Given a matrix $A \in \mathbb{C}^{m \times n}$, we can have the following decomposition

$$A = QR$$

where $Q \in \mathbb{C}^{m \times m}$ is unitary, and $R \in \mathbb{C}^{m \times n}$ is upper triangular.
Classical Gram-Schmidt

For simplicity, assume for the time being that $A$ has full column rank.

Set

$$q_1 = \frac{a_1}{\|a_1\|_2}, \quad r_{11} = \|a_1\|_2$$

Consider finding $q_2$. Since

$$a_2 = r_{12}q_1 + r_{22}q_2,$$

and $q_1^Hq_2 = 0$, we have that

$$r_{12} = q_1^Ha_2.$$ 

Subsequently,

$$q_2 = \frac{y_2}{\|y_2\|_2}, \quad r_{22} = \|y_2\|_2$$

where

$$y_2 = a_2 - r_{12}q_1$$

Repeating the above steps, we have, for $k = 2, \ldots, n,$

$$r_{ik} = q_k^Ha_k, \quad i = 1, \ldots, k - 1$$

$$y_k = a_k - \sum_{i=1}^{k-1} r_{ik}q_i$$

$$q_k = \frac{y_k}{\|y_k\|_2}, \quad r_{kk} = \|y_k\|_2$$
To complete the above Gram-Schmidt (GS) procedure, we need
\[ y_k \neq 0, \quad \forall k \]
for all \( k \). This is true when \( A \) is of full column rank.

The GS procedure leads to a thin QR decomposition
\[ A = Q_1 R_1 \]
where \( Q_1 \in \mathbb{C}^{m \times n} \) has orthonormal columns, and \( R_1 \in \mathbb{C}^{n \times n} \) is upper triangular.

We can always find a matrix \( Q_2 \in \mathbb{C}^{(m-n) \times n} \) so that
\[ Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \]
is unitary.

We can then form a complete QR decomposition
\[ A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \]
**Theorem 9.1** If $A$ has full column rank, then the decomposition

$$A = Q_1R_1$$

where $R_1 \in \mathbb{C}^{n \times n}$ is upper triangular with positive diagonal entries, $Q_1 \in \mathbb{C}^{m \times n}$ is semi-unitary, is unique.

**Proof:** Since

$$A^H A = R_1^H R_1,$$

$R_1^H$ is also the Cholesky factor which is unique (cf., Theorem 7.3). It follows that $Q_1 = AR_1^{-1}$ is also unique.

---

Some properties resulting from the GS procedure:

**Property 9.1** For full column rank $A$,

$$\text{span}\{a_1, \ldots, a_k\} = \text{span}\{q_1, \ldots, q_k\}$$

for $k = 1, \ldots, n$.

**Property 9.2** For full column rank $A$,

$$R(A) = R(Q_1), \quad R_\perp(A) = R(Q_2)$$

**Property 9.3** For full column rank $A$, $R_1$ is nonsingular.
For rank deficient \( A \), a QR pair can be constructed. In the GS procedure, do the following the modification

\[ q_k = 0, \quad r_{kk} = 0 \quad \text{if} \quad y_k = 0. \]

for \( k = 1, \ldots, n. \)

Let \( Q'_1 \in \mathbb{C}^{m \times r} \) be a matrix containing the nonzero \( q_k \)'s. Let \( Q'_2 = [ q'_{r+1}, \ldots, q'_n ] \) be a matrix so that \( [ Q'_1 \ Q'_2 ] \) is unitary.

Replace the first zero column vector \( q_k \) by \( q'_{r+1} \), the 2nd \( q_k \) by \( q'_{r+2} \), and so on. A QR decomposition will then be obtained.

**Application: Full column rank LS**

Consider again the LS problem

\[
\min_{x \in \mathbb{C}^n} \|Ax - b\|_2^2
\]

where \( A \) is of full column rank.
Let

\[ Q^H b = \begin{bmatrix} Q_1^H \\ Q_2^H \end{bmatrix} b = \begin{bmatrix} c \\ d \end{bmatrix} \]

We have that

\[ \|Ax - b\|^2_2 = \|Q^H(Ax - b)\|^2_2 \]

\[ = \left\| \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x - \begin{bmatrix} c \\ d \end{bmatrix} \right\|^2_2 \]

\[ = \|R_1 x - c\|^2_2 + \|d\|^2_2 \]

Hence, the LS solution is

\[ x_{LS} = R_1^{-1} c \]

In practice, the classical GS procedure is numerically unstable.

There are several other ways of finding QR factors. In this course we consider

- modified GS
- Householder transform
- Givens rotations.
Modified GS

Partition

\[ R_1 = \begin{bmatrix} r_1^T \\ \vdots \\ r_n^T \end{bmatrix} \]

The decomposition of \( A \) can be expressed as

\[ A = QR = \sum_{i=1}^{n} q_i r_i^T \]

At stage \( k \), consider

\[ A - \sum_{i=1}^{k-1} q_i r_i^T = \sum_{i=k}^{n} q_i r_i^T \]

\[ = \begin{bmatrix} 0 & \cdots & 0 & r_{kk} q_k \end{bmatrix}, \ r_{k+1,k} q_k + r_{kk} q_k, \cdots \]

Thus,

\[ y_k = (A - \sum_{i=1}^{k-1} q_i r_i^T) e_k \]

\[ q_k = y_k / \|y_k\|_2 \]

\[ r_k^T = q_k^H (A - \sum_{i=1}^{k-1} q_i r_i^T) \]

Unlike the classical GS, the modified GS has been found to be numerically very stable.
Householder QR

It is easier to look at the real-value case; the idea in the complex case is more or less the same.

To see how the Householder QR algorithm works, it is necessary to understand the concepts of reflection matrices and Householder transformation.

A matrix is called a reflection matrix if it is given by

\[ H = I - 2P \]

where \( P \in \mathbb{R}^{m \times m} \) is an orthogonal projection matrix.

A reflection matrix is symmetric, and orthogonal.

Let \( x = Px + P_\perp x \). We have that

\[ Hx = -Px + P_\perp x \]

where the component \( Px \) is reflected.
**Householder transformation:**

The Householder transformation is to find a matrix, denoted by $H$ so that

$$Hx = \beta e_1,$$

i.e., the elements of $Hx$ is eliminated except for the 1st.

Let

$$P = \nu (\nu^T \nu)^{-1} \nu^T$$

where $\nu$ is to be determined.

Now,

$$Hx = x - \frac{2\nu^T x}{\nu^T \nu} \nu \in \text{span}\{x, \nu\}$$

If $Hx = \beta e_1$ is what we want, then

$$\nu = x + \alpha e_1$$

for some coefficient $\alpha$.

If we set $\alpha = \|x\|_2$, then it can be verified that

$$Hx = -\|x\|_2 e_1$$

Note that a similar effect occurs when $\alpha = -\|x\|_2$.
**Householder QR:**

Let $H_1 \in \mathbb{R}^{m \times m}$ be the Householder transformation for $a_1$. Then,

$$A_1 = H_1 A$$

$$= \begin{bmatrix} x & x & \ldots & x \\ 0 & x & \ldots & x \\ \vdots & x & \ldots & x \\ 0 & 0 & \ldots & x \end{bmatrix}$$

Let

$$H_2 = \begin{bmatrix} 1 & 0 \\ 0 & H_2 \tilde{H} \end{bmatrix}$$

where $H_2 \tilde{H} \in \mathbb{R}^{m-1 \times m-1}$ is the Householder transform for $A_1(2: m, 2)$. Then,

$$A_2 = H_2 A_1$$

$$= \begin{bmatrix} x & x & x & \ldots & x \\ 0 & x & x & \ldots & x \\ \vdots & 0 & x & \ldots & x \\ \vdots & \vdots & x & \ldots & x \\ 0 & 0 & 0 & \ldots & x \end{bmatrix}$$
It follows that at stage $k$,

$$A_k = H_k A_{k-1}$$

where

$$H_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \tilde{H}_k \end{bmatrix},$$

and $\tilde{H}_k \in \mathbb{R}^{(m-k+1) \times (m-k+1)}$ is the Householder transformation for $A_{k-1}(k : m, k)$.

The process continues for $n - 1$ stages. We then obtain

$$A_{n-1} = H_{n-1} \ldots H_2 H_1 A$$

The matrix $A_{n-1} \triangleq R$ is upper triangular.

Since each $H_k$ is orthogonal, $\prod_{i=1}^{n-1} H_i \triangleq Q$ is also orthogonal.
Givens Rotations

Again, in the Givens QR decomposition, it is instructive to look at the real case.

Let

\[ J_{ik}(\theta) = \begin{bmatrix} I & c(\theta) & s(\theta) \\ -s(\theta) & I & c(\theta) \end{bmatrix} \]

where \( c(\theta) = \cos(\theta) \), \( s(\theta) = \sin(\theta) \), the \((i, k)\) entry is \(-s(\theta)\), and so on.

A Givens rotation \( J_{ik}(\theta) \) is orthogonal.

For example,

\[
J_{21}(\theta)J_{21}^T(\theta) = \begin{bmatrix} c(\theta) & s(\theta) \\ -s(\theta) & c(\theta) \end{bmatrix} \begin{bmatrix} c(\theta) & -s(\theta) \\ s(\theta) & c(\theta) \end{bmatrix} = \begin{bmatrix} c^2(\theta) + s^2(\theta) & -c(\theta)s(\theta) + c(\theta)s(\theta) \\ -c(\theta)s(\theta) + c(\theta)s(\theta) & c^2(\theta) + s^2(\theta) \end{bmatrix} = \begin{bmatrix} I \\ I \end{bmatrix}
\]
Consider $J_{21}(\theta)A$:

\[
J_{21}(\theta)A = \begin{bmatrix}
c(\theta) & s(\theta) \\
-s(\theta) & c(\theta)
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} & \cdots \\
0 & a_{22} & \cdots \\
\vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
\times & \times & \cdots \\
-s(\theta)a_{11} + c(\theta)a_{21} & \times & \cdots \\
\vdots & \vdots & \ddots
\end{bmatrix}
\]

Set $\theta = \tan^{-1}(a_{21}/a_{11})$. Then

\[-s(\theta)a_{11} + c(\theta)a_{21} = 0.
\]

For notational simplicity, let $J_{ik} = J_{ik}(\theta)$ where $\theta$ is chosen to annihilate the $(i, k)$ entry of the transformed matrix.

By performing a sequence of Givens rotations to annihilate the lower triangular parts of $A$, we obtain

\[
\underbrace{J_{n,n-1} \cdots (J_{n,2} \cdots J_{32})(J_{n1} \cdots J_{21})}_{Q^T} A = R
\]

which is upper triangular.