Orthonormal Bases

Say $\vec{q}_1, \ldots, \vec{q}_n$ in $\mathbb{R}^m$ are orthogonal if

$$\boxed{\vec{q}_i \cdot \vec{q}_j = 0 \iff \vec{q}_i^{\,T} \vec{q}_j = 0 \quad \text{for all } i \neq j}$$

We say they're orthonormal if

$$\boxed{\vec{q}_i^{\,T} \vec{q}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}}$$

Write $\underset{m \times n}{Q} = \begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix}$

$$\overset{n \times n}{\underset{n \times m \;\; m \times n}{Q^T Q}} = \begin{bmatrix} - & \vec{q}_1 & - \\ & \vdots & \\ - & \vec{q}_n & - \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{q}_1 & \cdots & \vec{q}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} \vec{q}_1^{\,T}\vec{q}_1 & \vec{q}_1^{\,T}\vec{q}_2 & \cdots & \vec{q}_1^{\,T}\vec{q}_n \\ \vec{q}_2^{\,T}\vec{q}_1 & \ddots & & \vdots \\ \vdots & & & \\ \vec{q}_n^{\,T}\vec{q}_n & \cdots & & \vec{q}_n^{\,T}\vec{q}_n \end{bmatrix}$$

if $\vec{q}_i$'s are orthonormal

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & 0 \\ 0 & & & 0 & 1 \end{bmatrix} = I$$

Thus $\boxed{Q \text{ has orthonormal columns} \iff Q^T Q = I \text{ (identity)}}$

Special case: $m=n$, so Q is $n\times n$ (square)

then $\boxed{Q^TQ = I \iff Q^{-1} = Q^T}$ (inverse = transpose)

When an $n\times n$ matrix Q satisfies $Q^{-1} = Q^T$,

it is called an <u>orthogonal matrix</u>.

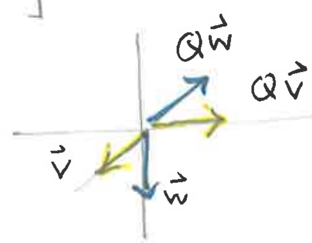
Examples

1)
$$Q = \begin{array}{ccc} \vec{q}_1 & \vec{q}_2 & \vec{q}_3 \\ \end{array}$$
$$Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

This is an orthogonal matrix. Note it's a permutation matrix.

$$\rightarrow \quad Q^{-1} = Q^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$Q\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ x \\ y \end{bmatrix}$$

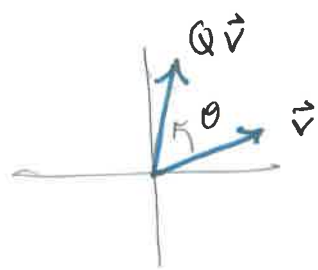


2)
$$\begin{array}{cc} \vec{q}_1 & \vec{q}_2 \end{array}$$
$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

(θ fixed real #)

$\vec{q}_1 \cdot \vec{q}_2 = \cos\theta(-\sin\theta) + \sin\theta(\cos\theta) = 0.$

$\vec{q}_1 \cdot \vec{q}_1 = \cos^2\theta + \sin^2\theta = 1$, similar for $\vec{q}_2$.



$$\rightarrow \quad Q^{-1} = Q^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

An orthogonal matrix $\underset{m\times n}{Q}$ preserves lengths of vectors:

$$\|Q\vec{x}\|^2 = (Q\vec{x})\cdot(Q\vec{x}) = (Q\vec{x})^T(Q\vec{x}) = \vec{x}^TQ^TQ\vec{x}$$

$Q^TQ = I$ since Q orthog.

$$= \vec{x}^T\vec{x}$$

(We also used $(AB)^T = B^TA^T$ for any matrices $\underset{m\times n}{A}, \underset{n\times k}{B}$)

$$= \vec{x}\cdot\vec{x} = \|\vec{x}\|^2$$

$\rightarrow$ $\boxed{\|Q\vec{x}\| = \|\vec{x}\| \text{ for any vector } \vec{x} \text{ in } \mathbb{R}^n}$

Suppose we want to project onto

$$\text{span}(\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n) \subset \mathbb{R}^m$$

where the $\vec{q}_i$'s are orthonormal. We let

$$A = Q = \begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix}$$

Then we form the projection matrix:

$$P = A(A^TA)^{-1}A^T$$

$$= Q(Q^TQ)^{-1}Q^T$$

$Q^TQ = I$

$$= QQ^T$$

$\rightarrow$ $\boxed{P = QQ^T}$ Projections are easy with orthonormal bases!

To project $\vec{b}$ onto $C(Q) = \text{span}(\vec{q}_1, \ldots, \vec{q}_n)$ take $P\vec{b}$:

$$P\vec{b} = QQ^T\vec{b} = \begin{bmatrix} | & & | \\ \vec{q}_1 & \cdots & \vec{q}_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & \vec{q}_1^{\,T} & — \\ & \vdots & \\ — & \vec{q}_n^{\,T} & — \end{bmatrix} \vec{b}$$

$$= \vec{q}_1 \vec{q}_1^{\,T} \vec{b} + \vec{q}_2 \vec{q}_2^{\,T} \vec{b} + \ldots + \vec{q}_n \vec{q}_n^{\,T} \vec{b}$$

$$= \vec{q}_1 (\vec{q}_1 \cdot \vec{b}) + \vec{q}_2 (\vec{q}_2 \cdot \vec{b}) + \ldots + \vec{q}_n (\vec{q}_n \cdot \vec{b})$$

$$\rightarrow \quad \boxed{P\vec{b} = \text{proj}_{\vec{q}_1}(\vec{b}) + \text{proj}_{\vec{q}_2}(\vec{b}) + \ldots + \text{proj}_{\vec{q}_n}(\vec{b})}$$

where $\text{proj}_{\vec{q}_i}(\vec{b})$ is projection of $\vec{b}$ onto the line $\text{span}(\vec{q}_i)$.

(Recall: $\text{proj}_{\vec{q}_i}(\vec{b}) = \dfrac{\vec{q}_i \vec{q}_i^{\,T} \vec{b}}{\vec{q}_i^{\,T} \vec{q}_i}$, but denominator is 1 since $\vec{q}_i \cdot \vec{q}_i = 1$)

Example Project $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto the subspace

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ in } \mathbb{R}^3 \;\middle|\; x+y+z=0 \right\}$$

$\vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\vec{q}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$ Can check $\vec{q}_1, \vec{q}_2$ is an orthormal basis of W

$$\text{proj}_{\vec{q}_1}(\vec{b}) = (\vec{q}_1 \cdot \vec{b})\, \vec{q}_1 = \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \vec{q}_1 = \frac{1}{\sqrt{2}} \vec{q}_1$$

$$\text{proj}_{\vec{q}_2}(\vec{b}) = (\vec{q}_2 \cdot \vec{b})\,\vec{q}_2 = \left(\frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix} \cdot \begin{bmatrix}1\\0\\0\end{bmatrix}\right)\vec{q}_2 = \frac{1}{\sqrt{6}}\vec{q}_2$$

Thus $\text{proj}_W(\vec{b}) = \text{proj}_{\vec{q}_1}(\vec{b}) + \text{proj}_{\vec{q}_2}(\vec{b})$

$$= \frac{1}{\sqrt{2}}\vec{q}_1 + \frac{1}{\sqrt{6}}\vec{q}_2 = \frac{1}{2}\begin{bmatrix}1\\-1\\0\end{bmatrix} + \frac{1}{6}\begin{bmatrix}1\\1\\-2\end{bmatrix}$$

$$= \begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}$$

Takeaway: theory of projections is very straightforward if we use orthornormal bases.

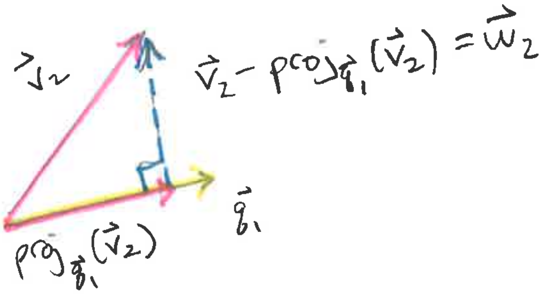
So we are led to the following:

Given a basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ of a subspace can we turn it into an orthonormal basis?

Let's try. Let $\vec{w}_1 = \vec{v}_1$. Make it a unit vector:

$$\vec{q}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|}$$

Now we want $\vec{w}_2$ orthogonal to $\vec{q}_1$.



We take

$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{q}_1}(\vec{v}_2)$$

$$= \vec{v}_2 - \vec{q}_1\vec{q}_1^{\,T}\vec{v}_2$$

Now $\vec{w}_2 \perp \vec{q}_1$ but $\vec{w}_2$ may not be unit length.

So we take $\vec{q}_2 = \dfrac{w_2}{\|\vec{w}_2\|}$. We then continue this

process. The result is the following.

Gram-Schmidt Orthogonalization

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ be independent vectors spanning $W \subset \mathbb{R}^m$. Then we can obtain an orthonormal basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n$ of W as follows:

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{q}_1 = \vec{w}_1 / \|\vec{w}_1\|$$

$$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{q}_1}(\vec{v}_2) \qquad \vec{q}_2 = \vec{w}_2 / \|\vec{w}_2\|$$

$$\vec{w}_3 = \vec{v}_3 - \text{proj}_{\vec{q}_1}(\vec{v}_3) - \text{proj}_{\vec{q}_2}(\vec{v}_3) \qquad \vec{q}_3 = \vec{w}_3 / \|\vec{w}_3\|$$

$$\vdots \qquad \vdots$$

$$\vec{w}_n = \vec{v}_n - \text{proj}_{\vec{q}_1}(\vec{v}_n) - \dots - \text{proj}_{\vec{q}_{n-1}}(\vec{v}_n) \qquad \vec{q}_n = \vec{w}_n / \|\vec{w}_n\|$$

Example

Let $\vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$

Find the orthonormal basis associated to these vectors using Gram-Schmidt.

$\vec{w}_1 = \vec{v}_1 \qquad \vec{q}_1 = \dfrac{\vec{w}_1}{\|\vec{w}_1\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ (here $\|\vec{w}_1\| = 1$)

$\vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{q}_1}(\vec{v}_2) = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\,\vec{q}_1 = \vec{v}_2 - 2\vec{q}_1$

$\left(\vec{q}_1 \cdot \vec{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = 2 \right) \qquad = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

and $\vec{q}_2 = \dfrac{\vec{w}_2}{\|\vec{w}_2\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ (again, $\|\vec{w}_2\| = 1$)

Finally, $\vec{w}_3 = \vec{v}_3 - \text{proj}_{\vec{q}_1}(\vec{v}_3) - \text{proj}_{\vec{q}_2}(\vec{v}_3)$

$= \vec{v}_3 - (\vec{q}_1 \cdot \vec{v}_3)\,\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_3)\,\vec{q}_2$

$= \vec{v}_3 - (-1)\,\vec{q}_1 - (0)\,\vec{q}_2$

$= \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$

and $\vec{q}_3 = \dfrac{\vec{w}_3}{\|\vec{w}_3\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$

Thus we obtain $Q = \begin{bmatrix} | & | & | \\ \vec{q}_1 & \vec{q}_2 & \vec{q}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$