbf{u}_n - \sum_{i=1}^{n-1} \frac{(\mathbf{u}_n, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2}\,\mathbf{v}_i.
\end{aligned}$$

Slide 22

Least-squares approximation

Let V, W be vector spaces and $A : V \to W$ be linear. Given $\mathbf{b} \in W$ then the linear equation $A\mathbf{x} = \mathbf{b}$ either has a solution $\mathbf{x}$ or it has no solutions.

Suppose now that there is an inner product in W, say $(\ ,\ )_W$, with associated norm $\|\ \|_W$. Then there exists a notion of approximate solution, given as follows.

Slide 23

Least-squares approximation

Definition 7 (Approximate solution) *Let V, W be vector spaces and let $(\ ,\ )_W$, $\|\ \|_W$ be an inner product and its associate norm in W. Let $A : V \to W$ be linear, and $\mathbf{b} \in W$ be an arbitrary vector. An approximate solution to the linear equation*

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

is a vector $\hat{\mathbf{x}} \in V$ such that

$$\|A\hat{\mathbf{x}} - \mathbf{b}\|_W \leq \|A\mathbf{x} - \mathbf{b}\|_W, \quad \forall, \mathbf{x} \in V.$$

Remark: $\hat{\mathbf{x}}$ is also called least-squares approximation, because $\hat{\mathbf{x}}$ makes the number

$$\|A\hat{\mathbf{x}} - \mathbf{b}\|_W^2 = [(A\mathbf{x})_1 - b_1]^2 + \cdots + [(A\mathbf{x})_m - b_m]^2$$

as small as possible, where $m = \dim(W)$.

Slide 24

Least-squares approximation

Theorem 9 *Let V, W be vector spaces and let $(\ ,\ )_W$, $\|\ \|_W$ be an inner product and its associate norm in W. Let $A : V \to W$ be linear, and $\mathbf{b} \in W$ be an arbitrary vector.*

If $\hat{\mathbf{x}} \in V$ satisfies that

$$(A\hat{\mathbf{x}} - \mathbf{b}) \perp \mathit{Range}(A),$$

then $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

Slide 25

Least-squares approximation

Proof: The hypothesis $(A\hat{\mathbf{x}} - \mathbf{b}) \perp \text{Range}(A)$ implies that for all $\mathbf{x} \in V$ holds

$$\begin{aligned} A\mathbf{x} - \mathbf{b} &= A\hat{\mathbf{x}} - \mathbf{b} + A\mathbf{x} - A\hat{\mathbf{x}}, \\ &= (A\hat{\mathbf{x}} - \mathbf{b}) + A(\mathbf{x} - \hat{\mathbf{x}}). \end{aligned}$$

The two terms on the right hand side are orthogonal, by hypothesis, then Pythagoras theorem holds, so

$$\|A\mathbf{x} - \mathbf{b}\|_W^2 = \|A\hat{\mathbf{x}} - \mathbf{b}\|_W^2 + \|A(\mathbf{x} - \hat{\mathbf{x}})\|_W^2 \geq \|A\hat{\mathbf{x}} - \mathbf{b}\|_W^2,$$

so $\hat{\mathbf{x}}$ is a least-squares solution. □

Slide 26

Least-squares approximation

Theorem 10 *Let $I\!R^n$, $(\ ,\ )_n$, and $I\!R^m$, $(\ ,\ )_m$ be the Euclidean inner product spaces and $A : I\!R^n \to I\!R^m$ be linear, identified with an $m \times n$ matrix. Fix $\mathbf{b} \in I\!R^m$. Then,*

$$\hat{\mathbf{x}} \in I\!R^n \ \textit{ is solution of } A^T A\hat{\mathbf{x}} = A^T\mathbf{b} \Leftrightarrow (A\hat{\mathbf{x}} - \mathbf{b}) \perp \ \textit{Col}(A).$$

Proof: Let $\hat{\mathbf{x}}$ such that $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$. Then,

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}, \Leftrightarrow A^T(A\hat{\mathbf{x}} - \mathbf{b}) = 0, \Leftrightarrow (\mathbf{a}_i, (A\hat{\mathbf{x}} - \mathbf{b}))_m = 0,$$

for all $\mathbf{a}_i$ column vector of A, where we used the notation $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$. Therefore, the condition $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ is equivalent to $(A\hat{\mathbf{x}} - \mathbf{b}) \perp \text{Col}(A)$. □

Slide 27

Least-squares approximation

The previous two results can be summarized in the following one:

Theorem 11 *Let $\mathbb{R}^n$, $(\ ,\ )_n$, and $\mathbb{R}^m$, $(\ ,\ )_m$ be the Euclidean inner product spaces and $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear, identified with an $m \times n$ matrix. Fix $\mathbf{b} \in \mathbb{R}^m$.*

If $\hat{\mathbf{x}} \in \mathbb{R}^n$ is solution of $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$, then $\hat{\mathbf{x}}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$.