## Review

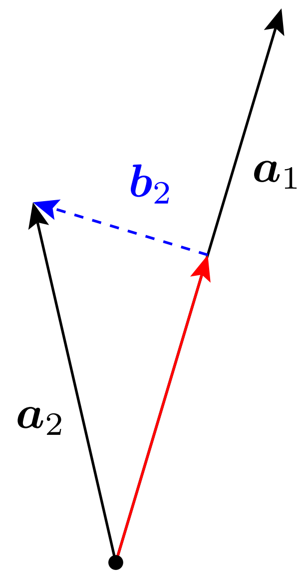
- Vectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are orthonormal if

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

- **Gram–Schmidt** orthonormalization:

Input: basis $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ for $V$.

Output: orthonormal basis $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ for $V$.




$$\begin{aligned} \boldsymbol{b}_1 &= \boldsymbol{a}_1, & \boldsymbol{q}_1 &= \frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|} \\ \boldsymbol{b}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1, & \boldsymbol{q}_2 &= \frac{\boldsymbol{b}_2}{\|\boldsymbol{b}_2\|} \\ \boldsymbol{b}_3 &= \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \langle \boldsymbol{a}_3, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2, & \boldsymbol{q}_3 &= \frac{\boldsymbol{b}_3}{\|\boldsymbol{b}_3\|} \\ &\vdots \end{aligned}$$

**Example 1.** Apply Gram–Schmidt to the vectors $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

**Solution.**

$$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \qquad \boldsymbol{q}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right\rangle \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \qquad \boldsymbol{q}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

$$\boldsymbol{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{q}_1 \right\rangle \boldsymbol{q}_1 - \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{q}_2 \right\rangle \boldsymbol{q}_2 = \ldots = \frac{1}{9}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \qquad \boldsymbol{q}_3 = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

We obtained the orthonormal vectors $\frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$.

**Theorem 2.** The columns of an $m \times n$ matrix $Q$ are orthonormal
$\Longleftrightarrow \quad Q^T Q = I$ (the $n \times n$ identity)

**Proof.** Let $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ be the columns of $Q$.

They are orthonormal if and only if $\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$

Armin Straub
astraub@illinois.edu

All these inner products are packaged in $Q^TQ = I$:

$$\begin{bmatrix} — & \boldsymbol{q}_1^T & — \\ — & \boldsymbol{q}_2^T & — \\ & \vdots & \end{bmatrix} \begin{bmatrix} | & | & \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{bmatrix}$$

□

**Definition 3.** An **orthogonal matrix** is a square matrix $Q$ with orthonormal columns.

It is historical convention to restrict to square matrices, and to say orthogonal matrix even though "orthonormal matrix" might be better.

An $n \times n$ matrix $Q$ is orthogonal $\iff$ $Q^TQ = I$

In other words, $Q^{-1} = Q^T$.

**Example 4.** $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ is orthogonal.

In general, all permutation matrices $P$ are orthogonal.

Why? Because their columns are a permutation of the standard basis.

And so we always have $P^TP = I$.

**Example 5.** $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

$Q$ is orthogonal because:

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$

  Just to make sure: why length 1? Because $\left\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$.

- Alternatively: $Q^TQ = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

**Example 6.** Is $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ orthogonal?

No, the columns are orthogonal but not normalized.

But $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is an orthogonal matrix.

Just for fun: a $n \times n$ matrix with entries $\pm 1$ whose columns are orthogonal is called a *Hadamard matrix* of size $n$.

A size 4 example: $\begin{bmatrix} H & H \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Continuing this construction, we get examples of size $8, 16, 32, \ldots$

It is believed that Hadamard matrices exist for all sizes $4n$.

But no example of size 668 is known yet.

Armin Straub
astraub@illinois.edu

## The QR decomposition (flashed at you)

- Gaussian elimination in terms of matrices: $A = LU$
- Gram–Schmidt in terms of matrices: $A = QR$

Let $A$ be an $m \times n$ matrix of rank $n$. (columns independent)

Then we have the **QR decomposition** $A = QR$,

- where $Q$ is $m \times n$ and has orthonormal columns, and
- $R$ is upper triangular, $n \times n$ and invertible.

Idea: Gram–Schmidt on the columns of $A$, to get the columns of $Q$.

**Example 7.** Find the QR decomposition of $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$.

**Solution.** We apply Gram–Schmidt to the columns of $A$:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \boldsymbol{q}_1$$

$$\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}, \boldsymbol{q}_1 \right\rangle \boldsymbol{q}_1 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \boldsymbol{q}_2$$

$$\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \left\langle \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \boldsymbol{q}_1 \right\rangle \boldsymbol{q}_1 - \left\langle \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \boldsymbol{q}_2 \right\rangle \boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \boldsymbol{q}_3$$

Hence: $Q = [\ \boldsymbol{q}_1 \ \ \boldsymbol{q}_2 \ \ \boldsymbol{q}_3\ ] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

To find $R$ in $A = QR$,

note that $Q^T A = Q^T Q R = R$.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}$$

Summarizing, we have

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}.$$

**Recipe**. In general, to obtain $A = QR$:

- Gram–Schmidt on (columns of) $A$, to get (columns of) $Q$.
- Then, $R = Q^T A$.

The resulting $R$ is indeed upper triangular, and we get:

$$\begin{bmatrix} | & | & \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} | & | & \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots \\ | & | & \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1^T\boldsymbol{a}_1 & \boldsymbol{q}_1^T\boldsymbol{a}_2 & \boldsymbol{q}_1^T\boldsymbol{a}_3 & \cdots \\ & \boldsymbol{q}_2^T\boldsymbol{a}_2 & \boldsymbol{q}_2^T\boldsymbol{a}_3 & \\ & & \boldsymbol{q}_3^T\boldsymbol{a}_3 & \\ & & & \ddots \end{bmatrix}$$

It should be noted that, actually, no extra work is needed for computing $R$: all the inner products in $R$ have been computed during Gram–Schmidt.

(Just like the LU decomposition encodes the steps of Gaussian elimination, the QR decomposition encodes the steps of Gram–Schmidt.)

## Practice problems

**Example 8.** Complete $\frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ to an orthonormal basis of $\mathbb{R}^3$.

(a) by using the FTLA to determine the orthogonal complement of the span you already have

(b) by using Gram–Schmidt after throwing in an independent vector such as $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

**Example 9.** Find the QR decomposition of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Armin Straub
astraub@illinois.edu