

---

**Linear Analysis**  
**Lecture 16**

---

# The QR Factorization

Recall the Gram-Schmidt orthogonalization process. Let  $V$  be an inner product space, and suppose  $a_1, \dots, a_n \in V$  are linearly independent. Define  $q_1, \dots, q_n$  inductively, as follows: set

$$\begin{aligned} p_1 &= a_1, & q_1 &= p_1 / \|p_1\|, \\ p_j &= a_j - \sum_{i=1}^{j-1} \langle a_j, q_i \rangle q_i & \text{for } 2 \leq j \leq n, & \quad \text{and} \\ q_j &= p_j / \|p_j\|. \end{aligned}$$

For  $1 \leq j \leq n$ ,

$$q_j \in \text{Span}\{a_1, \dots, a_j\},$$

so each  $p_j \neq 0$  by the lin. indep. of  $\{a_1, \dots, a_n\}$ . Thus each  $q_j$  is well-defined.

We have  $\{q_1, \dots, q_n\}$  is an orthonormal basis for  $\text{Span}\{a_1, \dots, a_n\}$ . Also

$$a_k \in \text{Span}\{q_1, \dots, q_k\} \quad 1 \leq k \leq n,$$

so  $\{q_1, \dots, q_k\}$  is an orthonormal basis of  $\text{Span}\{a_1, \dots, a_k\}$ .

# The QR Factorization

Define

$$r_{jj} = \|p_j\| \quad \text{and} \quad r_{ij} = \langle a_j, q_i \rangle \quad \text{for} \quad 1 \leq i < j \leq n,$$

we have:

$$\begin{aligned} a_1 &= r_{11} q_1, \\ a_2 &= r_{12} q_1 + r_{22} q_2, \\ a_3 &= r_{13} q_1 + r_{23} q_2 + r_{33} q_3, \\ &\vdots \\ a_n &= \sum_{i=1}^n r_{in} q_i. \end{aligned}$$

Set

$$A = [a_1 \ a_2 \ \dots \ a_n], \quad R = [r_{ij}], \quad \text{and} \quad Q = [q_1 \ q_2 \ \dots \ q_n],$$

where  $r_{ij} = 0$ ,  $i > j$ . Then  $A = QR$ , where  $Q$  is unitary and  $R$  is upper triangular.

## The QR Factorization: Remarks

- (1) If  $a_1, a_2, \dots$  is a lin. indep. sequence, apply Gram-Schmidt to obtain an orthonormal sequence  $q_1, q_2, \dots$  such that  $\{q_1, \dots, q_k\}$  is an orthonormal basis for  $\text{Span}\{a_1, \dots, a_k\}$ ,  $k \geq 1$ .
- (2) If the  $a_j$ 's are lin. dep., for some value(s) of  $k$ ,

$$a_k \in \text{Span}\{a_1, \dots, a_{k-1}\}, \quad \text{so} \quad p_k = 0.$$

The process can be modified by setting  $r_{kk} = 0$ , not defining a new  $q_k$  for this iteration and then proceeding as usual. We end up with orthogonal  $q_j$ 's. Then for  $k \geq 1$ , the vectors  $\{q_1, \dots, q_k\}$  form an orthonormal basis for  $\text{Span}\{a_1, \dots, a_{\ell+k}\}$  where  $\ell$  is the number of  $r_{jj} = 0$ . Again we obtain  $A = QR$ , but now  $Q$  may not be square.

# The QR Factorization: Remarks

- (3) The classical Gram-Schmidt algorithm has poor computational performance due to the accumulation of round-off error. The computed  $q_j$ 's are not orthogonal:  $\langle q_j, q_k \rangle$  is small for  $j \neq k$  and  $|k - j|$  small, but not for  $|k - j|$  big.

## Classic Gram-Schmidt

```
For  $j = 1, \dots, n$  do
   $p := a_j$ 
  For  $i = 1, \dots, j - 1$  do
     $r_{ij} = \langle a_j, q_i \rangle$ 
     $p := p - r_{ij}q_i$ 
   $r_{jj} := \|p\|$ 
   $q_j := p/r_{jj}$ 
```

## Modified Gram-Schmidt

```
For  $j = 1, \dots, n$  do
   $p := a_j$ 
  For  $i = 1, \dots, j - 1$  do
     $r_{ij} = \langle p, q_i \rangle$ 
     $p := p - r_{ij}q_i$ 
   $r_{jj} = \|p\|$ 
   $q_j := p/r_{jj}$ 
```

The only difference is in the computation of  $r_{ij}$ , we orthogonalize the accumulated partial sum for  $p_j$  against each  $q_i$  successively.

# The QR Factorization

**Proposition.** Suppose  $A \in \mathbb{C}^{m \times n}$  with  $m \geq n$ . Then

$$\exists \text{ unitary } Q \in \mathbb{C}^{m \times m} \quad \text{upper triangular } R \in \mathbb{C}^{m \times n}$$

for which  $A = QR$ . If  $\tilde{Q} \in \mathbb{C}^{m \times n}$  denotes the first  $n$  columns of  $Q$  and  $\tilde{R} \in \mathbb{C}^{n \times n}$  denotes the first  $n$  rows of  $R$ , then

$$A = QR = [\tilde{Q} \ *] \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} = \tilde{Q}\tilde{R}.$$

Moreover

- (a) We may choose an  $R$  to have nonnegative diagonal entries.
- (b) If  $A$  is of full rank, then we can choose  $R$  with positive diagonal entries, in which case the condensed factorization  $A = \tilde{Q}\tilde{R}$  is unique
- (c) If  $A$  is of full rank, the condensed factorization  $A = \tilde{Q}\tilde{R}$  is essentially unique: if  $A = \tilde{Q}_1\tilde{R}_1 = \tilde{Q}_2\tilde{R}_2$ , then there is a unitary diagonal matrix  $D \in \mathbb{C}^{n \times n}$  for which  $\tilde{Q}_2 = \tilde{Q}_1 D^H$ , rescaling the columns of  $\tilde{Q}_1$ , and  $\tilde{R}_2 = D\tilde{R}_1$ , rescaling the rows of  $\tilde{R}_1$ .

# The QR Factorization: Proof

If  $A$  has full rank, apply the Gram-Schmidt.  
Define

$$\tilde{Q} = [q_1, \dots, q_n] \in \mathbb{C}^{m \times n} \quad \text{and} \quad \tilde{R} = [r_{ij}] \in \mathbb{C}^{n \times n}$$

as above, so

$$A = \tilde{Q}\tilde{R}.$$

Extend  $\{q_1, \dots, q_n\}$  to an orthonormal basis  $\{q_1, \dots, q_m\}$  of  $\mathbb{C}^m$ , and set

$$Q = [q_1, \dots, q_m] \quad \text{and} \quad R = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} \in \mathbb{C}^{m \times n}, \quad \text{so } A = QR.$$

As  $r_{jj} > 0$  in the G-S process, we have (b).

Uniqueness follows by induction passing through the G-S process again, noting that at each step we have no choice.

(c) follows easily from (b).

## The QR Factorization: Remarks

- (1) If  $A \in \mathbb{R}^{m \times n}$ , everything can be done in real arithmetic, so, e.g.,  $Q \in \mathbb{R}^{m \times m}$  is orthogonal and  $R \in \mathbb{R}^{m \times n}$  is real, upper triangular.
- (2) In practice, there are more efficient and better computationally behaved ways of calculating the  $Q$  and  $R$  factors. The idea is to create zeros below the diagonal (successively in columns 1, 2, ...) as in Gaussian Elimination, except we now use Householder transformations (which are unitary) instead of the unit lower triangular matrices  $L_j$ .
- (3) A  $QR$  factorization is also possible when  $m < n$ .

$$A = Q[R_1 \ R_2] ,$$

where  $Q \in \mathbb{C}^{m \times m}$  is unitary and  $R_1 \in \mathbb{C}^{m \times m}$  is upper triangular.



# The QR Factorization

Every  $A \in \mathbb{C}^{m \times n}$  has a  $QR$ -factorization, even when  $m < n$ . Indeed, if

$$\text{rank}(A) = k,$$

there always exist

$Q \in \mathbb{C}^{m \times k}$  with orthonormal columns,

$R \in \mathbb{C}^{k \times n}$  full rank upper triangular,

and a permutation matrix  $P \in \mathbb{R}^{n \times n}$  such that

$$(*) \quad AP = QR.$$

Moreover, if  $A$  has rank  $n$  (so  $m \geq n$ ), then  $R \in \mathbb{C}^{n \times n}$  is nonsingular. On the other hand, if  $m < n$ , then

$$R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix},$$

where  $R_1 \in \mathbb{C}^{k \times k}$  is nonsingular.

Finally, if  $A \in \mathbb{R}^{m \times n}$ , then the same facts hold, but now both  $Q$  and  $R$  can be chosen to be real matrices.

# The QR-Factorization and Orthogonal Projections

Let  $A \in \mathbb{C}^{m \times n}$  have condensed QR-factorization

$$A = \tilde{Q}\tilde{R}.$$

Then by construction the columns of  $\tilde{Q}$  form an orthonormal basis for the range of  $A$ . Hence  $P = \tilde{Q}\tilde{Q}^H$  is the orthogonal projector onto the range of  $A$ .

Similarly, if the condensed QR-factorization of  $A^H$  is

$$A^H = \tilde{Q}_1\tilde{R}_1,$$

then  $P_1 = \tilde{Q}_1\tilde{Q}_1^H$  is the orthogonal projector onto  $\text{ran}(A^H) = \ker(A)^\perp$ , and so

$$I - \tilde{Q}_1\tilde{Q}_1^H$$

is the orthogonal projector onto  $\ker(A)$ .

The QR-factorization can be computed using either Givens rotations or Householder reflections. Although, the approach via rotations is arguably more stable numerically, it is more difficult to describe so we only illustrate the approach using Householder reflections.

## $QR$ using Householder Reflections

Recall that to any  $w \in \mathbb{C}^n$  we can associate the Householder reflection

$$U = I - 2 \frac{ww^*}{w^*w}$$

which reflects  $\mathbb{C}^n$  about the hyperplane  $\text{Span}\{w\}^\perp$ . Given a pair of non-zero vectors  $x$  and  $y$  with

$$\|x\|_2 = \|y\|_2, \quad \text{and} \quad x \neq y,$$

there is a Householder reflection such that  $y = Ux$ :

$$U = I - 2 \frac{(x - y)(x - y)^*}{(x - y)^*(x - y)}.$$

**Proof:**

$$\begin{aligned} Ux &= x - 2(x - y) \frac{\|x\|^2 - y^*x}{\|x\|^2 - 2y^*x + \|y\|^2} \\ &= x - 2(x - y) \frac{\|x\|^2 - y^*x}{2(\|x\|^2 - y^*x)} \\ &= y \end{aligned}$$

since  $\|x\| = \|y\|$ .

## QR using Householder Reflections

We now describe the basic *deflation* step in the QR-factorization. Suppose

$A_0 = \begin{bmatrix} \alpha_0 & a_0^T \\ b_0 & A_0 \end{bmatrix}$  and set  $\nu_0 = \left\| \begin{pmatrix} \alpha_0 \\ b_0 \end{pmatrix} \right\|_2$ . Let  $H_0$  be the Householder transformation that maps

$\begin{pmatrix} \alpha_0 \\ b_0^T \end{pmatrix} \mapsto \nu_0 e_1$ ;  $H_0 = I - 2 \frac{ww^T}{w^T w}$  where  $w = \begin{pmatrix} \alpha_0 \\ b_0 \end{pmatrix} - \nu_0 e_1 = \begin{pmatrix} \alpha_0 - \nu_0 \\ b_0 \end{pmatrix}$ .  
 Thus,  $H_0 A = \begin{bmatrix} \nu_0 & & \\ & a_1^T & \\ 0 & & A_1 \end{bmatrix}$ . A problem occurs if  $\nu_0 = 0$  or  $\begin{pmatrix} \alpha_0 \\ b_0 \end{pmatrix} = 0$ . In

this case, permute the offending column to the right bringing in the column of greatest magnitude. Now repeat with  $A_1$ .

If this method is implemented by always permuting the column of greatest magnitude into the current pivot column, then  $AP = QR$  gives a QR-factorization with the diagonal entries of  $R$  nonnegative and listed in the order of descending magnitude.

# QR for Solving Least Squares Problems

Suppose  $A \in \mathbb{C}^{m \times n}$ ,  $b \in \mathbb{C}^m$ , and  $m \geq n$  with  $\text{rank}(A) = n$ . Consider the least squares problem  $\min \|b - Ax\|^2$ . Let  $A = QR$  be a QR factorization of  $A$ , with condensed form  $Q_1 \tilde{R}$ , so that  $\tilde{R} \in \mathbb{C}^{n \times n}$  is nonsingular. Write  $Q = [Q_1 \ Q_2]$  where  $Q_2 \in \mathbb{C}^{m \times (m-n)}$ . Then

$$\begin{aligned}\|b - Ax\|^2 &= \|b - QRx\|^2 \\ &= \|Q^H b - Rx\|^2 \\ &= \left\| \begin{bmatrix} Q_1^H \\ Q_2^H \end{bmatrix} b - \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} x \right\|^2 \\ &= \left\| \begin{bmatrix} Q_1^H b - \tilde{R}x \\ Q_2^H b \end{bmatrix} \right\|^2 \\ &= \|Q_1^H b - \tilde{R}x\|^2 + \|Q_2^H b\|^2.\end{aligned}$$

Here  $\tilde{R} \in \mathbb{C}^{n \times n}$  is an invertible upper triangle matrix, so

$$x \text{ solves } \min_{x \in \mathbb{C}^n} \|b - Ax\|^2 \iff \tilde{R}x = Q_1^H b.$$

This system can be solved by back-substitution. Note that we only need  $Q_1$  and  $\tilde{R}$  to solve for  $x$ .

## QR for Solving Least Squares Problems

Next suppose  $A \in \mathbb{C}^{m \times n}$ ,  $b \in \mathbb{C}^m$ , and  $m < n$  with  $\text{rank}(A) = m$ , so that there are many solutions to  $Ax = b$ .

Consider the least squares problem

$$\min \|b - Ax\|^2 \quad .$$

In this case  $A$  is surjective, so there exists  $\bar{x} \in \mathbb{C}^n$  such that  $A\bar{x} = b$ . Let  $A^H = QR$  be a  $QR$  factorization of  $A^H$ , with condensed form

$$A = \tilde{Q}\tilde{R},$$

so that  $\tilde{R} \in \mathbb{C}^{n \times n}$  is nonsingular and  $\tilde{Q}^H\tilde{Q} = I_{m \times m}$ . Now solve

$$Ax = \tilde{R}^H\tilde{Q}^Hx = b.$$

To do this first solve

$$\tilde{R}^Hy = b$$

by forward substitution ( $\tilde{R}^H$  is nonsingular lower triangular), and set  $\bar{x} = \tilde{Q}y$ . Then

$$A\bar{x} = \tilde{R}^H\tilde{Q}^H\bar{x} = \tilde{R}^H\tilde{Q}^H\tilde{Q}y = \tilde{R}^Hy = b \quad .$$

Also,  $\bar{x}$  is the least norm solution to  $Ax = b$  ! (Why?)