MA-110 Linear Algebra and Differential Equations

Rekha Santhanam

Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai - 76

February 19, 2024
Lecture 20 D3

Orthogonal and Orthonormal Sets: Summary

Defn. A set of *non-zero* vectors $\{v_1, \cdots, v_k\} \subset \mathbb{R}^n$, is said to be an **orthogonal set** if $v_i^T v_j = 0$ **for all** $i, j = 1, \cdots n, \; i \neq j$.

Examples: $\{(1,3,1), (-1,0,1)\} \subset \mathbb{R}^3$,
$\{(2,1,0,-1), (0,1,0,1), (-1,1,0,-1)\} \subseteq \mathbb{R}^4$,
$\left\{(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}})\right\} \subseteq \mathbb{R}^3$, $\{e_1, \cdots, e_n\} \subseteq \mathbb{R}^n$.
Of these, the last two examples have all unit vectors (vectors of length one).

Defn. An orthogonal set $\{v_1, \cdots, v_k\} \subseteq \mathbb{R}^n$ with all unit vectors, i.e., $\|v_i\| = 1$ for all i, is called an **orthonormal set**.

Note: If $\{v_1, \cdots, v_k\}$ is an orthogonal set, then $\{u_1, \cdots, u_k\}$ is orthonormal, for $u_i = v_i/||v_i||$.

Exercise: If $S = \{v_1, \ldots, v_k\}$ is an orthogonal set, then v_k is orthogonal to each $v \in \text{Span}\{v_1, \ldots, v_{k-1}\}$.

Theorem: An orthogonal set in $\mathbb{R}^n$ is linearly independent.

Defn. A square matrix A whose column vectors form an orthonormal set is called an orthogonal matrix.

Orthogonal Matrices: Examples

Examples: 1. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 2. $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

3. $\frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$. 4. $\frac{1}{2}\begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$.

Orthogonal Basis

Defn. A basis $\mathscr{B} = \{v_1, \ldots, v_k\}$ of a subspace V of $\mathbb{R}^n$ is an orthogonal basis if it is an orthogonal set, i.e., ${v_i}^T v_j = 0$ for $i \neq j$. Furthermore, if $||v_i|| = 1$ for each i, then $\mathscr{B}$ is an orthonormal basis (or o.n.b.) of V.

Example: Consider the bases of $\mathbb{R}^2$:
$\mathscr{B}_1 = \{w_1 = (8,0)^T, w_2 = (6,3)^T\}$,
$\mathscr{B}_2 = \{(8,0)^T, (0,3)^T\}$ and
$$\mathscr{B}_3 = \left\{ \left(\frac{8}{\sqrt{8^2+0^2}}, 0 \right)^T, \left(0, \frac{3}{\sqrt{0^2+3^2}} \right)^T \right\}.$$

Then $\mathscr{B}_1$ is not orthogonal, $\mathscr{B}_2$ is an orthogonal basis, but not an orthonormal basis, and $\mathscr{B}_3$ is an orthonormal basis of $\mathbb{R}^2$.

Note: If $\{u_1, \ldots, u_k\} \subseteq \mathbb{R}^n$ is an orthonormal set, then it is an o.n.b. of $V = \text{Span}\{u_1, \ldots, u_k\}$.

Importance of Orthogonal Basis

Example : The set $\mathscr{B} = \{v_1 = (-1,1)^T, v_2 = (1,1)^T\}$ is a orthogonal basis of $\mathbb{R}^2$.

- Find $[v]_{\mathscr{B}} = (a,b)^T$:

$$v = av_1 + bv_2 = a(-1,1)^T + b(1,1)^T$$

$$v_1^T v = (-1,1)v = a(-1,1)(-1,1)^T = 2a = a||v_1||^2$$

Then $a = \dfrac{v_1^T v}{2} = \dfrac{v_1^T v}{||v_1||^2}$ and $b = \dfrac{v_2^T v}{2} = \dfrac{v_2^T v}{||v_2||^2}$

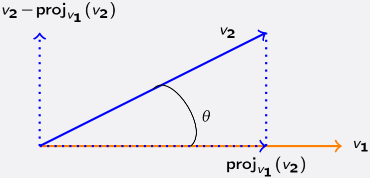
General Case: If $\mathscr{B} = \{v_1, \ldots, v_n\}$ is an o.n.b of V, then $[v]_{\mathscr{B}} = (c_1, \ldots, c_n)^T$, where $c_j = v_j^T v$.

Moreover, if $T : V \to V$ is linear, and $[T]_{\mathscr{B}}^{\mathscr{B}} = [a_{ij}]$, then $[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} [T(v_1)]_{\mathscr{B}} & \cdots & [T(v_n)]_{\mathscr{B}} \end{pmatrix} \Rightarrow a_{ij} = \_\_$.

Think!

When does T map orthogonal sets to orthogonal sets?

Every subspace of $\mathbb{R}^n$ has an orthogonal basis.
To construct one, we can start with any basis and modify it (Gram-Schmidt process).
First we see what happens in $\mathbb{R}^2$.



To construct an orthogonal basis in $\mathbb{R}^n$, we need to know how to find $\text{proj}_{v_1}(v_2)$ in $\mathbb{R}^n$.

Orthogonal Projections in $\mathbb{R}^n$

If $v(\neq 0), w \in \mathbb{R}^n$, then $\text{proj}_v(w)$, is a multiple of v and $w - \text{proj}_v(w)$ is orthogonal to v. Thus

$$\begin{aligned} \text{proj}_v w &= av \text{ for some } a \in \mathbb{R} \\ v^T(w - \text{proj}_v w) &= 0 \\ v^T w - v^T av = 0 &\Leftrightarrow a = \frac{v^T w}{v^T v} \end{aligned}$$

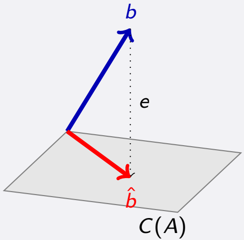
Therefore $\boxed{\text{proj}_v(w) = \left(\frac{v^T w}{v^T v}\right) v.}$

Example. If $w = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$ and $v = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^T$, then the orthogonal projection of w on Span$\{v\}$ is given by

$$\text{proj}_v(w) = \left(\frac{v^T w}{v^T v}\right) v = \frac{6}{14}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Application: This can be used to “solve” an inconsistent system of equations.

Suppose system $Ax = b$ is inconsistent, i.e. $b \notin C(A)$. The error $E = \|Ax - b\|$ is the distance from b to $Ax \in C(A)$.



We want the least square solution $\hat{x}$ which minimizes E, i.e., we want to find $\hat{b}$ closest to b such that $A\hat{x} = \hat{b}$ is a consistent system.
Therefore, $\hat{b} = \text{proj}_{C(A)}(b)$ and $A\hat{x} = \hat{b}$.

The error vector $e = b - A\hat{x}$ must be perpendicular to $C(A)$, which is also the row space of A^T.
So, e must be in the left null space of A, $N(A^T)$, i.e.,

$$A^T(b - A\hat{x}) = 0 \text{ or } \boxed{A^TA\hat{x} = A^Tb}$$

Therefore, to find $\hat{x}$, we need to solve $A^TA\hat{x} = A^Tb$.

Linear Least Squares and Projections

Let A be $m \times n$. Then $A^T A$ is a symmetric $n \times n$ matrix.

- $\boxed{N(A^T A) = N(A)}$.

Proof. $Ax = 0 \Rightarrow A^T Ax = 0$. So, $N(A) \subseteq N(A^T A)$.
For the other inclusion, take $x \in N(A^T A)$.
$A^T Ax = 0 \Rightarrow x^T(A^T Ax) = (Ax)^T(Ax) = \|Ax\|^2 = 0$
$\Rightarrow Ax = 0$,i.e., $x \in N(A)$.

- Since $N(A) = N(A^T A)$, by rank-nullity theorem,
rank$(A) = n - \dim\ (N(A)) =$ rank$(A^T A)$.

- A has linearly independent columns $\Leftrightarrow$ rank$(A) = n \Leftrightarrow$ rank$(A^T A) = n \Leftrightarrow A^T A$ is invertible.

- If rank$(A) = n$, then the least square solution of $Ax = b$ is given by $A^T A\hat{x} = A^T b \Rightarrow \boxed{\hat{x} = (A^T A)^{-1} A^T b}$ and the orthogonal proj. of b on $C(A)$ is $\boxed{\hat{b} = A\hat{x} = Pb}$, where $\boxed{P = A(A^T A)^{-1} A^T}$ is the projection matrix. **Ques:** Is $P^2 = P$?

Linear Least Squares: Example

Example: Find the least square solution to the system

$$\begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix} x = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \quad (Ax = b)$$

We need to solve $A^T A \hat{x} = A^T b$. Now $A^T b = \begin{pmatrix} -4 \\ 11 \end{pmatrix}$ and

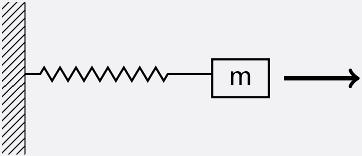
$$A^T A = \begin{pmatrix} 6 & -11 \\ -11 & 22 \end{pmatrix}.$$

$$[A^T A \mid A^T b] = \begin{pmatrix} 6 & -11 & | & -4 \\ -11 & 22 & | & 11 \end{pmatrix} \to \begin{pmatrix} 6 & -11 & | & -4 \\ 0 & 11/6 & | & 11/3 \end{pmatrix}.$$

Therefore $\hat{x}_2 = 2$, and $\hat{x}_1 = 3$.

Exercise: Find the projection matrix P, and check that $Pb = A\hat{x}$.

Reading Slide: Linear Least Squares: Application



Hooke's Law states that displacement x of the spring is directly proportional to the load (mass) applied, i.e., $m = kx$.

A student performs experiments to calculate spring constant k. The data collected says for loads $4, 7, 11$ kg applied, the displacement is $3, 5, 8$ inches respectively. Hence we have:

$$\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} k = \begin{pmatrix} 4 \\ 7 \\ 11 \end{pmatrix} \qquad (ak = b).$$

Clearly the data is inconsistent.
Allowing for various errors, how do we find an estimate for k?
The method of least squares allows us to find a consistent system "close" to this one!
Exercise: Estimate k using the method of least squares.

Question: We want to find the best line $y = C + Dx$ which fits the given data and gives least square error.
Data: $(x, y) = (-2, 4)$, $(-1, 3)$, $(0, 1)$, and $(2, 0)$.

The system $\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 1 \\ 0 \end{pmatrix} \quad (Ax = b)$

is inconsistent.

Find the least square solution by solving $A^T A \hat{x} = A^T b$.

Question: Find the best quadratic curve $y = C + Dx + Ex^2$ which fits the above data and gives least square error.

Hint. The first row of the matrix A in this case will be $[1 \ \ -2 \ \ 4]$.

Gram-Schmidt Process

If the set of vectors $v_1, \ldots, v_r$ in $\mathbb{R}^n$ are linearly independent, then we can find an orthonormal set of vectors $q_1, \ldots, q_r$ such that $\text{Span}\{v_1, \ldots, v_r\} = \text{Span}\{q_1, \ldots, q_r\}$.

First find an orthogonal set.

Let $w_1 = v_1$, $w_2 = v_2 - \text{proj}_{w_1}(v_2)$. Then $w_1 \perp w_2$ and

$\text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$.

Let $c_1 w_1 + c_2 w_2$ be the projection of v_3 on $\text{Span}\{w_1, w_2\}$. Then

$(v_3 - c_1 w_1 - c_2 w_2) \perp w_1$ and $(v_3 - c_1 w_1 - c_2 w_2) \perp w_2$.
$\Rightarrow w_1^T (v_3 - c_1 w_1 - c_2 w_2) = 0 \Rightarrow c_1 w_1 = \text{proj}_{w_1}(v_3)$ and similarly $c_2 w_2 = \text{proj}_{w_2}(v_3)$. Therefore,

$$w_3 = v_3 - \text{proj}_{\text{Span}\{w_1, w_2\}}(v_3) = v_3 - \left(\frac{w_1^T v_3}{\|w_1\|^2}\right) w_1 - \left(\frac{w_2^T v_3}{\|w_2\|^2}\right) w_2.$$

$$\boxed{\text{Span}\{v_1, v_2, v_3\} = \text{Span}\{w_1, w_2, w_3\} \text{ and } w_1^T w_3 = 0, w_2^T w_3 = 0.}$$

By induction,

$$w_r := v_r - \mathrm{proj}_{\mathrm{Span}\{w_1,\ldots w_{r-1}\}}(v_r) =$$
$$v_r - \mathrm{proj}_{w_1}(v_r) - \mathrm{proj}_{w_2}(v_r) - \cdots - \mathrm{proj}_{w_{r-1}}(v_r)$$
$$= v_r - \frac{w_1^T v_r}{\|w_1\|^2} w_1 - \frac{w_2^T v_r}{\|w_2\|^2} w_2 - \cdots - \frac{w_{r-1}^T v_r}{\|w_{r-1}\|^2} w_{r-1}$$

Now take $q_1 = \dfrac{w_1}{\|w_1\|}$, $q_2 = \dfrac{w_2}{\|w_2\|}$, …, $q_r = \dfrac{w_r}{\|w_r\|}$. Then $\{q_1, \ldots, q_r\}$ is an orthonormal set and $W = \mathrm{Span}\{v_1, \ldots, v_r\} = \mathrm{Span}\{w_1, \ldots, w_r\} = \mathrm{Span}\{q_1, \ldots, q_r\}$.

In particular, $\{q_1, q_2, \ldots, q_r\}$ is an *orthonormal basis* for W.

Exercise: Show that if $\{w_1, \ldots, w_r\}$ is an orthogonal set, then

$$\boxed{\mathrm{proj}_{\mathrm{Span}\{w_1,\ldots w_{i-1}\}}(v_i) = \mathrm{proj}_{w_1}(v_i) + \mathrm{proj}_{w_2}(v_i) + \cdots + \mathrm{proj}_{w_{i-1}}(v_i).}$$

Gram-Schmidt Method: Example

Q: Let $S = \left\{ v_1 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} -5 \\ 1 \\ 5 \\ -7 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} \right\}$ and $W =$ Span(S). Find an orthonormal basis for W.

Exercise: First verify that $\{v_1, v_2, v_3\}$ are linearly independent. (Check that rank of $\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is 3). Hence S is a basis of W.