SC223 - Linear Algebra

Aditya Tatu

Lecture 38

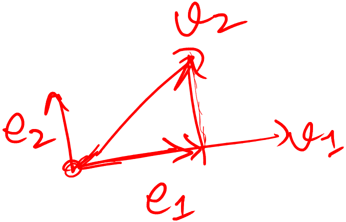
November 21, 2023

Summary of Lecture 38

- **Proposition 23:** In a IPS $(V, \langle\cdot,\cdot\rangle)$, a set of n non-zero orthogonal vectors $\{v_1, \ldots, v_n\}$ is linearly independent.
- **Proposition 24:** (Pythagoras Theorem): In an IPS $(V, \langle\cdot,\cdot\rangle)$, if $\langle x, y\rangle = 0$, $||x+y||^2 = ||x||^2 + ||y||^2$.
- **Orthogonal Decomposition:** Let $x, y \neq \theta \in V . x = \frac{\langle x,y\rangle}{\langle y,y\rangle} \cdot y + \underbrace{\left(x - \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\right)}_{w}$, such that $w \perp y$.
- **Proposition 25** (Cauchy-Schwartz inequality): In an IPS $(V, \langle\cdot,\cdot\rangle)$, $|\langle x, y\rangle| \leq ||x||\ ||y||$.

Gram-Schmidt Procedure

- **Proposition 26:** (Gram-Schmidt Procedure): Let $\{v_1, \ldots, v_m\}$ be a list of linearly independent vectors. Then there exists a list of orthonormal vectors $\{e_1, \ldots, e_m\}$ such that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\}), \forall j = 1, \ldots, m$.
- Let $e_1 = \frac{v_1}{||v_1||}$. Define $e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||v_2 - \langle v_2, e_1 \rangle e_1||}$.
- Similarly, $e_3 = \frac{v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)}{||v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)||}$, and $e_k = \frac{v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i}{||v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i||}$.
- Observe that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\})$.



Gram-Schmidt Procedure

- **Proposition 26:** (Gram-Schmidt Procedure): Let $\{v_1, \ldots, v_m\}$ be a list of linearly independent vectors. Then there exists a list of orthonormal vectors $\{e_1, \ldots, e_m\}$ such that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\}), \forall j = 1, \ldots, m$.
- Let $e_1 = \frac{v_1}{||v_1||}$. Define $e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||v_2 - \langle v_2, e_1 \rangle e_1||}$.
- Similarly, $e_3 = \frac{v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)}{||v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)||}$, and $e_k = \frac{v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i}{||v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i||}$.
- Observe that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\})$.
- It is easy to see that $e_1 \perp e_2$. Assume that $\{e_1, \ldots, e_j\}$ are orthonormal.

$$\langle e_1, e_2 \rangle = \left\langle \frac{v_1}{||v_1||}, \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||\quad \cdot\cdot \quad||} \right\rangle = 0$$

$$= \left\langle e_1, \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||\quad \cdot \quad||} \right\rangle$$

Gram-Schmidt Procedure

$w = a_1 v_1 + \cdots + a_m v_m$

- **Proposition 26:** (Gram-Schmidt Procedure): Let $\{v_1, \ldots, v_m\}$ be a list of linearly independent vectors. Then there exists a list of orthonormal vectors $\{e_1, \ldots, e_m\}$ such that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\}), \forall j = 1, \ldots, m$.
- Let $e_1 = \frac{v_1}{||v_1||}$. Define $e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{||v_2 - \langle v_2, e_1 \rangle e_1||}$.
- Similarly, $e_3 = \frac{v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)}{||v_3 - (\langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2)||}$, and $e_k = \frac{v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i}{||v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i||}$.
- Observe that $span(\{v_1, \ldots, v_j\}) = span(\{e_1, \ldots, e_j\})$.
- It is easy to see that $e_1 \perp e_2$. Assume that $\{e_1, \ldots, e_j\}$ are orthonormal.
- Then $\forall l = 1, \ldots j$, with $\tilde{e_{j+1}} = v_{j+1} - \sum_{i=1}^{j} \langle v_{j+1}, e_i \rangle e_i$

$$\langle e_{j+1}, e_l \rangle = \frac{1}{||\tilde{e_{j+1}}||} \left(\langle v_{j+1}, e_l \rangle - \sum_{i=1}^{j} \langle v_{j+1}, e_i \rangle \langle e_i, e_l \rangle \right)$$
$$= \frac{1}{||\tilde{e_{j+1}}||} \left(\langle v_{j+1}, e_l \rangle - \langle v_{j+1}, e_l \rangle \right) = 0$$

$$e_{j+1} = \frac{\tilde{e}_{j+1}}{||\tilde{e}_{j+1}||}$$

$$\langle v_{j+1} - \sum \underline{\langle v_{j+1}, e_i \rangle} e_i, e_l \rangle$$

$$w = \sum a_i v_i$$

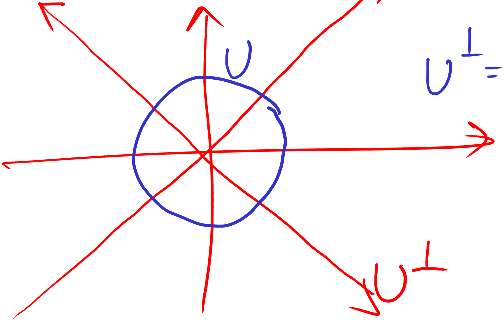
$$\langle w, v_k \rangle = \sum a_i \langle v_i, v_k \rangle.$$

Orthogonal Complement

- Let V be a FD IPS and let U be a subset of V. The **Orthogonal Complement** of U is defined as

$$U^{\perp} = \{v \in V \mid \underbrace{\langle v, u \rangle = 0, \forall u \in U\}}_{U}$$

$V = \mathbb{R}^2$



$U^{\perp} = \{(0,0)\}.$

Orthogonal Complement

- Let V be a FD IPS and let U be a subset of V. The **Orthogonal Complement** of U is defined as

$$U^{\perp} = \{v \in V \mid \langle v, u \rangle = 0, \forall u \in U\}$$

- **Proposition 27:** Irrespective of whether U is a subspace of V or not, $U^{\perp}$ is a subspace.

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.

$w \in U^{\perp} \Rightarrow \langle w, u \rangle = 0, \ \forall u \in U.$

$w \in U \Rightarrow \langle w, w \rangle = 0 = \|w\|^2 \Rightarrow w = \theta.$

$V = U + U^{\perp}$

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.

$P_U \leftarrow V = U \oplus W$

$P_U \leftarrow V = U \oplus U^{\perp}$, $U^{\perp} \neq W$.

$P_U \rightarrow$ O.P.

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^\perp$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^\perp$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:

1. $\text{Range}(P_U) =$

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$,

$P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:

1. $\text{Range}(P_U) = U$
2. $\text{Null}(P_U) =$

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:

1. Range$(P_U) = U$
2. Null$(P_U) = U^{\perp}$
3. Idempotent: $(P_U)^2 =$

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:

1. Range$(P_U) = U$
2. Null$(P_U) = U^{\perp}$
3. Idempotent: $(P_U)^2 = P_U$
4. (Conjugate) Symmetric: If $V = \mathbb{R}^n$ (or $\mathbb{C}^n$), $P_U^T =$

$V = \mathbb{C}^n$, $\langle x, y \rangle = y^* x$.

Let $v_1, v_2 \in \mathbb{C}^n$, $v_1 = u_1 + w_1$, $u_1 \in U$, $w_1 \in U^{\perp}$

$v_2 = u_2 + w_2$, $u_2 \in U$, $w_2 \in U^{\perp}$

$v_2^* P_U v_1$

$= \underline{\langle P_U v_1, v_2 \rangle} = \langle u_1, u_2 + w_2 \rangle = \langle u_1, u_2 \rangle$

$\langle v_1, P_U v_2 \rangle = \langle u_1, u_2 \rangle = v_2^* P_U^* v_1 \Rightarrow P_U^* = P_U$

Orthogonal Projection

- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:
  1. $\text{Range}(P_U) = U$
  2. $\text{Null}(P_U) = U^{\perp}$
  3. Idempotent: $(P_U)^2 = P_U$
  4. (Conjugate) Symmetric: If $V = \mathbb{R}^n$ (or $\mathbb{C}^n$), $P_U^T = P_U$ ($P_U^* = P_U$).

Orthogonal Projection

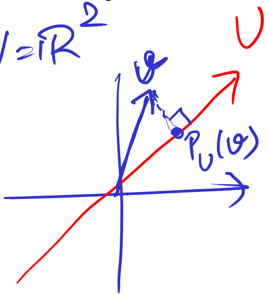
- Let U be a subspace of FD IPS V, and $V = U \oplus U^{\perp}$.
- Define $P_U \in \mathcal{L}(V)$ as $\forall v \in V$, if $v = u + w, u \in U, w \in U^{\perp}$, $P_U(v) = u$.
- P_U is said to be the *Orthogonal Projection Operator on* U.
- It has the following properties:

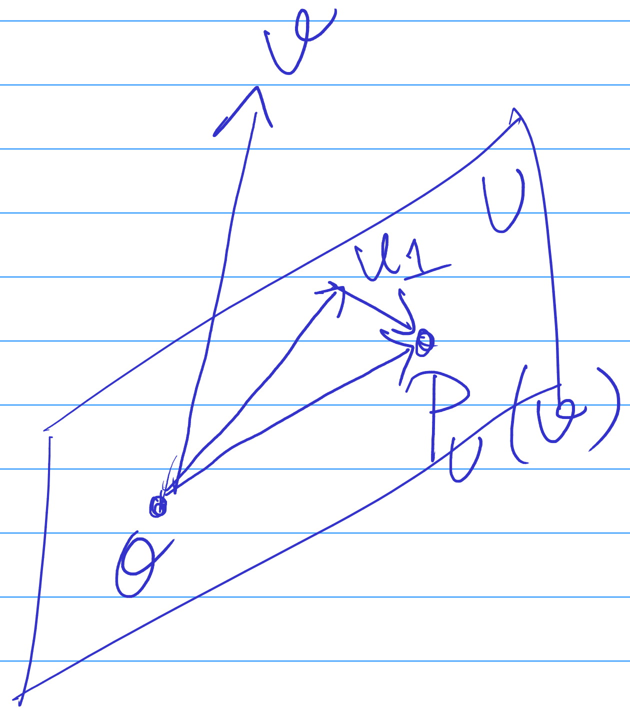
1. $\text{Range}(P_U) = U$
2. $\text{Null}(P_U) = U^{\perp}$
3. Idempotent: $(P_U)^2 = P_U$
4. (Conjugate) Symmetric: If $V = \mathbb{R}^n$ (or $\mathbb{C}^n$), $P_U^T = P_U$ ($P_U^* = P_U$).
5. $\forall v \in V, P_U(v) = \arg\min_{u \in U} ||u - v||^2$.

$V = \mathbb{R}^2$

$$x^* = \operatorname*{argmin}_{x \in \mathbb{R}^n} f(x)$$

$$\min_{x \in \mathbb{R}^n} f(x)$$





$\|u_1 - v\|^2 \qquad \|P_U(v) - v\|^2$

$P_U(v) = u_1 + u_2$

$\|u_1 - v\|^2 = \|P_U(v) - u_2 - v\|^2$

$\geq \|P_U(v) - v\|^2 + \|u_2\|^2$

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.
- Solution: $x^* = \arg\min_{x \in \mathbb{R}^n} ||Ax - b||^2$

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.
- Solution: $x^* = \arg\min_{x \in \mathbb{R}^n} ||Ax - b||^2$
- $Ax^* = P_{C(A)}(b)$

End of Class

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.
- Solution: $x^* = \arg\min_{x \in \mathbb{R}^n} ||Ax - b||^2$
- $Ax^* = P_{C(A)}(b)$, i.e., $\langle b - P_{C(A)}(b), Ax \rangle = 0, \forall x \in \mathbb{R}^n$.

Least Squares

- Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.
- Solution: $x^* = \arg\min_{x \in \mathbb{R}^n} ||Ax - b||^2$
- $Ax^* = P_{C(A)}(b)$, i.e., $\langle b - P_{C(A)}(b), Ax \rangle = 0, \forall x \in \mathbb{R}^n$.
- Let $a_{*i}, i = 1, \ldots, n$ denote the n column vectors. Then $\langle b - P_{C(A)}(b), a_{*i} \rangle = 0, \forall i = 1, \ldots, n$.

Least Squares

- Let $A \in \mathbb{R}^{m\times n}$ be a matrix with $rank(A) = n$.
- Solve for $x \in \mathbb{R}^n$ in $Ax = b$ such that $b \notin C(A)$.
- Solution: $x^* = \arg\min_{x\in\mathbb{R}^n} ||Ax - b||^2$
- $Ax^* = P_{C(A)}(b)$, i.e., $\langle b - P_{C(A)}(b), Ax\rangle = 0, \forall x \in \mathbb{R}^n$.
- Let $a_{*i}, i = 1, \ldots, n$ denote the n column vectors. Then $\langle b - P_{C(A)}(b), a_{*i}\rangle = 0, \forall i = 1, \ldots, n$.
-

$$\langle Ax^*, a_{*i}\rangle = \langle b, a_{*i}\rangle$$

Using Euclidean inner product

$$(a_{*i})^T Ax^* = (a_{*i})^T b, \forall i = 1, \ldots, n$$

$$A^T Ax^* = A^T b \Rightarrow x^* = \underbrace{(A^T A)^{-1} A^T}_{A^\dagger} b,$$

where $A^\dagger$ is known as the *Pseudo-inverse* of the matrix A.

Adjoint of an operator

● **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

Adjoint of an operator

- **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

- Let $V = \mathbb{C}^n$ be the IPS with IP $\langle x, y \rangle = y^* x$.

Adjoint of an operator

● **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

● Let $V = \mathbb{C}^n$ be the IPS with IP $\langle x, y \rangle = y^* x$. For $A \in \mathbb{C}^{n \times n}$, what is A^*?

Adjoint of an operator

- **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

- Let $V = \mathbb{C}^n$ be the IPS with IP $\langle x, y \rangle = y^* x$. For $A \in \mathbb{C}^{n \times n}$, what is A^*?
- Let $V = \{f \in \mathcal{C}^1[-1, 1] \mid f(-1) = f(1)\}$ be an IPS with IP: $\langle f, g \rangle = \int_{-1}^{1} f(x) \cdot g(x)\, dx$.

Adjoint of an operator

- **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

- Let $V = \mathbb{C}^n$ be the IPS with IP $\langle x, y \rangle = y^* x$. For $A \in \mathbb{C}^{n \times n}$, what is A^*?
- Let $V = \{f \in \mathcal{C}^1[-1,1] \mid f(-1) = f(1)\}$ be an IPS with IP: $\langle f, g \rangle = \int_{-1}^{1} f(x) \cdot g(x)\ dx$.What is $\frac{d}{dx}^*$?

Adjoint of an operator

● **Definition:** (Adjoint of an operator) Let $T \in \mathcal{L}(U)$, where U is an IPS with IP $\langle \cdot, \cdot \rangle$. Then adjoint of T, denoted by T^* is defined as

$$\forall x, y \in U, \langle Tx, y \rangle = \langle x, T^* y \rangle$$

● Let $V = \mathbb{C}^n$ be the IPS with IP $\langle x, y \rangle = y^* x$. For $A \in \mathbb{C}^{n \times n}$, what is A^*?

● Let $V = \{f \in \mathcal{C}^1[-1, 1] \mid f(-1) = f(1)\}$ be an IPS with IP: $\langle f, g \rangle = \int_{-1}^{1} f(x) \cdot g(x)\, dx$.What is $\frac{d}{dx}^*$?

● **Definition:** (Self-Adjoint Operator). Let $T \in \mathcal{L}(U)$, where $(U, \langle \cdot, \cdot \rangle)$ is an IPS. If $T = T^*$, then T is called a Self-adjoint operator.