MA-106 Linear Algebra

H. Ananthnarayan

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

19th February 2018
D1 - Lecture 20

Random Attendance

1. 17D070059 Yash Sharma
2. 170050002 Mashkaria Satvik Mehulbhai
3. 170050005 Yateesh Agrawal Absent
4. 170050024 Chitrank Gupta
5. 170050057 Arpit Menaria
6. 170050059 Saurav Yadav
7. 170050091 Amanaganti Rohan Ganesh
8. 170050018 Shubhamkar Bajrang Ayare Absent
9. 170050108 Ujjval Goury
10. 170070010 Soumya Chatterjee
11. 170070017 Ojas Sanjiv Thakur
12. 170070018 Himanshu Baheti
13. 170070046 Rishabh Gopichand Ramteke
14. 170070051 Koustav Jana
15. 17D070013 Paras Vijay Bodake Absent
16. 17D070024 Prajwal Dnyaneshwar Kamble Absent

Recall: Inner product on $\mathbb{R}^n$

- $\boxed{v \cdot w = v^T w}$ for $v, w \in \mathbb{R}^n$ defines an inner product on $\mathbb{R}^n$.
- The *norm* (or length) of $v \in \mathbb{R}^n$ is $\boxed{\|v\| = \sqrt{v \cdot v}.}$
- Vectors $v, w \in \mathbb{R}^n$ are *orthogonal* to each other if $v \cdot w = 0$.
- The subset $\{v_1, \cdots, v_k\} \subseteq \mathbb{R}^n$ is an *orthogonal set* if $v_i \neq 0$ for all i and $v_i^T v_j = 0$ for $i \neq j$.
- An orthogonal set $\{v_1, \cdots, v_k\} \subseteq \mathbb{R}^n$ is *orthonormal* if each v_i is a unit vector.

e.g., if $\{v_1, \cdots, v_k\}$ is an orthogonal set, then $\{v_1/||v_1||, \cdots, v_k/||v_k||\}$ is orthonormal.

Orthogonal Matrices

- An orthogonal subset of $\mathbb{R}^n$ is linearly independent.
- If the columns of an $m \times n$ matrix A form an orthonormal set, then $A^T A = I_n$.
In particular, if $m = n$, then A is invertible, $A^{-1} = A^T$ (why?), and $\det(A) = \pm 1$.
- A square matrix Q whose column vectors form an orthonormal set is called an *orthogonal* matrix.
- If Q is an $n \times n$ orthogonal matrix, then for each $v \in \mathbb{R}^n$, $||Qv|| = ||v||$.
In particular, the only (real) eigenvalues of Q, if they exist, are ± 1.

Orthogonal Basis

Defn. A basis $\mathcal{B} = \{v_1, \ldots, v_k\}$ 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 each v_i is a unit vector, then $\mathcal{B}$ is an *orthonormal basis* (or o.n.b.) of V.

Example: Consider the bases of $\mathbb{R}^2$:
$\mathcal{B}_1 = \{w_1 = (8,0)^T, w_2 = (6,3)^T\}$, $\mathcal{B}_2 = \{(8,0)^T, (0,3)^T\}$ and

$$\mathcal{B}_3 = \left\{ \left(\frac{8}{\sqrt{8^2+0^2}}, 0 \right)^T, \left(0, \frac{3}{\sqrt{0^2+3^2}} \right)^T \right\}.$$

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

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

Importance of Orthogonal Basis

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

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

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

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

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

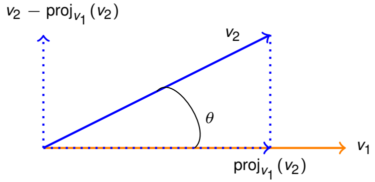
General Case: If $\mathcal{B} = \{v_1, \ldots, v_n\}$ is an o.n.b of V, then $[v]_{\mathcal{B}} = (c_1, \ldots c_n)^T$, where $c_j = v_j^T v$.

Moreover, if $T : V \to V$ is linear, and $[T]_{\mathcal{B}}^{\mathcal{B}} = [a_{ij}]$, then $[T]_{\mathcal{B}}^{\mathcal{B}} = \big([T(v_1)]_{\mathcal{B}} \quad \cdots \quad [T(v_n)]_{\mathcal{B}}\big) \Rightarrow a_{ij} = \_\_$.

Q: When does T map orthogonal sets to orthogonal sets?

Orthogonal Basis and Projections

- Every subspace of $\mathbb{R}^n$ has an orthogonal basis.
- An orthogonal basis will help us to define projections in $\mathbb{R}^n$.
- Using projections, we will discuss an application to approximation: Method of linear least squares.



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: $\mathbb{R}^n$

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

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

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

Example. If $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, then find the orthogonal projection of w on $\text{Span}\{v\}$.

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

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 $\mathrm{Span}\{v_1, \ldots, v_r\} = \mathrm{Span}\{q_1, \ldots, q_r\}$.

First find an orthogonal set.

Let $w_1 = v_1$, $w_2 = v_2 - \mathrm{proj}_{w_1}(v_2)$.

Then $w_1 \perp w_2$ and $\mathrm{Span}\{v_1, v_2\} = \mathrm{Span}\{w_1, w_2\}$.

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

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

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

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

Gram-Schmidt Process Contd.

By induction,

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

Now take $q_1 = \frac{w_1}{\|w_1\|}$, $q_2 = \frac{w_2}{\|w_2\|}$, …, $q_r = \frac{w_r}{\|w_r\|}$.

Then $\{q_1, \dots, q_r\}$ is an orthonormal set and
$W = \text{Span}\{v_1, \dots, v_r\} = \text{Span}\{w_1, \dots, w_r\} = \text{Span}\{q_1, \dots, q_r\}$.
In particular, $\{q_1, q_2, \dots, q_r\}$ is an *orthonormal basis* for W.

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

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

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 $W = \text{Span}(S)$. Find an orthonormal basis for W.

Exercise: 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. Use Gram-Schmidt method,

$$w_1 = v_1,\ w_2 = v_2 - \left(\frac{w_1^T v_2}{\|w_1\|^2} \right) w_1$$

$$\Rightarrow \quad w_2 = v_2 - \left(\frac{-15 + 1 - 5 - 21}{9 + 1 + 1 + 9} \right) w_1 = v_2 - \left(\frac{-40}{20} \right) w_1$$

$$\Rightarrow \quad w_2 = v_2 + 2w_1. = \begin{pmatrix} 1 & 3 & 3 & -1 \end{pmatrix}^T.$$

Observe: $v_1,\ v_2 \in \text{Span}\{w_1, w_2\},\ w_1,\ w_2 \in \text{Span}\{v_1, v_2\} \Rightarrow$ $\text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$.

Gram-Schmidt Method: Example

Recall $w_1 = \begin{pmatrix} 3 & 1 & -1 & 3 \end{pmatrix}^T$, $w_2 = \begin{pmatrix} 1 & 3 & 3 & -1 \end{pmatrix}^T$, and $v_3 = \begin{pmatrix} 1 & 1 & -2 & 8 \end{pmatrix}^T$. (Check $w_1^T w_2 = 0$).

Now $w_3 = v_3 - \left(\dfrac{w_1^T v_3}{\|w_1\|^2} \right) w_1 - \left(\dfrac{w_2^T v_3}{\|w_2\|^2} \right) w_2$.

$$w_3 = v_3 - \left(\frac{3+1+2+24}{20} \right) w_1 - \left(\frac{1+3-6-8}{20} \right) w_2$$

$$\Rightarrow \quad w_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 3 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 3 \end{pmatrix}.$$

Check $w_1^T w_3 = 0 = w_2^T w_3$; Span$\{v_1, v_2, v_3\}$ = Span$\{w_1, w_2, w_3\}$.

Hence $\{w_1, w_2, w_3\}$ is an orthogonal basis of W.

An orthonormal basis for W is $\left\{ \dfrac{1}{\sqrt{20}} w_1, \dfrac{1}{\sqrt{20}} w_2, \dfrac{1}{\sqrt{20}} w_3 \right\}$.

Reading: QR Factorization

Let $A = \begin{pmatrix} v_1 & \dots & v_r \end{pmatrix}$ be an $n \times r$ matrix of rank r. Then $v_1, \dots, v_r$ are linearly independent vectors in $\mathbb{R}^n$.

By the Gram-Schmidt method, we get an orthonormal basis $\{q_1, \dots, q_r\}$ of $C(A)$, where $q_i = \dfrac{w_i}{\|w_i\|}$ and $w_1 = v_1$,

$$w_k = v_k - \left(\frac{w_1^T v_k}{\|w_1\|^2}\right) w_1 - \cdots - \left(\frac{w_{k-1}^T v_k}{\|w_{k-1}\|^2}\right) w_{k-1}.$$

Let $Q = \begin{pmatrix} q_1 & \dots & q_r \end{pmatrix}$. $\boxed{\text{How are } A \text{ and } Q \text{ related?}}$

$\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{w_1, \dots, w_k\} = \text{Span}\{q_1, \dots, q_k\}\ \forall k$.

If $v_k = c_1 q_1 + \ldots + c_k q_k$, then
$c_1 = q_1^T v_k, \quad c_2 = q_2^T v_k, \quad \dots, \quad c_k = q_k^T v_k$.

$$\text{Hence } v_k = (q_1^T v_k) q_1 + \ldots + (q_k^T v_k) q_k.$$

Reading: QR factorization

$$v_k = (q_1^T v_k)q_1 + \ldots + (q_k^T v_k)q_k \quad \text{for each } k.$$

Therefore

$$\begin{pmatrix} v_1 & v_2 & \ldots & v_r \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \ldots & q_r \end{pmatrix} \begin{pmatrix} q_1^T v_1 & q_1^T v_2 & & q_1^T v_r \\ 0 & q_2^T v_2 & & q_2^T v_r \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & q_r^T v_r \end{pmatrix}$$

i.e. $\boxed{A = QR}$, where the columns of Q form an orthonormal set and R is an invertible $r \times r$ matrix. **Q:** Why is R invertible?
This is called QR-factorization of A.

• If A is invertible $n \times n$, then $\boxed{A = QR}$, where Q is an orthogonal matrix and R is an invertible upper triangular matrix, both are $n \times n$ matrices.