

---

Numerical Analysis Hilary Term 2021  
Lecture 3: QR Factorization

---

**Definition:** a square real matrix  $Q$  is **orthogonal** if  $Q^T = Q^{-1}$ . This is true if, and only if,  $Q^T Q = I = Q Q^T$ .

**Example:** the permutation matrices  $P$  in LU factorization with partial pivoting are orthogonal.

**Proposition.** The product of orthogonal matrices is an orthogonal matrix.

**Proof.** If  $S$  and  $T$  are orthogonal,  $(ST)^T = T^T S^T$  so

$$(ST)^T(ST) = T^T S^T S T = T^T (S^T S) T = T^T T = I.$$

**Definition:** The **scalar (dot)(inner) product** of two vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

in  $\mathbb{R}^n$  is

$$x^T y = y^T x = \sum_{i=1}^n x_i y_i \in \mathbb{R}$$

**Definition:** Two vectors  $x, y \in \mathbb{R}^n$  are **orthogonal** if  $x^T y = 0$ . A set of vectors  $\{u_1, u_2, \dots, u_r\}$  is an **orthogonal set** if  $u_i^T u_j = 0$  for all  $i, j \in \{1, 2, \dots, r\}$  such that  $i \neq j$ .

**Lemma.** The columns of an orthogonal matrix  $Q$  form an orthogonal set, which is moreover an orthonormal basis for  $\mathbb{R}^n$ .

**Proof.** Suppose that  $Q = [q_1 \ q_2 \ \dots \ q_n]$ , i.e.,  $q_j$  is the  $j$ th column of  $Q$ . Then

$$Q^T Q = I = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} [q_1 \ q_2 \ \dots \ q_n] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Comparing the  $(i, j)$ th entries yields

$$q_i^T q_j = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

Note that the columns of an orthogonal matrix are of length 1 as  $q_i^T q_i = 1$ , so they form an orthonormal.

To see that it forms a basis, let  $x \in \mathbb{R}^n$  be any vector. One has  $x = QQ^T x = Qc$  where  $c = Q^T x$ , so  $x = \sum_{i=1}^n c_i q_i$ .

**Lemma.** If  $u \in \mathbb{R}^n$ ,  $P$  is  $n$ -by- $n$  orthogonal and  $v = Pu$ , then  $u^T u = v^T v$ .

**Proof.**  $v^T v = (Pu)^T (Pu) = (u^T P^T)(Pu) = u^T (P^T P)u = u^T u$ .

**Definition:** The **outer product** of two vectors  $x$  and  $y \in \mathbb{R}^n$  is

$$xy^T = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_n \end{bmatrix},$$

an  $n$ -by- $n$  matrix (notation:  $xy^T \in \mathbb{R}^{n \times n}$ ). More usefully, if  $z \in \mathbb{R}^n$ , then

$$(xy^T)z = xy^T z = x(y^T z) = \left( \sum_{i=1}^n y_i z_i \right) x.$$

**Definition:** For  $w \in \mathbb{R}^n$ ,  $w \neq 0$ , the **Householder** reflector  $H(w) \in \mathbb{R}^{n \times n}$  is the matrix

$$H(w) = I - \frac{2}{w^T w} w w^T.$$

**Proposition.**  $H(w)$  is a symmetric orthogonal matrix.

**Proof.**

Symmetry is straightforward to verify. For orthogonality,

$$\begin{aligned} H(w)H(w)^T &= \left( I - \frac{2}{w^T w} w w^T \right) \left( I - \frac{2}{w^T w} w w^T \right) \\ &= I - \frac{4}{w^T w} w w^T + \frac{4}{(w^T w)^2} w (w^T w) w^T \\ &= I. \end{aligned}$$

□

**Lemma.** Given  $u \in \mathbb{R}^n$ , there exists a  $w \in \mathbb{R}^n$  such that

$$H(w)u = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix} \equiv v,$$

say, where  $\alpha = \pm \sqrt{u^T u}$ .

**Remark:** Since  $H(w)$  is an orthogonal matrix for any  $w \in \mathbb{R}^n$ ,  $w \neq 0$ , it is necessary for the validity of the equality  $H(w)u = v$  that  $v^T v = u^T u$ , i.e.,  $\alpha^2 = u^T u$ ; hence our choice of  $\alpha = \pm \sqrt{u^T u}$ .

**Proof.** Take  $w = \gamma(u - v)$ , where  $\gamma \neq 0$ . Recall that  $u^T u = v^T v$ . Thus,

$$\begin{aligned} w^T w &= \gamma^2 (u - v)^T (u - v) = \gamma^2 (u^T u - 2u^T v + v^T v) \\ &= \gamma^2 (u^T u - 2u^T v + u^T u) = 2\gamma u^T (\gamma(u - v)) \\ &= 2\gamma w^T u. \end{aligned}$$

So

$$H(w)u = \left( I - \frac{2}{w^T w} w w^T \right) u = u - \frac{2w^T u}{w^T w} w = u - \frac{1}{\gamma} w = u - (u - v) = v.$$

□

Now if  $u$  is the first column of the  $n$ -by- $n$  matrix  $A$ ,

$$H(w)A = \left[ \begin{array}{c|ccc} \alpha & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right], \text{ where } \times = \text{general entry.}$$

Similarly for  $B$ , we can find  $\hat{w} \in \mathbb{R}^{n-1}$  such that

$$H(\hat{w})B = \left[ \begin{array}{c|ccc} \beta & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array} \right]$$

and then

$$\left[ \begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & H(\hat{w}) & \\ 0 & & & \end{array} \right] H(w)A = \left[ \begin{array}{cc|ccc} \alpha & \times & \times & \cdots & \times \\ 0 & \beta & \times & \cdots & \times \\ \hline 0 & 0 & & & \\ 0 & 0 & & C & \\ \vdots & \vdots & & & \\ 0 & 0 & & & \end{array} \right].$$

Note

$$\left[ \begin{array}{cc} 1 & 0 \\ 0 & H(\hat{w}) \end{array} \right] = H(w_2), \text{ where } w_2 = \left[ \begin{array}{c} 0 \\ \hat{w} \end{array} \right].$$

Thus if we continue in this manner for the  $n - 1$  steps, we obtain

$$\underbrace{H(w_{n-1}) \cdots H(w_3) H(w_2) H(w)}_{Q^T} A = \left[ \begin{array}{cccc} \alpha & \times & \cdots & \times \\ 0 & \beta & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma \end{array} \right] = (\square).$$

The matrix  $Q^T$  is orthogonal as it is the product of orthogonal (Householder) matrices, so we have constructively proved that

**Theorem.** Given any square matrix  $A$ , there exists an orthogonal matrix  $Q$  and an upper triangular matrix  $R$  such that

$$A = QR$$

- Notes:**
1. This could also be established using the Gram–Schmidt Process.
  2. If  $u$  is already of the form  $(\alpha, 0, \dots, 0)^T$ , we just take  $H = I$ .

---

3. It is not necessary that  $A$  is square: if  $A \in \mathbb{R}^{m \times n}$ , then we need the product of (a)  $m-1$  Householder matrices if  $m \leq n \implies$

$$\begin{pmatrix} \square \\ \square \end{pmatrix} = A = QR = \begin{pmatrix} \square \\ \square \end{pmatrix} \begin{pmatrix} \square \\ \square \end{pmatrix}$$

or (b)  $n$  Householder matrices if  $m > n \implies$

$$\begin{pmatrix} \square \\ \square \end{pmatrix} = A = QR = \begin{pmatrix} \square \\ \square \end{pmatrix} \begin{pmatrix} \square \\ \square \end{pmatrix}. \tag{1}$$

This  $m > n$  case is particularly important, and we note that one can also write

$$\begin{pmatrix} \square \\ \square \end{pmatrix} = A = QR = \begin{pmatrix} \square \\ \square \end{pmatrix} \begin{pmatrix} \square \\ \square \end{pmatrix}.$$

This is called the *thin* QR factorization, wherein  $Q \in \mathbb{R}^{m \times n}$  has orthonormal columns and has the same size as  $A$ ; by contrast, in (1)  $Q$  is square orthogonal, and (1) is called the *full* QR.