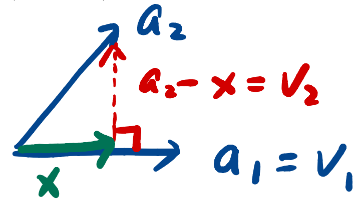
Lecture 23: Quick review from previous lecture

- (Gram-Schmidt Process)
  Suppose that $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are linearly independent.

Turn $\mathbf{a}_1, \cdots, \mathbf{a}_n$ to orthogonal vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$:



$$\mathbf{v}_1 = \mathbf{a}_1$$

$$\mathbf{v}_2 = \mathbf{a}_2 - \overbrace{\frac{\langle \mathbf{a}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1}^{= x}$$

$$\mathbf{v}_3 = \mathbf{a}_3 - \frac{\langle \mathbf{a}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{a}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

$$\vdots$$

$$\mathbf{v}_n = \mathbf{a}_n - \frac{\langle \mathbf{a}_n, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{a}_n, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|^2}\mathbf{v}_{n-1}$$

$\langle v_i, v_j \rangle = 0, \quad i \neq j.$

Today we will discuss

- Sec. 4.2 The Gram-Schmidt process.
- Sec. 4.3 The Orthogonal Matrices

- Lecture will be recorded -

Example 2. We know that $\{1, x, x^2\}$ forms a basis for $\mathcal{P}^{(2)}([0,1])$, the space of polynomials of degree ≤ 2 on $[0,1]$. Let's turn them into an orthonormal basis, with respect to the usual L^2 inner product.

By Gram-Schmidt process, we turn them into orthogonal vectors.

$$p_1 = 1$$

$$p_2 = x - \frac{\langle x, p_1 \rangle}{\|p_1\|^2} p_1$$

$$= x - \frac{1}{2}.$$

- $\langle x, p_1 \rangle = \int_0^1 x \, 1 \, dx = \frac{1}{2}$

  $\|p_1\|^2 = \langle p_1, p_1 \rangle = \int_0^1 1 \, dx = 1$

$$p_3 = x^2 - \frac{\langle x^2, p_1 \rangle}{\|p_1\|^2} p_1 - \frac{\langle x^2, p_2 \rangle}{\|p_2\|^2} p_2.$$

- $\langle x^2, p_1 \rangle = \int_0^1 x^2 \cdot 1 \, dx = 1/3.$

  $\langle x^2, p_2 \rangle = \int_0^1 x^2 (x - \frac{1}{2}) \, dx = \frac{1}{12}$

  $\|p_1\|^2 = 1$

  $\|p_2\|^2 = \int_0^1 (x - \frac{1}{2})^2 \, dx = \int_0^1 x^2 - x + \frac{1}{4} \, dx = \frac{1}{12}$

Then $p_3 = x^2 - \frac{\frac{1}{3}}{1} 1 - \frac{\frac{1}{12}}{\frac{1}{12}} (x - \frac{1}{2})$

$$= x^2 - x + \frac{1}{6}$$

Now we get orthogonal basis $\{p_1, p_2, p_3\}.$

Thus, an orthonormal basis is

$\|p_2\|^2 = \frac{1}{12}$

$\|p_2\| = \sqrt{\frac{1}{12}}$

$$\left\{ \frac{p_1}{\|p_1\|}, \frac{p_2}{\|p_2\|}, \frac{p_3}{\|p_3\|} \right\}.$$

Exercise find $\|p_3\|$.

$$= \left\{ 1, \frac{x - \frac{1}{2}}{\frac{1}{\sqrt{12}}}, \frac{p_3}{\|p_3\|} \right\}$$

$\sqrt{12}\,(x-\frac{1}{2})$

Example 3. Let the subspace $W \subset \mathbb{R}^4$ consisting of all vectors orthogonal to $\mathbf{a} = (1,1,1,0)^T$. Find an orthonormal basis for W under the standard dot product.

(1) Find $W = \left\{ \vec{x} = \begin{pmatrix} a\\ b\\ c\\ d \end{pmatrix} \in \mathbb{R}^4 \;\middle|\; \langle \vec{x}, \vec{a} \rangle = 0 \right\}$.

$$0 = \langle \vec{x}, \vec{a} \rangle = a+b+c = \underbrace{\begin{bmatrix} \boxed{1} & 1 & 1 & 0 \end{bmatrix}}_{1\times 4} \begin{bmatrix} a\\ b\\ c\\ d \end{bmatrix} \quad \text{(linear system)}$$

Free variables: b, c, d

$a = -b-c$,

$$W = \left\{ \begin{pmatrix} -b-c\\ b\\ c\\ d \end{pmatrix} \;\middle|\; b, c, d \in \mathbb{R} \right\}.$$

A basis of W is:

$b=1,\ c=d=0$ $\quad$ $c=1,\ d=b=0$ $\quad$ $d=1,\ b=c=0$

$$\left\{ \underset{w_1}{\begin{pmatrix} -1\\ 1\\ 0\\ 0 \end{pmatrix}}, \underset{w_2}{\begin{pmatrix} -1\\ 0\\ 1\\ 0 \end{pmatrix}}, \underset{w_3}{\begin{pmatrix} 0\\ 0\\ 0\\ 1 \end{pmatrix}} \right\} \quad \text{(NOT orthogonal yet)}$$

(2) Using G-S process

$$v_1 = w_1 = \begin{pmatrix} -1\\ 1\\ 0\\ 0 \end{pmatrix}$$

$$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1 = \begin{pmatrix} -1/2\\ -1/2\\ 1\\ 0 \end{pmatrix}.$$

$$v_3 = w_3 - \underbrace{\frac{\langle w_3, v_1 \rangle}{\|v_1\|^2}}_{=0} v_1 - \underbrace{\frac{\langle w_3, v_2 \rangle}{\|v_2\|^2}}_{0} v_2 .$$

$$= \begin{pmatrix} 0\\ 0\\ 0\\ 1 \end{pmatrix}$$

Then $\{v_1, v_2, v_3\}$ is an orthogonal basis

So, O.N.B. is $\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \frac{v_3}{\|v_3\|} \right\}$

$$= \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} -1\\ 1\\ 0\\ 0 \end{pmatrix}, \frac{1}{\|v_2\|} \begin{pmatrix} -1/2\\ -1/2\\ 1\\ 0 \end{pmatrix}, \begin{pmatrix} 0\\ 0\\ 0\\ 1 \end{pmatrix} \right\}$$

4.3 Orthogonal Matrices

↓ Exercise find ||v_n||.

Definition: A square matrix A is called an **orthogonal matrix** if it satisfies

$$A^T A = AA^T = I.$$

Recall: We call X is the inverse of A if

$$XA = AX = I. \quad \text{Denote } X = A^{-1}$$

Some Properties about an orthogonal matrix A:

Fact 1: Let A be an **orthogonal matrix**.

(1) The inverse of A is

$$A^T = A^{-1}.$$

(2) The solution to the linear system $A\mathbf{x} = \mathbf{v}$ is $\mathbf{x} = A^T\mathbf{v}$.[a]

(3) $\det(A) = \pm 1$.

[a] Thus there is NO need to apply Gaussian elimination to solve this system

(2). $Ax = v$. $A^T(Ax) = A^T v$

$A^T(Ax) = Ix = x$ since $A^TA = I$.

(3) $A^TA = I \Rightarrow \underbrace{\det I}_{=1} = \det(A^TA) = \det A^T \det A = (\det A)^2$.

$\det A = 1, -1.$ #

Moreover, we have

Fact 2: A square matrix A is an **orthogonal matrix** if and only if its columns form an **orthonormal basis** on $\mathbb{R}^n$ with respect to the Euclidean dot product.

[To see this:] $A = [v_1 \cdots v_n]_{n\times n}$. Since A is orthogonal,

$$I = A^TA = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} [v_1 \cdots v_n] = \begin{bmatrix} v_1^Tv_1 & v_1^Tv_2 & \cdots & v_1^Tv_n \\ \vdots & \vdots & & \vdots \\ v_n^Tv_1 & v_n^Tv_2 & \cdots & v_n^Tv_n \end{bmatrix}$$

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = \begin{bmatrix} v_1^Tv_1 & v_1^Tv_2 & \cdots & v_1^Tv_n \\ \vdots & \vdots & & \vdots \\ v_n^Tv_1 & v_n^Tv_2 & \cdots & v_n^Tv_n \end{bmatrix} \Rightarrow \begin{array}{l} \|v_i\| = 1, \ 1 \le i \le n. \\ v_i^T v_j = 0, \ i \ne j. \end{array}$$

$\Rightarrow \{v_i\}$ is O.N.B.

Fact 3: If A is orthogonal, so is A^T (since $(A^T)^T = A$).

* This implies that the column vectors of A^T (they are row vectors of A) also from an orthonormal basis of $\mathbb{R}^n$.

Example 1. Permutation matrix $P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ is orthogonal.

Method 1: $P^TP = PP^T = I_3$

Method 2: Checking (rows) columns are O.N.B.

* *One can easily see that the rows (columns) of the above matrix form an orthonormal basis for* $\mathbb{R}^3$.

Example 2.

(1) Is $B = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ orthogonal? NO since $\det B \neq 1$.

$v_1 \quad v_2$

(2) How to turn B in (1) into an orthogonal matrix?

$\langle v_1, v_2 \rangle = 0$. Then $\{v_1, v_2\}$ is orthogonal.

Fact 2 $\longrightarrow$ Turn them into orthonormal basis by

$q_1 = \frac{v_1}{\|v_1\|}$, $q_2 = \frac{v_2}{\|v_2\|}$. Then $B \longrightarrow Q = [q_1 \ q_2]$ orthogonal matrix

$= \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$.

Fact 4: If A and B are **orthogonal matrices**, then AB is **orthogonal** too.

[To see this:] $(AB)^T AB = B^T A^T AB \overset{\text{since } A \text{ is orthogonal}}{=} B^T I B$

$= B^T B$

$= I$ since B is orthogonal!

$(AB)(AB)^T = I$ (Exercise).

Fact 5: If A is orthogonal, then the matrix A preserve length in the sense that

$$\|A\mathbf{x}\| = \|\mathbf{x}\| \qquad \text{for all } x \in \mathbb{R}^n \qquad \text{(Homework problem)},$$

where $\|\cdot\|$ denotes the 2-norm.

§ The QR Factorization

Q: How do we turn the matrix $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ to an orthogonal matrix?

Answer: This can be achieved by applying the Gram-Schmidt equation.

Let's revisit the Gram-Schmidt process:

Let

$$A = [\mathbf{a}_1 \cdots \mathbf{a}_n] \text{ be } n \times n \text{ nonsingular matrix}$$

where $\mathbf{a}_j$ is the i^{th} column vector of A. Thus, $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are linearly independent.

Step 1: Turn $\mathbf{a}_1, \cdots, \mathbf{a}_n$ to orthogonal vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$:

Gram-Schmidt process

$$\left\{\begin{aligned}
\mathbf{v}_1 &= \mathbf{a}_1 \\
\mathbf{v}_2 &= \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{a}_3 - \frac{\langle \mathbf{a}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{a}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_n &= \mathbf{a}_n - \frac{\langle \mathbf{a}_n, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{a}_n, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|^2}\mathbf{v}_{n-1}
\end{aligned}\right.$$

Then

$$\underbrace{\left(\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right)}_{A} \underbrace{\longrightarrow}_{GramSchmidt} \underbrace{\left(\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{array}\right)}_{V}$$

Step 2: Normalize $\mathbf{v}_1, \cdots, \mathbf{v}_n$ to get orthonormal vectors $\mathbf{q}_1, \cdots, \mathbf{q}_n$:
That is,

$$\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}$$

Then a nonsingular matrix A is turned to an orthogonal matrix Q:

$$\underbrace{\left(\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right)}_{A \text{ (nonsingular)}} \underbrace{\longrightarrow}_{GramSchmidt+Normalization} \underbrace{\left(\begin{array}{c|c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{array}\right)}_{Q \text{ (orthogonal)}}$$

To answer the earlier question:

Example 3. Now let's turn the matrix $A = \begin{array}{cc} a_1 & a_2 \\ \begin{pmatrix} 1 \\ 2 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{array}$ to an orthogonal matrix.

$v_1 = a_1$

$v_2 = a_2 - \frac{\langle a_2, v_1 \rangle}{\|v_1\|^2} v_1$

$q_i = v_i / \|v_i\|$

$A \longrightarrow Q = [\, q_1 \quad q_2 \,]$

(Continue Next time)

Q: From above example, we have seen that $A \to Q$. Indeed, we will be able to decompose

$$A = Q\blacksquare.$$

What is this matrix $\blacksquare$ that encodes all processes turning A to Q? This matrix $\blacksquare$ is actually upper triangular.

Let's figure this out.

Rewrite Step 1 above as follows:

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{a}_1 \\
\mathbf{v}_2 &= \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{a}_3 - \frac{\langle \mathbf{a}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{a}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_n &= \mathbf{a}_n - \frac{\langle \mathbf{a}_n, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{a}_n, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|^2}\mathbf{v}_{n-1}
\end{aligned}$$

Poll Question 1: Then Gram-Schmidt process can turn every **linearly independent** vectors into mutually orthogonal vectors?

A) Yes
B) No