# The QR factorization

**Nicholas Hu** · Last updated on 2025-03-14

Let $A \in \mathbb{C}^{m \times n}$. The **QR factorization** is a factorization of $A$ as $QR$, where $Q \in \mathbb{C}^{m \times m}$ is unitary and $R \in \mathbb{C}^{m \times n}$ is (rectangular) upper triangular.[1] We will show below that such a factorization always exists by describing three different methods to compute it.

When $A$ has full column rank, we have $a_j = \sum_{i \le j} r_{ij} q_i$ for each $j$, so $\operatorname{span}\{a_j\}_{j \le k} \subseteq \operatorname{span}\{q_j\}_{j \le k}$ for each $k$. As these subspaces are both $k$-dimensional, they must be equal, which also implies that the diagonal entries of $R$ are nonzero. Moreover, if $\hat{Q}$ denotes the left $m \times n$ submatrix of $Q$ and $\hat{R}$ denotes the upper $n \times n$ submatrix of $R$, we have the **thin/reduced QR factorization** $A = \hat{Q}\hat{R}$.

The thin QR factorization of a full column rank matrix is nearly unique in the sense that if $A = \tilde{Q}\tilde{R}$ for some $\tilde{Q} \in \mathbb{C}^{m \times n}$ with orthonormal columns and some upper triangular $\tilde{R} \in \mathbb{C}^{n \times n}$, then $\tilde{Q} = \hat{Q}D$ and $\hat{R} = D\tilde{R}$ for some diagonal matrix $D$ whose diagonal entries have unit modulus. This follows from the observation that $D := \hat{Q}^*\tilde{Q} = \hat{R}\tilde{R}^{-1} = \hat{R}^{-*}\tilde{R}^*$ must be both upper and lower triangular. Thus, if we specify a (complex) sign for each diagonal entry of $\hat{R}$, the factorization is unique.

## Gram–Schmidt orthogonalization

Suppose that $(a_j)_{j \ge 1}$ is a sequence of vectors in a Hilbert space $V$. **Gram–Schmidt orthogonalization** defines an *orthogonal* sequence of vectors $(b_j)_{j \ge 1}$ in $V$ such that $\mathcal{A}_k := \operatorname{span}\{a_j\}_{j \le k} = \mathcal{B}_k := \operatorname{span}\{b_j\}_{j \le k}$ for each $k$. To wit, let $\operatorname{proj}_b := \operatorname{proj}_{\operatorname{span}\{b\}}$ for $b \in H$; that is,

$$\operatorname{proj}_b a = \begin{cases} \frac{\langle a, b \rangle}{\langle b, b \rangle} b & \text{if } b \ne 0, \\ b & \text{if } b = 0. \end{cases}$$

We then inductively define

$$b_j := a_j - \sum_{i<j} \operatorname{proj}_{b_i} a_j.$$

Assuming that $\{b_j\}_{j<k}$ is orthogonal for a given $k$, we then have $\langle b_k, b_j \rangle = \langle a_k - \sum_{i<k} \operatorname{proj}_{b_i} a_k, b_j \rangle = \langle a_k - \operatorname{proj}_{b_j} a_k, b_j \rangle = 0$ for all $j < k$, which shows that $\{b_j\}_{j \le k}$ is orthogonal. Moreover, if $\mathcal{A}_{k-1} = \mathcal{B}_{k-1}$, then $b_k \in a_k - \mathcal{B}_{k-1} = a_k - \mathcal{A}_{k-1} \subseteq \mathcal{A}_k$ and $a_k \in b_k + \mathcal{B}_{k-1} \subseteq \mathcal{B}_k$, so $\mathcal{A}_k = \mathcal{B}_k$.

To compute a QR factorization of $A$, we can apply Gram–Schmidt orthogonalization to the columns of $A =: [a_1 \quad \cdots \quad a_n]$ as follows. For each $j \le m$, we inductively define $b_j := a_j - \sum_{i<j} \operatorname{proj}_{q_i} a_j$ if $j \le n$ *and* the right-hand expression is *nonzero*; otherwise, we select an *arbitrary nonzero* $b_j \in \mathcal{B}_{j-1}^{\perp}$. In either case, we then define $q_j := \frac{b_j}{\|b_j\|}$. We thereby obtain an orthonormal basis $\{q_j\}_{j \le m}$ of $\mathbb{C}^m$ such that $a_j = \sum_{i \le \min\{j,\, m\}} r_{ij} q_i$ for some $r_{ij} \in \mathbb{C}$, as required.

## Modified Gram–Schmidt orthogonalization

In Gram–Schmidt orthogonalization, we define $b_j = (I - \sum_{i<j} \operatorname{proj}_{b_i})a_j$. Since the $b_i$ are orthogonal, this can equivalently be written as $b_j = (I - \operatorname{proj}_{b_{j-1}}) \cdots (I - \operatorname{proj}_{b_2})(I - \operatorname{proj}_{b_1})a_j$, so computationally speaking, the projection operator $I - \operatorname{proj}_{b_i}$ can be applied to all $a_j$ with $i < j$ (assuming there are finitely many of them) as soon as $b_i$ is generated. The resulting algorithm is known as **modified Gram–Schmidt orthogonalization** and exhibits greater numerical stability than "classical" Gram–Schmidt orthogonalization.

# Householder reflections

Suppose that $v$ is a nonzero vector in a Hilbert space $V$. The **reflection operator** across the hyperplane $\{v\}^\perp$ is defined for $x \in H$ by

$$\operatorname{refl}_v x := (I - 2 \operatorname{proj}_v)\, x = x - \frac{2\langle x, v\rangle}{\langle v, v\rangle} v.$$

Since $\operatorname{proj}_v$ is idempotent and self-adjoint, $\operatorname{refl}_v$ is involutory and self-adjoint and therefore unitary.

A **Householder reflection** is a reflection operator $H : \mathbb{C}^d \to \mathbb{C}^d$ that zeroes out all components of some vector $x$ except for its first component $x_1$; we assume that the other components are not already all zeroes. In other words, $Hx = \alpha e_1$ for some $\alpha \in \mathbb{C}$, where $e_1 := \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top$ and $x \notin \operatorname{span}\{e_1\}$.

As $H$ is unitary and self-adjoint, we must have $|\alpha| = \|x\|$ and $\langle Hx, x\rangle = \alpha\overline{x_1} \in \mathbb{R}$, which implies that $\alpha = \pm \operatorname{sign}(x_1)\|x\|$ (unless $x_1 = 0$, in which case the only constraint is $|\alpha| = \|x\|$). Since $\operatorname{refl}_w x = \alpha e_1$ if and only if $\frac{2\langle x,w\rangle}{\langle w,w\rangle} w = x - \alpha e_1$, using the **Householder vector** $v := x - \alpha e_1$ guarantees that $H := \operatorname{refl}_v$ satisfies $Hx = \alpha e_1$. A conventional choice of $\alpha$ in this context is $\alpha = -\operatorname{sign}(x_1)\|x\|$ so as to maximize $\|v\|^2 = 2(\|x\|^2 \mp |x_1|\|x\|)$ for the sake of numerical stability.

To compute a QR factorization of $A$, we can apply Householder reflections successively to introduce zeroes below the diagonal in each column of $A$. More precisely, we can find a Householder reflection $H \in \mathbb{C}^{m\times m}$ such that

$$HA = \begin{bmatrix} \alpha & b^\top \\ & A' \end{bmatrix},$$

where $\alpha \in \mathbb{C}$, $b \in \mathbb{C}^{n-1}$, and $A' \in \mathbb{C}^{(m-1)\times(n-1)}$ (allowing $H = I$ if the subdiagonal entries in the first column of $A$ are already zero). Now supposing inductively that $A'$ has a QR factorization $Q'R'$, we obtain the factorization

$$A = \underbrace{H^* \begin{bmatrix} 1 & \\ & Q' \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \alpha & b^\top \\ & R' \end{bmatrix}}_{R}.$$

# Givens rotations

Given $a, b \in \mathbb{C}$, consider the problem of finding a $U \in \mathrm{SU}(2)$ and an $r \in \mathbb{C}$ such that $U \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$. We have

$$U = \begin{bmatrix} c & s \\ -\overline{s} & \overline{c} \end{bmatrix}, \quad \text{where } |c|^2 + |s|^2 = 1$$

and $ac + bs = r$, $b\overline{c} - a\overline{s} = 0$. Since $U$ is unitary, we must have $r = \omega\sqrt{|a|^2 + |b|^2}$ for some $\omega \in \mathbb{C}$ with $|\omega| = 1$, and assuming that $r \neq 0$ (which is to say that $a$ and $b$ are not both zero), we obtain

$$c = \frac{\overline{a}}{\overline{r}} = \frac{\omega\overline{a}}{\sqrt{|a|^2 + |b|^2}}, \quad s = \frac{\overline{b}}{\overline{r}} = \frac{\omega\overline{b}}{\sqrt{|a|^2 + |b|^2}}.$$

A conventional choice in this context is $\omega = \mathbf{sign}(a)$, along with $U = I$ (and $r = 0$) in the case $a = b = 0$.

Thus, if $a$ and $b$ are the $i^{\text{th}}$ and $j^{\text{th}}$ components of some $x \in \mathbb{C}^m$, where $i < j$, the **Givens rotation**

$$G := \begin{bmatrix} I_{i-1} & & & & \\ & c & & s & \\ & & I_{(j-1)-i} & & \\ & -\overline{s} & & \overline{c} & \\ & & & & I_{m-j} \end{bmatrix}$$

is a unitary matrix such that the $j^{\text{th}}$ component of $Gx$ is zero. (In the real-valued setting, $G$ is indeed a rotation in the $x_i$-$x_j$ plane.) Such rotations can evidently be applied to compute a QR factorization of $A$ by introducing zeroes below the diagonal of $A$ one at a time.

---

1. If $A \in \mathbb{R}^{m \times n}$, a QR factorization is defined analogously; i.e., with $Q$ orthogonal. ↩